



UNIVERSITY OF
BIRMINGHAM

EUROSLA 3₂

European Second Language Association

Book of Abstracts



Birmingham, 30th August – 2nd September 2023

Table of Contents

Organising committee

Scientific committee

Plenary presentations

Roundtable presentations

Doctoral workshops

Oral presentations

Colloquium

Poster presentations

Award winners

Sponsors

Organising Committee

Bene Bassetti Chair (University of Modena and Reggio Emilia, and School of Education, University of Birmingham)

Akira Murakami Deputy Chair (Department of English Language and Linguistics)

Dagmar Divjak (Department of Modern Languages)

Andrea Krott (School of Psychology)

Petar Milin (Department of Modern Languages)

Roundtable organisers:

Maria Reraki (formerly School of Education, University of Birmingham, currently Edge Hill University)

Robin Thompson (School of Psychology)

Scientific Committee

Annika Andersson	Emmanuelle Labeau	Anita Thomas
Arnaud Arslangul	Monique Lambert	Elsa Tragant
Panos Athanasopoulos	Batia Laufer	Danijela Trenkic
Dalila Ayoun	Pascale Leclercq	Koen Van Gorp
Elisabet Pladevall Ballester	Mayya Levkina	Jan Vanhove
Júlia Barón	Christina Lindqvist	Josje Verhagen
Malgorzata Barras	Sarah Liszka	Georges Daniel Veronique
Bene Bassetti	Ana Llinares	Marjolijn Verspoor
Giulia Berchio	Fanny Forsberg Lundell	Karina Vidal
Raphael Berthele	Emma Marsden	Christiane von Stutterheim
Ute Bohnacker	Heather Marsden	Alexandra Vraciu
Audrey Bonvin	Maisa Martin	Stuart Webb
Teresa Cadierno	Kevin Mcmanus	Verena Wecker
Susanne Carroll	Paul Meara	Magdalena Wrembel
M. Luz Celaya	Marije Michel	Clare Wright
Saveria Colonna	Petar Milin	Stefanie Wulff
Kathy Conklin	Imma Miralpeix	Anita Thomas
Ian Cunnings	Rosamond Mitchell	Elsa Tragant
Samantha Curle	Joan C. Mora	Danijela Trenkic
Nathalie Dherbey	Maria Mos	Koen Van Gorp
Simona Pekarek Doehler	Carmen Muñoz	Jan Vanhove
Bryan Donaldson	Akira Murakami	Josje Verhagen
Amanda Edmonds	Florence Myles	Georges Daniel Veronique
Nicholas Ellis	Marianne Nikolov	Marjolijn Verspoor
Soeren Wind Eskildsen	Rakel Österberg	Karina Vidal
Anna Ewert	Gabriele Pallotti	Christiane von Stutterheim
Sara Feijoo	Magali Paquot	Alexandra Vraciu
Claudia Felser	Johanne Paradis	Stuart Webb
Maria Del Pilar Garcia Mayo	Mirosław Pawlak	Verena Wecker
Gaëtanelle Gilquin	Christos Pliatsikas	Magdalena Wrembel
Aline Godfroid	Gregory J. Poarch	Clare Wright
Gisela Granena	Philippe Prévost	
Jonas Granfeldt	Andrea Revesz	
Marianne Gullberg	Leah Roberts	
Ayse Gurel	Karen Roehr-Brackin	
Gisela Håkansson	Inès Saddour	
Julia Herschensohn	Susan Sayehli	
Holger Hopp	Andrea Schalley	
Alex Housen	Monika Schmid	
Martin Howard	Norbert Schmitt	
Rosa M Jiménez Catalán	Krister Schönström	
Elma Kerz	Rob Schoonen	
Judit Kormos	Raquel Serrano	
Andrea Krott	Rachel Shively	
Folkert Kuiken	Frida Splendido	
Tanja Kupisch	Dieter Thoma	

Plenary presentations

The crucial role of learner emotions in foreign language performance and acquisition

Jean-Marc Dewaele, Birkbeck, University of London

I will talk about an emerging area of research in the field of foreign language learning, which was triggered by the introduction of Positive Psychology (Dewaele & MacIntyre, 2014; MacIntyre & Gregersen, 2012). It has focused on the role of positive and negative foreign language learner emotions in the classroom and their effect on performance, beyond the traditional focus on foreign language anxiety. Positive emotions such as Foreign Language Enjoyment and Peace of Mind have been shown to be positively linked to Willingness to Communicate and academic achievement (Botes et al., 2022; Zhou et al., 2022) while Foreign Language Classroom Anxiety and Foreign Language Boredom have the opposite effect (Dewaele et al., 2023). Teacher behaviour in the FL class has been shown to shape learners' emotions. Establishing a positive emotional atmosphere in the classroom has not only linguistic benefits for learners but can also boost their happiness.

References

- Botes, E., Dewaele, J.-M., & Greiff, S. (2022). Taking stock: A meta-analysis of the effects of Foreign Language Enjoyment. *Studies in Second Language Learning and Teaching* 12(2), 205-232.
- Dewaele, J.-M., Botes, E., & Meftah, R. (2023). A Three-Body Problem: The effects of Foreign Language Anxiety, Enjoyment, and Boredom on Academic Achievement. *Annual Review of Applied Linguistics*
doi:10.1017/S0267190523000016
- Dewaele, J.-M., & MacIntyre, P.D. (2014). The two faces of Janus? Anxiety and enjoyment in the foreign language classroom. *Studies in Second Language Learning Teaching*, 4, 237-274.
- MacIntyre, P.D., & Gregersen, T. (2012). Emotions that facilitate language learning: The positive broadening power of the imagination. *Studies in Second Language Learning and Teaching*, 2, 193-213.
- Zhou, Li, Dewaele, J.-M., Lochtmann, K., & Xi, Yiheng (2021). Foreign language peace of mind: A positive emotion drawn from the Chinese EFL learning context. *Applied Linguistics Review*
<https://doi.org/10.1515/applirev-2021-0080>

Scaling up SLA research in the classroom: Digitally supported randomized controlled field studies

Detmar Meures, University of Tübingen

Second Language Acquisition research has uncovered a range of factors that influence acquisition, with different strands of SLA highlighting different aspects - from the Input and Noticing Hypotheses to the Output and the Interaction Hypotheses, to name just a few of the ways research has focused on cognitive, linguistic, and social factors shaping language learning.

Arguably, any ecologically valid data on authentic language learning results from a combination of these different factors - but since different levels of granularity of data at different time scales are at stake, it is difficult for research to target the interaction of the factors shaping such authentic data. In addition, interaction effects require a substantial amount of data to be reliably identified, especially in noisy authentic data. A factor making this even more challenging is the important role that individual differences play in SLA. As a result, many of the SLA hypotheses and potentially impactful research issues, such as aptitude treatment effects, remain empirically understudied and with limited impact on real-life teaching and learning.

But where could large scale longitudinal data on authentic learning processes and products come from? Over 20 million school children in Europe learn a foreign language in upper secondary schools! Yet children in the school setting are relatively underresearched in SLA, with legal and practical issues making it difficult to stage interventions and collect data at scale. Given the increasing use of digital tools in school, with a substantial boost during the Coronapandemic, a growing number of language learning contexts leave digital traces, be it when searching for and reading texts or when practicing with online activities. Adding AI methods to the mix to support the automatic analysis of language and learner modeling, it becomes possible to enhance digital practice environments so that they offer clear advantages to learners and teachers: from individually adaptive selection of motivating and input-enhanced materials to adaptive activity sequencing and automated scaffolded feedback during practice (Ruiz, Rebuschat & Meurers 2023). Complementing the practical benefits facilitating large-scale use, such digital environments can in principle be used to conduct year-long SLA experiments fully embedded in regular school contexts and provide detailed information on learners, learning processes, and products.

In this talk, I present our first steps in this direction, reporting on several randomized controlled field trials carried out in German schools using the Intelligent Tutoring System "FeedBook" to provide practice opportunities for high school students as part of the regular English classes. The first field study (Meurers et al 2019) confirmed the effectiveness of specific scaffolded feedback provided during practice and illustrates some solutions to the challenges that arise when scaling up experiments to settings that entail a substantial loss of control, including the use of learning analytics to interpret learning process data (Hui, Rudzewitz, Meurers, in press). We then report on the first results from a second year-long field study in the Interact4School project exploring motivational feedback and a learner dashboard integrating practice as pre-task activities in a task context with criterial feedback. As an outlook we then sketch two current studies: The just completed DigBinDiff study illustrating the integration of ambulatory assessment of working memory to support adaptive activity sequencing based on linguistic and cognitive factors, and the AI2Teach study that will introduce a teacher dashboard and teacher training to link the individualized practice with the teacher-orchestrated classroom. Negotiations are currently under way to potentially support all schools in the state of Baden-Württemberg. While the studies conducted so far only scratch the surface of the SLA research issues that could be investigated in such a setting, we hope it will encourage systematic collaboration on large-scale experimentation in authentic education settings, which we believe has the potential to empirically substantiate and advance complex SLA research hypotheses and highlight the relevance of SLA research for improving education.

References

- Meurers, D., De Kuthy, K., Nuxoll, F., Rudzewitz, B., & Ziai, R. (2019). Scaling up intervention studies to investigate real-life foreign language learning in school. *Annual Review of Applied Linguistics*, 39, 161-188.
- Hui, B., Rudzewitz, B., & Meurers, D. (in press). Learning processes in interactive CALL systems: Linking automatic feedback, system logs, and learning outcomes. *Language Learning & Technology*.
- Ruiz, S., Rebuschat, P., & Meurers, D. (2023). Supporting Individualized Practice through Intelligent CALL. In Suzuki, Y. (Ed.) *Practice and automatization in second language research* (119-143). Routledge. <https://doi.org/10.4324/9781003414643-7>

Investigating second language speaking and writing processes: A task-based perspective.

Andrea Revesz, University College London

In the past few decades, task-based language teaching (TBLT) has received an increasing amount of attention from instructed SLA researchers as a means to promote L2 communicative ability in L2 instructed contexts. This enhanced interest has also led to a growth of research on TBLT, with much of the extant research dedicated to investigating the characteristics of task-based speech and written production. As a result, considerable progress has been made in describing and understanding the behaviours of L2 speakers and writers when they carry out task-based work. To date, however, much less is known about the processes in which L2 learners engage when they perform oral and written tasks.

In this talk, I will argue that, to promote L2 theory-building and advance task-based pedagogical practices, it is crucial that researchers devote more attention to task-generated processes. Then, I will briefly review methods available to study task-based processes and highlight that, with a view to achieving a fuller and more complete understanding of neuro-cognitive processes during task work, it is best to triangulate various research techniques.

To demonstrate, I will describe and discuss some of my and collaborators' past and ongoing work on task-based speech and written production processes through mixed methods, including subjective (questionnaires, stimulated recalls) and objective (dual-task methodology, keystroke-logging, eye-tracking, and fMRI) tools. To begin with, I will focus on work that explored links between task complexity and task-based processes employing behavioural tools. Then, I will move onto ongoing work investigating pausing behaviours during L2 speech and writing by combining behavioural and neuroimaging techniques.

I will end the talk with summarising the main findings from this line of work drawing on my own and others' work. I will also outline some recommendations for future neuro-cognitively oriented research into L2 speaking and writing processes.

Multidimensionality and determinism of individual differences in -lingualism research: Heritage Language Bilingualism as a case in point

Jason Rothman, UiT The Arctic University of Norway

While there is great conformity in the grammars of speakers of individual languages and a set of variables that underlie how/why this comes to be, there is also a spectrum of individual differences that defines linguistic performance and competence outcomes across all types of speakers (not only bi-/multi-lingual ones, but, yes, L1-dominant users as well). Individual differences are not random. Rather, they are governed by a unique set of dynamic variables we do not quite yet fully understand. The present talk will revolve around two central points falling out from the importance of better understanding the systematicity behind individual differences: (i) the determinism of various internal and external factors—as well as their interactions—contributing to the acquisition and processing of heritage languages specifically, and thus, by extension language more generally and (ii) questioning the utility, if not appropriateness, of default aggregate comparisons as the norm, especially in bi-multilingual research where particular varieties are too liberally assumed as an appropriate comparative baseline.

While Heritage Language (HL) bilinguals acquire their HLs naturalistically in early childhood, studies over the past three decades not only typically show differences to L1-dominant homeland users, but those that examine/discuss individual level data almost inevitably document high degrees of variation at the individual level. In the present talk, I will present the landscape of HL studies that endeavors to contextualize, investigate and capitalize on the theoretical value of understanding such individual differences—presenting in such a course work that sidesteps the fallacy inherent to monolingual-to-bilingual comparisons for many—but not all—questions currently asked. In doing so, we will discuss “-lingualism” as a scale and, thus, the determinism of context (opportunities for outcomes for a particular type) and why it is (at least partially) explanatory for outcome variation at the individual level for all, but especially for HL bilinguals. Showcasing research from our group with several language pairings, across various age ranges, using a wide array of empirical methodologies and focusing on various domains of grammar—mainly studies from the large 4-year HeLPiNG (Heritage Language Proficiency in their Native Grammars) grant—, we will see that systematicity defines individual difference outcomes in HL acquisition and processing. Yet, as we might expect, there is no one-size-fits-all (set of) variable(s) that is explanatory for everything. Rather, as we should expect, the (weighting and/or interaction of) variables that regress to cover individual differences depend on what (the domain of language), who (the profiles of HL bilinguals) and where (e.g., important differences of location that delimit exposure/usage) we are investigating. This discussion sheds light on multiple levels of trending differences: between so-called monolinguals to HL bilinguals, between distinct populations of HL bilinguals, e.g., European to North American to Asian contexts, and intra-group differences across individual HL bilinguals traditionally studied as a single aggregate.

Roundtable

The role of first language, working memory and reading anxiety in second language reading: Implications for teaching and assessment for language learners with specific learning differences

Judit Kormos, Lancaster University

There is ample research evidence to suggest that first language literacy skills and second language reading comprehension are strongly inter-related. Lower first language literacy skills and smaller capacity of working memory capacity are also known to be underlying causes of specific learning differences that have been shown to impact on second language learning outcomes. However, most research in this area has focussed on younger children or higher level university students. In this presentation I describe a study that has investigated how first language reading comprehension, phonological awareness and naming speed together with working memory capacity and reading anxiety predict second language reading performance among Hungary secondary school students. In this project I also examined how these first-language literacy-related and cognitive variables are associated with the time needed for participants to complete the reading comprehension test. The findings have implications for research on time adjustments for students with specific learning differences and for supporting students who might experience challenges in second language reading comprehension.

Accommodating learners with specific learning difficulties in L2 assessment: a straightforward matter?

Benjamin Kremmel, University of Innsbruck

Guidelines for good practice in assessment stipulate the importance of fair and equitable treatment of all candidates in the testing process.

Given the increased number of test takers with specific learning difficulties (SpLDs), both providers and researchers of language tests have started to give accessibility arrangements their long overdue attention over the last 15 years. While many test providers aim at providing appropriate testing conditions these days, and even Ministries of Education are finally issuing recommendations on the need for accommodating learners with SpLDs, there is still a dearth of empirical evidence when it comes to the actual implementations of such accommodations and their effectiveness.

In this talk, I will briefly introduce SpLDs and considerations of fairness in language assessment before discussing previous studies that have looked into assessment accommodations. In particular, I will draw on a study in the context of L2 listening assessment for young learners in Austria to illustrate that implementing accommodations might not always result in improvements in performance or support test-takers with SpLDs efficiently. The study investigated the potential of self-paced listening as an adjustment for learners with low L1 literacy skills, which are relevant proxies for learning difficulties in contexts where systems of official SpLD diagnoses are lacking. The research examined the impact of self-paced listening on comprehension scores, listening anxiety and strategy use. The presentation concludes with a call for increased research efforts to ensure that well-intended measures are evidence based and efficient in delivering fairness and equity in assessment.

Second language reading in people with literacy difficulties: Some lessons learned

Jackie Masterson, UCL Institute of Education

People with reading and/or spelling difficulties can experience poor academic and socio-emotional outcomes, and life chances can be compromised. We will discuss the difficulties of identification of literacy difficulties in the face of sparse materials for some languages. Potential underlying reasons for reading and spelling difficulties are discussed, together with their interaction with writing systems. We will look at evidence for different manifestations of reading and spelling difficulties across a person's languages, as well as some intervention studies that have resulted in effective outcomes. Areas for further development are identified.

Adults as facilitators of learning for children with reading difficulties in diverse language learning contexts.

Anna Tsakalaki. University of Reading

Abstract: Educators and family members regularly play a major role in supporting the literacy development of children with reading difficulties in school or at home. Schools that host learners with language backgrounds (L1) different from the language of instruction (L2) around the world increasingly recognise the linguistic diversity of their schooling population as a means of creating an inclusive ethos in their setting. Parents in multilingual families steadily realise the benefits of raising multilingual children and encouraging the use of L1 at home for social and learning purposes. However, levels of confidence in supporting literacy development of children falling in both categories vary among the adults tasked to facilitate their learning, especially when it comes to specific elements of language learning when specialised knowledge is required.

In this talk, I will examine the role of educators and adult family members in facilitating specific elements of language learning, namely reading, spelling and vocabulary for learners with reading difficulties coming from different language backgrounds. Selected data will be drawn from four recent studies done in Greece, the UK and a low-resource Sub-Saharan country with primary school-aged children and adults supporting them. Three key factors empowering adults to support children's development of literacy-related skills in school and at home will be discussed: a) confidence in identifying and supporting learners' specific literacy needs, b) understanding of the interactions between L1 and L2, and c) making room for differentiation in learning. Implications of these findings for further research in the role of adults in supporting multilingual children with reading difficulties, educational practice and policy will be discussed.

Doctoral workshops

The Role of Executive Functions and Task Features in Multiple Text Reading

Hatice Akgün (Boğaziçi University) and Gülcan Erçetin (Boğaziçi University).

Although the importance of multiple text reading (MTR) for academic study among university-level students has been underlined by a number of researchers (Goldman, 2011, Schoor et al., 2020), very few studies investigate the effect of task features on the integration processes and reading comprehension strategy use (Primor and Katzir, 2018) and the role of executive functions (Tarchi et al., 2021) involved in MTR. For this reason, the aim of the present study is twofold. First, it will investigate the effect of item types (i.e. literal vs inferential questions), rhetorical relations among the texts (i.e. conflicting vs complementary) and the explicitness/implicitness of these relations on the integration processes based on the integrated framework of multiple text use developed by List and Alexander (2018) and strategy use in L2 MTR. Secondly, the study will examine the role of executive functions, namely inhibition and working memory (WM) in L2 MTR research. The data will be collected from 30 undergraduate students in an English medium university through eye-tracking technology and retrospective verbal protocols. To measure L2 MTR, one set of reading task will be created and this set will include three different texts which have overlapping or conflicting arguments on the same topic. Types of textual rhetorical relations will consist of conflicting, complementary, and overlapping relations across the texts and the explicitness/implicitness of these relations will be manipulated by inserting and removing connectors and/or signaling phrases. Tasks will consist of questions that measure inferential and literal reading. Eye movement data and verbal protocol data will be used to elucidate the reading processes based on the task features. To measure inhibition, antisaccade and flanker tests will be implemented and to measure WM span, symmetry span and rotation span tasks will be used. The scores from these tests will be used to find out if there is a relationship between MTR performance and the cognitive individual differences. Investigating the effect of task features and the role of EFs in relation to MTR is significant in terms of MTR instruction and assessment, improving the existing theoretical models in MTR, and understanding the cognitive individual differences in MTR. References Goldman, S. R. (2011). Choosing and using multiple information sources: Some new findings and emergent issues. *Learning and Instruction*, 21(2), 238–242. List, A., & Alexander, P. A. (2018). Toward an Integrated Framework of Multiple Text Use. *Educational Psychologist*, 54(1), 20-39. Primor, L., & Katzir, T. (2018). Measuring Multiple Text Integration: A Review. *Conceptual Analysis*, 9(2294), 1-16. Schoor, C., Hahnel, C., Mahlow, N., Klagges, J., Kroehne, U., Goldhammer, F., & Artelt, C. (2020). Multiple Comprehension of University Students. In O. Zlatkin-Troitschanskaia, H. A. Pant, M. Toepper, & C. Lautenbach (Eds.), *Student learning in German higher education*. (pp. 221-240). Springer VS. Wiesbaden. Tarchi, C., Ruffini, C., & Pecini, C. (2021). The Contribution of Executive Functions When Reading Multiple Texts: A Systematic Literature Review. *Frontiers in Psychology*, 12.

Fundamentals of word learning: The interaction of phonological knowledge, perception and production

Sophie Bennett (Lancaster University), Susana Correia (NOVA University Lisbon, Linguistics Research Centre), Padraic Monaghan (Lancaster University) and Patrick Rebuschat (Lancaster University).

Sophie Bennett(1), Susana Correia(2), Padraic Monaghan (1) & Patrick Rebuschat(1, 3) (1) Lancaster University, UK (2) NOVA University Lisbon - Linguistics Research Centre, Portugal (3) University of Tübingen, Germany

Two key factors drive successful second language (L2) word learning. The first is phonological knowledge of the target language for perception and production skills (e.g., Flege & Bohn, 2021). Non-native speakers may not have the phonological knowledge to learn words containing unfamiliar, non-native sounds (e.g., Flege & Bohn, 2021). The second is the practice modality of these developing skills. According to the skill specificity theory, performance of a given skill is determined by practice of that skill in question (e.g., Li and DeKeyser, 2017) yet few studies have researched the influence of production on L2 learning. The present study investigates the interaction of both phonological knowledge and practice modality in the acquisition of European Portuguese novel words.

This preregistered study emulates the natural conditions of implicit word learning through a cross-situational word-learning paradigm (e.g., Tuninetti, Mulak, & Escudero, 2020). In the coming months, 80 Portuguese and 80 English native speakers aged 18-30 will learn the form and meaning of 8 auditorily-presented Portuguese pseudowords through this paradigm online. Their accuracy in comprehending the word-meaning association will be measured on a trial-by-trial basis. Half of the participants will also rehearse the words for every trial through vocal repetition (the “listen+repeat” group). Their accuracy in producing the form will be measured alongside their comprehension score. The effect of comprehension and production practice on word learning, versus that of comprehension practice alone will be compared. To measure the effect of phonological knowledge, half of the pseudowords contain a nasal vowel (e.g., [lõsu]) which are “familiar” to the Portuguese natives but “unfamiliar” to the English natives since they will have no experience of Portuguese. The other half contain familiar sounds to both groups (e.g., [davu]). The participants’ production and comprehension scores for familiar versus unfamiliar words can thus be compared. The robustness of this word learning will be evaluated through a comprehension-based test (word recognition) and a production-based test (picture naming).

Practice modality is predicted to drive the participants’ production and comprehension accuracy at test; the listen-only group will perform better than the listen+repeat group in the comprehension test and vice versa in the production test (e.g., Li and DeKeyser, 2017). The participants’ phonological knowledge will also affect their production and comprehension scores where the Portuguese group will outperform the English group in the unfamiliar condition, but not in the familiar condition (e.g., Flege & Bohn, 2021). The project has important theoretical implications for L2 acquisition, as well as practical insight into how effective word-learning activities are implemented remotely.

References Flege, J., & Bohn, O. (2021). The Revised Speech Learning Model (SLM-r). In R. Wayland (Ed.), *Second Language Speech Learning: Theoretical and Empirical Progress* (pp. 3-83). Cambridge: Cambridge University Press.

Li, M., & DeKeyser, R. (2017). Perception practice, production practice, and musical ability in L2 Mandarin tone-word learning. *Studies in Second Language Acquisition*, 39(4), 593–620.

Tuninetti, A., Mulak, K. E., & Escudero, P. (2020). Cross-Situational Word Learning in Two Foreign Languages: Effects of Native Language and Perceptual Difficulty. *Frontiers in Communication*, 5(109), 1–12.

Questions for the audience 1. How can accuracy scores of specific L2 linguistic features (e.g., the production accuracy of a given European Portuguese nasal vowel) be predicted when those features have never been studied in L2 acquisition before in any language?

2. The participants will hear the auditory stimuli and record their speech via the online research platform, Gorilla. How can researchers try to control for differences in the quality of the participants’ own audio and recording equipment online in a between-subjects design?

Contact: s.c.bennett@lancaster.ac.uk

Language Attrition and Ultimate Attainment of Adjectives in Near-native Spanish Grammars

Andrea Calpe Álvarez (Universitat de les Illes Balears).

This study examines whether near-native L2 Spanish adult speakers with more than 10 years of exposure to Spanish in an immersion context can achieve ultimate attainment of properties related to adjectival order within Spanish DP, and whether there is evidence of attrition effects in their first language. We test their knowledge of uninterpretable and interpretable features related to overt noun raising in the Spanish DP and the possible readings for qualitative adjectives regarding their position with respect to the head noun. In doing so, we evaluate the predictions of the Interpretability accounts (Tsimpili & Dimitrakopoulou, 2007 a.o.) against Full Accessibility approaches (e.g., White 2003) including the Feature Reassembly Hypothesis (Lardiere, 2009 a.o.) as well as the possible influence of the L2's grammar in L1's structures. (Putnam et al., 2019 a.o.). Germanic and Romance languages exhibit syntactical and semantic differences of adjective-noun order: English and German lack overt noun movement, and are thus parametrically different from Spanish and Portuguese. Therefore, Germanic qualitative adjectives appear in prenominal position with an ambiguous reading between restrictive and non-restrictive meanings, whereas in Romance languages, where both positions are possible, each one corresponds to one reading (Cinque, 2010; Demonte, 2008). We compare three experimental groups (L1 English (n=50); L1 German (n=50) and L1 Portuguese (n=50)), previously assessed by independent scrutinized nativelikeness test measures based on Abrahamsson & Hyltenstam's (2009) against a native Spanish control group (n=100) in a combination of offline and online tasks. The offline tasks consist on a semantic interpretation task to choose a paraphrase of the meaning based on the adjectival position and a collocation task to place the adjective pre or postnominally based on a provided context. For the online tasks, word-by-word non-cumulative self-paced reading and production tasks testing grammaticality and adjective position will be used. Based on current minimalist theory, we hypothesize that Germanic speakers lack the feature-checking that triggers overt noun movement and they must acquire it, while Portuguese speakers have similar morpholexical assemblages and all the relevant features are available in their L1. Hence, Portuguese near-native speakers of Spanish would have an advantage in acquiring obligatory noun raising in Spanish and are expected to perform more similarly to natives than their English and German counterparts if transfer obtains. As for attrition effects, we would expect that, if the features involved in overt noun movement are interpretable, they will appear more vulnerable to attrition, so Portuguese speakers in the experimental group may show divergences in their responses compared to those from their monolingual counterparts.

Selected references

- Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language. *Language learning*, 59(2), 249-306.
- Cinque, G. (2010). *The syntax of adjectives: A comparative study*. MIT Press
- Demonte, V. (2008). Meaning-form correlations and the order of adjectives in Spanish in L. McNally, & C. Kennedy (Eds.), *Adjectives and Adverbs: Syntax, semantics, and discourse* (pp.71-100). OUP.
- Putnam, M., Perez-Cortes, S., & Sánchez, L. (2019). Language attrition and the Feature Reassembly Hypothesis in. M.S. Schmid & B. Köpke (Eds.), *The Oxford Handbook of Language Attrition* (pp. 18-24). OUP.

Multilingualism in the Workplace: An Emic Exploration of Workers' Experiences and Perceptions of Language Learning at FishCom

Ina Celise Sortland (Norwegian School of Economics).

The topic of this PhD project is language practices and language learning in the multilingual workplace. The project explores the case of FishCom (pseudonym), one of Norway's leading suppliers of farmed salmon, with over 500 employees. The topic of the present workshop is related to the first paper in the dissertation addressing the following research question: How do L2-users of Norwegian perceive the conditions for learning languages at the workplace and how do these experiences align/misalign with the company's language learning intentions and policies?

Prior research has explored the language practices and lived experiences of workers in multilingual workplaces (e.g., Angouri, 2013), as well as in situ language learning that occurs in these environments (e.g., Creese & Blackledge, 2019; Svennevig, 2018). Despite this, there remains a gap regarding workers' perception of language learning in the workplace. This is a significant issue to address, as workplace second language education is a growing trend in Scandinavia, driven by the belief that interaction with L1-speakers of the target language promotes language acquisition. Previous studies have provided insights into language learners with specific immigrant statuses who engage in formal workplace language education (Kirilova & Lønsmann, 2020, p. 41). The present project, however, shed light on a diverse and understudied group of working immigrants with varying levels of formal education in Norwegian, who already has gained access to the labor market without taking part in workplace language education.

The study employs qualitative methods, including interviews with both L1- and L2-user of Norwegian, leaders and other workers with key functions. For the PhD project, I have also conducted ethnographic observations and collected video recordings of naturally occurring interaction. A thematic analysis of the interview data will be conducted (Pavlenko, 2007) with a sensitivity to historical, political, and social circumstances. The interview data will hold a prominent position in the first paper as they will provide valuable insight into the workers' perceptions and experiences. In the workshop, I will present interview data and solicit feedback on the analytic approach.

The project intends to contribute to the growing body of literature on the role of language learning in real-life context. By taking an emic perspective, the study will shed light on workers' experiences, and provide a nuanced understanding of the impact of multilingualism on language practices and learning.

References: Angouri, Jo (2013) The multilingual reality of the multinational workplace: language policy and language use. *Journal of Multilingual and Multicultural Development*, 34:6, 564- 581,

Creese, A. & Blackledge, A. (2019) Translanguaging and Public Service Encounters: Language Learning in the Library. *The Modern Language Journal*, 103:4, 800-814

Kirilova, M. & Lønsmann, D. (2020) Danish – the key to the labour market? Ideologies about language use and language learning in two work contexts in Denmark. *Nordand*, 15:1, 37-57

Pavlenko, A. (2007) Autobiographic Narratives as Data in Applied Linguistics. *Applied Linguistics*, 28:2, 163-188

Svennevig, J. (2018). "What's it called in Norwegian?" Acquiring L2 vocabulary items in the workplace. *Journal of Pragmatics* 126, 68-77.

What can MINERVA2 tell us about killing hope? Investigating L2 Collocational Processing as Memory Retrieval

Sydelle de Souza (School of Informatics, University of Edinburgh), Ivan Vegner (School of Informatics, University of Edinburgh), Leonidas Doulas (School of Philosophy, Psychology, and Language Sciences, University of Edinburgh) and Francis Mollica (School of Informatics, University of Edinburgh).

Collocations are semi-productive word combinations wherein one word is used literally and one other figuratively. They are characterized by an arbitrary restriction on substitution (e.g., kill hope, #murder hope) and are notoriously difficult for L2 speakers to acquire and use. This difficulty is clearly reflected in processing and can be traced to complex interactions between frequency (exposure to L2 input), congruency (availability in the L1), and semantic transparency (degree to which meaning can be inferred from the constituent words). Research shows that collocations incur a processing cost over productive combinations (e.g., eat cake). To put this into perspective, fully opaque and non-productive idioms (e.g., break the ice) enjoy a processing advantage over productive combinations. Evidently, collocations not only possess idiosyncratic meanings, but also exhibit idiosyncratic processing trends, presenting a curious case for psycholinguistics, and language acquisition at large. However, despite increasing interest among researchers in Applied Psycholinguistics, little is known about the underlying mechanisms of collocational processing that can account for this idiosyncrasy. The overarching aim of my thesis is to find a suitable explanatory model for L2 collocational processing. Based on evidence for analogical processes in L2 idiom and metaphor comprehension, I explore analogy as a plausible mechanism for collocations. Therefore, as an initial foray, this study attempts to explain trends in L2 collocational processing from the standpoint of memory retrieval—the first step in analogical processing. We parametrically explore the assumptions under which retrieval would be sufficient to explain the observed patterns of L2 learners' collocational processing. We use MINERVA2, a frequency-based, global-matching memory model, to simulate reaction times and compare them to data from 99 L1 and 230 L2 (L1 Portuguese) English speakers involving free combinations (e.g., eat cake ↔ 'comer bolos'), congruent (e.g., read minds ↔ 'ler mentes'), incongruent collocations (e.g., kick habits, no equivalent translation in Portuguese), and nonsense baselines (e.g., read cakes). We make assumptions about: (i) how similar items are encoded in memory, and (ii) their sensitivity to frequency. Using DistilBERT models trained on English and Portuguese corpora, we extract two vector representations for each item capturing the semantics of the word combination in English, and its literal Portuguese translation. This allows us to model congruency and semantic transparency. We populate the MINERVA2 memory based on phrasal frequency in English, Portuguese, and a weighted (60-40) mix of Portuguese and English frequencies to model an intermediate-advanced L2 English learner. We simulate memory retrieval patterns from both languages under the various frequency conditions. Under the assumptions that the L2 lexicon develops with respect to the L1 and the L2 lexicon is sensitive to L1 frequencies, we find that MINERVA2 can neatly account for processing trends in both L1 and L2 collocational processing (see figures 1 and 2). Based on this, I would like to brainstorm with the audience about manipulating frequency conditions to simulate levels of L2 proficiency and model its effects on collocational processing.

The Processing and Production of Pronouns in English as a Second Language

Linghui Diao (University of York).

Background and Purpose: Pronoun comprehension and production can be challenging for second language (L2) speakers, especially in contexts where the reference of the pronoun is ambiguous, for instance, where a pronoun can grammatically refer to either of two potential antecedents in the preceding discourse (e.g., Roberts et al., 2008). Most L2 acquisition studies of pronoun resolution have focused on speakers of a non-pro-drop language, like English, with obligatory overt pronouns, who are learning a pro-drop system, with null pronouns (e.g., Italian, Spanish, Greek). There have been fewer studies investigating pronoun resolution/use with L2 speakers from a pro-drop first language (L1) background acquiring the assumed less complex non-pro-drop system. Not only are the results of these studies mixed, but different tasks have led to different results, even with the same group(s) of speakers. The present study focuses on English as the target language. The L2 learners in this study have Mandarin Chinese as their native language, which, in contrast to English, is a pro-drop language that permits null pronouns. The purpose of the study is two-folded: 1. To investigate English subject pronoun resolution in comprehension and production by L2 speakers with Mandarin Chinese as their L1. 2. To explore the similarities and discrepancies in performance between the processing and production of English pronouns by L2 speakers from a Mandarin Chinese L1 background.

Methodology: The study will employ three methods: the visual world eye-tracking paradigm, an elicited production task and an acceptability judgement task. Two participant groups will be involved in this study: 1) forty-four L2 speakers of English with Mandarin Chinese as their L1, who have different levels of L2 language proficiency; 2) forty-four native speakers of English who will act as a control group. The study will employ 24 sets of experimental sentences, structured similarly to the materials used in the study conducted by Roberts et al. (2008). It will manipulate discourse complexity and discourse topic.

Expected Outcomes: The research predicts that when there are two possible referents in the discourse, L2 English speakers (compared to L1 speakers) will take more time to decide on the resolution of pronouns. For example, in the sentence "The doctor[female] and the nurse[female] are at the clinic. While the doctor is working with the children, she is singing a tune." This increased difficulty is expected, if the complexity of having multiple potential referents is a problem for L2ers (as per Cunnings et al., 2017). Additionally, when there is only one potential referent in the discourse, if L1 influence affects L2 English speakers, a topic shift condition like "The doctor[male] and the nurse[female] are at the clinic. While the doctor is working with the children, she is singing a tune." may actually be easier for L2ers to process than for L1ers, according to Roberts et al. (2008). Furthermore, if L2ers' production abilities lag behind their comprehension abilities, they will not show consistent patterns in choosing referents in real-time comprehension and non-real-time production tasks, as native speakers typically do.

Conclusion: Overall, this research will contribute to our understanding of L2 speakers' acquisition mechanisms and their learning process in pronoun usage, which can have implications for English language teaching and learning methods.

Code-Switching by Venetian-Spanish Bilingual in a Family Interaction: Longitudinal Case Study

Karina Fascinetto-Zago (Benemérita Universidad Autónoma de Puebla, Mexico / Leiden University, The Netherlands).

Code-Switching (CS) is an activity that may be observed in the speech -or writing- of bilinguals who go back and forth between their two languages in the same conversation (Deuchar 2013). The asymmetry and systematicity between two grammatical systems in contact were captured in the notions of Matrix Language (ML) and Embedded Language (EL). These initial notions were taken up in the Matrix Language Frame (MLF) model proposed by Myers-Scotton (2002). The MLF has been applied to analyze bilingual speech in adults and with some analytical modifications in children (e.g. Lanza 1997; Vihman 1998; Paradis, Nicoladis, and Genesee 2000) in different language pairs. This work focuses on the ML of intra-clause CS in children and adults in bilingual data Venetian (VEC)/Spanish (SPA). The VEC language is a minority language in Puebla, Mexico, since 1882, due to the arrival of Italian immigrants, most from the Veneto region, Northern Italy. The purpose of this work is the identification of the ML in the monolingual and bilingual clauses in a family interaction based on the MLF model. The ML of the bilingual clauses can be VEC or SPA. The ML of VEC monolingual clauses is VEC, while the ML of SPA monolingual clauses is SPA. We take the clause as the analysis unit and the ML as a dependent variable to achieve two goals: (1) to describe the relative role of the VEC/SPA languages in the bilingual speech of adults and children, and (2) to compare the ML in the bilingual and monolingual clauses of children and adults. Until now, we analyzed nine recordings with 8:48:23 hours of natural speech conversation in play situations in family interaction. The age range of the two younger children is 3;02,07-3;09,00, and of the older sister is 6;04,13-7;06,01. The MLF model has not been used to analyze the VEC language in contact with another language. Moreover, this is the first work where it is analyzed the VEC/SPA intra-clause CS in both the bilingual speech of adults and children. Preliminary results show that the ML of most of the bilingual clauses is VEC in both adults and children and that in the latter, there have been several cases where it has not been possible to determine the ML of the clause from the MLF model.

References: Deuchar, Margaret. 2013. "Code-Switching." In *The Encyclopedia of Applied Linguistics*, 10 Volume Set, edited by Carol A. Chapelle, 657–64. New York, NY, USA: Wiley. Lanza, Elizabeth. 1997. *Language Mixing in Infant Bilingualism: A Sociolinguistic Perspective*. Oxford: Clarendon Press. Myers-Scotton, Carol M. 2002. *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes*. Oxford: Oxford University Press. Paradis, Johanne, Elena Nicoladis, and Fred Genesee. 2000. "Early Emergence of Structural Constraints on Code-Mixing: Evidence from French–English Bilingual Children." *Bilingualism: Language and Cognition* 3 (3): 245–61. <https://doi.org/10.1017/s1366728900000365>. Vihman, Marilyn May. 1998. "A Developmental Perspective on Code-Switching Conversations between a Pair of Bilingual Siblings." *International Journal of Bilingualism* 2 (1): 45–84.

Metalinguistic reflection through linguistic contrast in the EFL classroom

Neus Frigolé Pujol (Universitat de Barcelona).

Languages are in constant interaction (Cook, 2001), and each new language modifies our general linguistic competence (Piccardo, 2013). Consequently, there is a need to move from a monolingual to a plurilingual approach to language teaching (Cenoz & Gorter, 2019). This need for fostering plurilingual and intercultural education is more evident facing the diverse linguistic repertoires in the language classroom. In the light of the above, this interventional study aims to expand participants' perspectives on teaching EFL plurilingually through the confrontation of their current beliefs with scientific evidence.

The intervention is divided into two stages. In the first, 42 English teachers from different schools in Catalonia have answered a questionnaire on their beliefs about plurilingualism, mediation and metalinguistic reflection. In the second, five teachers selected from Stage 1 will engage in a reflexive training process throughout an academic year. This reflexive training will be divided into three phases. In the first phase —self-analysis—, participants' Orienting Basis of Action (OBA; Gal'perin, 1992) will be analysed by means of a visual reflexive diary (VRD). This VRD will combine visual narratives (Kalaja et al., 2013) and narrative inquiry (Johnson & Golombek, 2002). In the second and third phases —contrast and redescription, respectively—, four sessions of a focus group will be conducted to compare participants' visions. In the third phase too, the participants will reflect upon activities aimed at fostering metalinguistic reflection. They will also examine their own classroom practices to see if their initial OBA has conceptually expanded (Engeström, 2011) as a result of the treatment.

This PhD proposes a new procedural and conceptual framework to promote a plurilingual and intercultural approach in the EFL classroom. However, we seek guidance on (1) whether the five participants in Stage 2 should answer the questionnaire again at the end of the intervention to see if their beliefs have changed because of the training process, and (2) the best way to analyse the VRD.

Interpretation and processing of relative clause attachment: the influence of L2 instructed exposure in bilinguals' L1

Elena García-Guerrero (University of Granada).

A recurrent concern in second language acquisition (SLA) has been the extent to which bilinguals' first language (L1) may influence their second language (L2). This unidirectional effect has been challenged by studies that demonstrate that the L1 may also experience readjustments in processing and comprehension, i.e., attrition, due to high L2 exposure and use (Chamorro et al., 2016; Dussias & Sagarra, 2007; Tsimpli et al., 2004). These investigations have analysed long-immersed bilinguals in naturalistic settings, but there is scarce data on the potential influence of other L2 exposure types like instructed settings. This dissertation will fill this gap by investigating an overlooked population: bilinguals in an L1 environment with high L2 formal exposure.

The effect of instructed exposure on bilinguals' L1 will be studied regarding a specific type of structural ambiguity, i.e., relative clause attachment (RCA). In (1), the relative pronoun who can refer to either the daughter, i.e., high attachment (HA) or the psychologist, i.e., low attachment (LA).

(1) Peter fell in love with the daughter of the psychologist who studied in California

These mechanisms appear to be language-specific: while Spanish favours HA (Dussias, 2003; Jegerski et al., 2016), English shows a tendency to LA (Cuetos & Mitchell, 1988; Fernández, 2002). Thus, the aim is to investigate: (1) the RCA disambiguation preferences of L1 Spanish-L2 English bilinguals in their L1, (2) whether this is modulated by their extensive L2 instructed exposure and (3) whether attrition may be found in processing only or also in offline comprehension.

To do so, three groups will be studied: (1) advanced L1 Spanish-L2 English bilingual students of English Studies in Spain, as well as (2) Spanish and (3) English monolinguals as control groups. Data will be collected via a picture selection task to address offline comprehension preferences, together with an auditory sentence-picture verification task and a visual-world eye-tracking experiment to test their processing/online preferences. All experiments have been designed following the visual-world paradigm (VWP) and using the same visual and auditory stimuli to ensure reliable comparison. Regarding the target sentences, the relative pronoun gender was manipulated as in (2) to create three conditions, i.e., HA-bias, LA-bias and ambiguous.

(2) Mira al fotógrafo(i) de la actriz(j) el cual(i) / la cual(j) / que(i/j) sujeta un globo felizmente

Look at the photographer (masc) of the actress (fem) who holds a balloon happily

An effect of L2 instructed exposure is expected on bilinguals' comprehension and processing of RCA in Spanish represented by higher acceptance of their L1 dispreferred condition, i.e., LA-bias condition, in all experiments. The use of the VWP, a methodology not previously employed to test RCA preferences, to investigate a population overlooked in previous research will make a relevant contribution to the understanding of L1 attrition in different contexts.

The presentation welcomes a discussion on the influence of language experience in bilinguals' L1 and particularly, on the extent to which the impact of L2 instructed exposure may be similar to that of naturalistic immersion. Feedback regarding the experimental design and methodology will also be appreciated.

English as an LX- Multilingualism, transfer and interactions with background profile

Bjørn H Handeland (University of Agder).

The aim of the study is to investigate second language English learning and effects of home language in young learners in Norway. Previous research on young learners and multilingualism in Norway has tended to focus either on learners' attitudes, experiences or language practices, and little research has considered how such background factors influence actual linguistic performance.

Multilingualism is generally associated with potential benefits including, but not limited to language learning, often related to higher use of metacognition and learning strategies. However, recent Norwegian studies point to this not always being the case in young learners with immigrant or 2nd generation immigrant backgrounds. Additionally, achieved marks after completed lower secondary education in Norway have shown the largest gap between majority and minority language speakers in the entire OECD area. This project combines extensive background data with several measures of proficiency in English morphosyntax in order to consider how linguistic background influences performance between groups. Two data collection tools were created for the project. Firstly, a questionnaire was developed to collect data on known languages and proficiency across all domains, language use in specific situations, general attitudes to language, language learning and communication. As all participants were at least bilingual, the questionnaire was designed to provide insights into participants' ratings of their abilities in all languages they knew, but also to gain understanding of the degree to which they were actively multilingual in daily situations and which language(s) were used in the home. Secondly, a three-part test comprising sentence completion tasks, free production tasks and error spotting was constructed. The test focussed on key aspects of English morphosyntax; subject-verb concord, aspect, word order, definiteness marking and use of prepositions. These variables were chosen after considering grammatical overlap between English and the four languages with the largest numbers of speakers in Norway, Norwegian, Arabic, Polish and Somali with the objective of detecting distinct patterns of transfer errors between different groups of learners. Data from 580 pupils in lower secondary school was collected through anonymous digital data-collection using SurveyXact. 303 female and 277 male pupils participated; average age was 12 years. 383 were speakers of Norwegian as the only home language and 197 were speakers of a wide variety of other home languages.

The study's analytical approach is to consider the extent to which background factors influence error patterns, and how salient background factors differ between the two participant groups. Data analysis is ongoing in a process of several stages where first the two questionnaire data sets have been factor analysed. This first analysis has considered all general background factors in addition to those questions relating specifically to Norwegian and English, not including the number of other languages known to the participants. Factor analysis and multiple regressions against error spotting data will be presented at the workshop.

Audio-visual input: A pathway to better perception and production?

Laura Hund (University Konstanz).

To date, only few studies have examined the potential benefits of audio-visual input for developing learners' speech perception and production. While Hutchinson & Dmitrieva (2022) found positive effects on the segmental level for production, Wisniewska & Mora (2020) have used global accent rating, showing that videos with and without caption can lead to improved pronunciation. However, the latter method, has two disadvantages: it is rather subjective and it remains unclear which features have contributed to the raters' perception of accentedness. Given Mora & Cerviño-Povedano (2019) findings that captions may enhance the processing of subtitled text depending on the rhythmic structure of the target language, a focus on specific properties, in particular at the suprasegmental level, seem pertinent to advance this line of research. This study investigates whether audio-visual input can enhance phonological skills by looking at specific phonological properties of L2 French learners with German as their L1. Two properties differing between German and French have been selected, one segmental, nasal vowels, and one suprasegmental, rhythm. I will address the following questions: (1) Do learners' L2 perception and production of nasal vowels and rhythm benefit from exposure to audio-visual input in the L2? (2) What is the effect of the viewing mode (captions vs no captions)? The hypothesis is that the caption group will outperform the other groups due to the facilitation of perceptual learning. Specifically, the use of captions may enable updating of phonological representations to more accurately correspond with the auditory input associated with the written form. A pretest-posttest design will be used. Adult German participants, who are advanced learners of French, will be exposed to a total of six episodes (approx. 5 hours) of French TV series. The 90 participants will be randomly assigned to one of the three conditions: (a) audio-visual input and captions; (b) only audio-visual input; and (c) no input (control group). All participants will be tested twice, in the pre-and posttest, on the two phenomena. To assess changes in production participants will do a delayed sentence production task. The speech samples will be acoustically analysed and compared to samples of a monolingual French control group (30 participants). An identification task will be used to assess possible benefits in perception. Furthermore, individual differences that might have an influence on the learning gains, like L2 proficiency and reading behaviour (Wisniewska & Mora, 2020), will be assessed with an Elicited Imitation Task and an eye tracking experiment to calculate a Reading Index of Dynamic Text.

References Hutchinson, A.E & Dmitrieva, O. (2022). Exposure to speech via foreign film and its effects on non-native vowel production and perception. *Journal of Phonetics*, 95, 101242. Mora, J., & Cerviño-Povedano, E. (2019). The effects of bimodal L2 input on the processing of function words by Spanish EFL learners: An eye-tracking study. In C. Herrero & I. Vanderschelden (Eds.), *Using film and media in the language classroom*, 76–91, *Multilingual Matters*. Wisniewska, N., & Mora, J. C. (2020). Can captioned video benefit second language pronunciation? *Studies in Second Language Acquisition*, 42(3), 1–26.

Questions for the discussants 1) Are there additional variables, that I should control for, besides L2 proficiency and reading behaviour? 2) How well do the chosen research methods match the research questions? 3) Is there a more suitable method to elicit speech samples, as compared to the delayed sentence production task?

The development of narrative structure in the oral narratives of L1 Russian/L2 Spanish children

Maria Jose Centelles Cuart (University of Barcelona), Elisa Rosado Villegas (University of Barcelona) and Iban Mañas Navarrete (University of the Balearic Islands).

Child narrative development largely depends on their linguistic and cognitive abilities (Berman, 1999; Strömquist & Day, 1993) which allow them to move from the "here-and-now" discourse, and handle "there-and-then" events (Hickmann, 2003). Research on first languages (L1) narrative development shows that nine-year-old children's narratives already include all the components of narrative macrostructure; however, the introduction of evaluative material (typical of adult narratives) continues to develop until late adolescence (Berman & Nir-Sagiv, 2007; Tolchinsky, 2014). These findings have motivated more research on narrative development by bilingual children (Gagarina, 2016; Verhoeven & Strömquist, 2001) and in second language (L2) learning (Karlsen et al. 2015; Pavlenko, 2006), showing that mastery of narrative structure is a very challenging process for non-native speakers, which requires specific didactic intervention. In this study we examine the development of narrative competence by L1 Russian children who learn Spanish as an L2. Participants (N=105) were divided into two age groups (10-11 and 12-13 years), and were asked to produce two semispontaneous oral narratives, in both languages. A total of 170 texts were assessed for: 1) narrative structure, 2) narrative complexity, and 3) inner states (Petersen et al., 2008). The main questions addressed in this study are: (i) To what extent does children's performance differ in terms of narrative macrostructure in both groups? (ii) To what extent do these same features differ in L1/Russian versus L2 Spanish? (iii) What is the developmental pattern of narrative complexity, i.e., more number of events; and references to inner states and evaluative elements? (iv) To what extent does this pattern differ in L1/Russian versus L2 Spanish?

References:

- Berman, R. (1999). Bilingual proficiency/proficient bilingualism: Insights from narrative texts. In G. Extra & L. Verhoeven (Eds.), *Bilingualism and migration* (pp. 187-208). Mouton De Gruyter.
- Berman, R., & Nir-Sagiv, B. (2007). Comparing narrative and expository text construction across adolescence: A developmental paradox. *Discourse Processes* 43(2), 79-120.
<https://doi.org/10.1080/01638530709336894>
- Gagarina, N. (2016). Narratives of Russian–German preschool and primary school bilinguals: Rasskaz and Erzaehlung. *Applied Psycholinguistics* 37, 91–122. <https://doi.org/10.1017/S0142716415000430>
- Hickmann, M. (2003). *Children's discourse: person, space and time across languages*. Cambridge: Cambridge University Press
- Pavlenko, A. (2006). Narrative Competence in a Second Language. In H. Byrnes, H. Weger-Guntharp, & K. Sprang (Eds.), *Educating for advanced foreign language capacities* (pp. 105-118). Georgetown University Press.
- Petersen, D. B., Gillam, S. L., & Gillam, R. B. (2008). Emerging procedures in narrative assessment: The index of narrative complexity. *Topics in Language Disorders* 28(2), 115–130.
<https://doi.org/10.1097/01.tld.0000318933.46925.86>
- Strömquist, S., & Day, D. (1993). On the development of narrative structure in child L2 and adult L2 acquisition. *Applied Psycholinguistics* 14(2), 135-158. <https://doi.org/10.1017/s0142716400009528>
- Tolchinsky, L. (2014). El desarrollo de la narración más allá de la primera infancia. In R. Barriga (Ed.), *Las narrativas y su impacto en el desarrollo lingüístico infantil* (pp. 253-281). Colegio de México.
- Verhoeven, L., & Strömquist, S. (2001). Narrative development in a multilingual context. John Benjamins.

Vocabulary development in low-educated adult learners of Norwegian L2

Anna-Marie Kjøde Olsen (Universitetet i Bergen).

This study investigates the multidimensional nature of incipient lexical development in Norwegian L2 in a group of immigrant learners. Data from participants with little or no prior knowledge of the target language and with curtailed formal schooling can offer valuable insights; there is a disproportionate reliance in SLA on academic learner samples, and a growing awareness of the need to include more diverse learner backgrounds (Ortega 2019, Young-Scholten 2013).

A key priority for beginning learners is lexical acquisition. However, due to the scarcity of longitudinal studies that accommodate the multidimensional nature of vocabulary knowledge, we have limited understanding of how different aspects develop and might interrelate. Vocabulary knowledge is commonly described using the breadth-depth dichotomy. The present study takes a component approach to depth, building on Nation's (2013) description of what knowledge of single words entails. Acquisition is assumed to occur incrementally across aspects of form, meaning and use, with some being acquired before others. The research questions, grounded in usage-based theories, are:

To what degree are the different aspects of word knowledge acquired over time? And in what ways are they related?

Two of the rare studies to have followed the concurrent growth of multiple aspects of vocabulary knowledge over time, are Schmitt (1998) and Dóczy & Kormos (2016), who traced the development of a small set of words in 3 advanced and 14 pre-intermediate participants respectively. Drawing on their methodological approaches, this study tracks learners' knowledge of 9 target words over a six month period. The words were selected to meet criteria including frequency, imageability, pedagogical relevance, word class and prototypicality. Participants were 8 Arabic L1 / early-L2 speakers (3 females, 5 males, aged 19-42) with 4 to 10 years of home country schooling who at the onset were three months into their compulsory Norwegian language instruction. Individual semi-structured interviews were used to elicit knowledge of meaning, form, associations, word class and usage.

Descriptive statistics and inferential statistics (e.g. Friedman's ANOVA), will be employed to detect patterns that might elucidate the development of and relationship between different aspects of word knowledge. Individual variation will be explored through statistical comparison and case studies of two select learners. Targeting an understudied population of beginning learners of a non-English L2, this study contributes to our understanding of the multidimensional nature of vocabulary knowledge. Moreover, it has implications for L2 pedagogy.

References Dóczy, B. & J. Kormos. (2016). Longitudinal developments in vocabulary knowledge and lexical organization. Oxford University Press. Nation, I. S. P. (2013). Learning vocabulary in another language (2nd ed.). Cambridge University Press. Ortega, Lourdes. (2019.) SLA and the study of equitable multilingualism. *The Modern Language Journal*, 103: 23–38. Schmitt, N. (1998). Tracking the Incremental Acquisition of Second Language Vocabulary: A Longitudinal Study. *Language Learning*, 48(2). 281–317. Young-Scholten, M. (2013). Low-educated immigrants and the social relevance of second language acquisition research. *Second Language Research*, 29(4). 441-454.

Multilingual speakers' utterance and cognitive fluency across languages: Combining speech production and eye-tracking

Elina Lehtilä (University of Turku).

Fluency is an essential part of L2 oral proficiency, and it has traditionally been examined with respect to three dimensions: temporal features of speech (utterance fluency), listener ratings of fluency (perceived fluency), and cognitive processes underlying speech production (cognitive fluency; Segalowitz, 2010). Recent studies suggest that speakers' L1 and L2 utterance fluency are connected (Duran-Karaoz & Tavakoli, 2020), and that these connections are affected by the speakers' proficiency level in the target language (Peltonen, 2018). However, examining learners' speech fluency across multiple target languages, combined with a cognitive fluency perspective, has thus far received little attention.

The present study addresses this gap by using eye-tracking (an indicator of cognitive fluency) with L1, L2, and L3 speech production (utterance fluency) in picture description tasks. Eye-tracking studies based on L1 object naming have suggested that combining eye-movement and speech data might help to disentangle disfluencies related to word preparation difficulties from others (e.g., self-monitoring processes; Pistono & Hartsuiker, 2021). With little prior empirical research on eye movements during spontaneous speech in the L2/L3, however, the study is methodologically innovative in its use of eye-tracking to examine the cognitive processes underlying disfluencies when speakers produce complex descriptions in different languages. The dataset (collected in spring 2023) consists of speech and eye-movement data from Finnish university students (N = 40), who speak Finnish as their L1, English as their L2 (at intermediate proficiency level), and Swedish as their L3 (at beginner proficiency level). In the experiment, the participants describe complex pictures in Finnish, English, and Swedish while their eye movements are recorded. The objective of the study is to examine the relationship between L1, L2, and L3 utterance fluency and cognitive fluency with differing levels of proficiency in the target languages.

Research questions are: What kinds of differences are there between utterance fluency measures (articulation rate and frequency and duration of pauses and corrections) across L1 Finnish, L2 English, and L3 Swedish? To what extent are the utterance fluency measures correlated across the three languages? Are there differences between eye-movement measures (average and first fixation duration and number of fixations) during picture description in the three languages? To what extent are the eye-movement measures correlated across the three languages?

The data analyses focusing on the correlations and differences between utterance fluency and eye-movement measures will be completed before the conference. Preliminary results will be discussed in the presentation, with the aim of identifying future research directions and potential other analysis methods that could be applied to the dataset.

References: Duran-Karaoz, Z., & Tavakoli, P. (2020). Predicting L2 fluency from L1 fluency behavior: The case of L1 Turkish and L2 English speakers. *Studies in Second Language Acquisition*, 42(4), 671–695.

Peltonen P. (2018). Exploring connections between first and second language fluency: A mixed methods approach. *The Modern Language Journal*, 102(4), 676–692. Pistono, A., & Hartsuiker, R. J. (2021). Eye-movements can help disentangle mechanisms underlying disfluency. *Language, Cognition and Neuroscience*, 36, 1038–1055. Segalowitz, N. (2010). *The cognitive bases of second language fluency*. Routledge.

The Study of Family Language Policy (FLP) in Hokkien-Speaking Diasporic Families in the UK

Zhonghan Lin (King's College London).

Increasing global mobility has led to the formation of transcultural families with linguistic super-diversity (Blommaert and Rampton, 2011), highlighting the fluidity between languages and cultures in social interaction (Leung and Valdes, 2019). With multiple languages spoken in one household, family language policy (FLP) has been shown to have influences on members' proficiency in different languages (Garcia, 2009), as it involves language use, and the extent to which cultural values and traditions are preserved through languages (King and Lanza, 2016). For diasporic families living in countries where their heritage language (HL) is not widely spoken (Nortier, 2008), 'family' becomes a critical domain for the continued use of the HL, and FLPs that encourage the use of the HL can help preserve the language and cultural traditions (King and Lanza, 2016). In spite of the extensive work carried out in understanding how diasporic parents construct their FLPs to manage their HL transmission and support the development of their children's second language (L2) abilities (Curdt-Christiansen, 2009), not many voices have been heard from the children themselves, and not all languages(communities) received equal attention. To gain a deeper understanding of language practices among minority-speaking diasporic families, and the extent to which each family member interprets the wider social changes that interact with their language beliefs and practices at home, this ethnographic study focuses on six Hokkien-speaking families in the UK with school-aged children. This study seeks to investigate patterns of language use in these families in everyday home interactions by collecting data that include audio recordings and observations of their everyday activities. This presentation will provide an overview of the literature on FLP before outlining the planned study and its methodology. It will then present some preliminary data and findings taking account of family members' beliefs and values, language proficiency, family structures and wider societal factors. The study will also explore the usefulness of FLP as a concept for the study of the language practices of minority-speaking diasporic families, and how might FLPs bridge the familial and the societal, parents and their offspring. The findings will contribute to a deeper understanding of the dynamics of a particular diasporic community and of other communities.

The Interrelationship Between Foreign Language Enjoyment, Foreign Language Classroom Anxiety and Beliefs about Language Learning

Yanan Lu (University of Birmingham) and Bene Bassetti (University of Birmingham).

Emotions are of great importance for second language (L2) learning and achievement (Dewaele & Li 2018). However, past research on L2 learning emotions has traditionally focused on negative emotions, especially language anxiety (Dewaele & Li 2018). The introduction of Positive Psychology in second language acquisition (SLA) research in 2012 (MacIntyre & Gregersen 2012) encouraged researchers to look beyond anxiety for a wide range of diverse L2 emotions. Positive and negative classroom emotions and their complex and dynamic interactions have been examined from various linguistic, educational and sociocultural contexts (Dewaele & Li 2018). However, little is known about the emotional experiences of the learners who learn Chinese as a foreign language. In addition, the relationship between emotion and cognitive individual differences is largely unaddressed in SLA literature. The present mixed-method study examines whether and to what extent learners' FLE and FLA are linked to their beliefs about language learning as well as a range of learner variables. A group of 107 British university students of Chinese as a Foreign Language completed the Beliefs about Language Learning Inventory (Horwitz 1988) and the short forms of the FLE scale (Botes et al., 2021) and the FLCAS scale (Horwitz et al., 1986), as well as the Self-perceived Linguistic Confidence Scale (SLCS, Pyun et al., 2014). Ten of them later participated in a one-to-one online semi-structured interview. Statistical analyses revealed that higher levels of FLE and lower levels of FLCA were linked to higher scores on learners' beliefs in their ability in learning Chinese well and their self-perceived linguistic confidence. Gender, course levels and study abroad experience had no effect on learners' levels of FLE and FLCA. The qualitative data revealed a more nuanced picture of learners' emotional experiences of learning Chinese. Their FLE mainly came from the sense of achievement, teachers' flexibility in teaching Chinese and peer students' emotional support. Their FLCA originated from the difficulty of Mandarin Chinese, peer pressure and their belief that aptitude is required to learn languages well.

Priming and the L2 acquisition of motion event constructions

Anna Michelotti (University of Mannheim) and Helen Engemann (University of Mannheim).

The way in which motion is expressed varies considerably across languages and has been shown to be a domain in which L2 learners have difficulties adapting to the target language norms (e.g., Laws et al., 2022).

According to Talmy's typology (Talmy, 1983), languages are divided in two categories according to the way they encode motion events. Satellite-framed languages (e.g., German) usually encode manner of motion in the verb stem. Path of motion, in turn, is usually encoded via prepositions and particles (1). On the other hand, verb-framed languages (e.g., Italian), usually encode path of motion in the verb stem, while manner is either omitted or expressed in the verbal periphery (2).

(1) Der Mann rennt ins Gebäude. THE MAN RUNS IN.THE BUILDING (2)

L' uomo entra nell' edificio correndo. THE MAN ENTERS IN.THE BUILDING RUNNING

By selecting one of these two structural options, the speaker foregrounds specific spatial information (e.g., either path or manner of motion). Therefore, acquiring the target-like ability to talk about motion in a foreign language does not only involve structural and lexical knowledge, but may also imply learning a new way to conceptualize space.

Given this situation, it is important to answer the following questions: how can the target-like use of motion events be acquired by speakers of typologically different languages? Is implicit learning possible given that the acquisition of these structures goes beyond lexical and structural competence?

In my PhD project, I want to investigate whether implicit learning, in the form of priming, has an impact on the production of motion events by a group of L2 learners of Italian whose native language is German. Additionally, I am interested in testing whether priming can be used as a tool for L2 teaching of motion events.

Priming can be described as the speaker's tendency to re-use linguistic structures they have previously heard. Different studies have shown that priming does not only occur at the structural level but also at the conceptual level (Bunger et al., 2013). If priming effects are long lasting, priming might also have the potential to be used to facilitate the acquisition of motion events by L2 learners.

Discussion questions: - Can priming be used as a tool to generate implicit learning of L2 motion events? - How to isolate the impact of semantic, structural, and conceptual priming on the implicit learning process of motion construction?

Bunger, A., Papafragou, A., & Trueswell, J. C. (2013). Event structure influences language production: Evidence from structural priming in motion event description. *Journal of Memory and Language*, 69(3), 299–323. <https://doi.org/10.1016/j.jml.2013.04.002>

Laws, J., Attwood, A., & Treffers-Daller, J. (2022).

Unlearning the boundary-crossing constraint: Processing instruction and the acquisition of motion event construal. *International Review of Applied Linguistics in Language Teaching*, 60(4), 1089–1118.

<https://doi.org/10.1515/iral-2020-0147> Talmy, L. (1983). How Language Structures Space. In H. L. Pick & L. P. Acredolo (Eds.), *Spatial Orientation: Theory, Research, and Application* (pp. 225–282). Springer US.

https://doi.org/10.1007/978-1-4615-9325-6_11

The effects and the effectiveness of different types of instruction on the acquisition of Dutch vowels by L1 speakers of Portuguese

Francisco Miguel Valada (Vrije Universiteit Brussel, Universidade NOVA de Lisboa).

This research aims at investigating how phonological instruction affects the acquisition of phonemic categories in a second language (L2). More specifically, it investigates how the impact of instruction is moderated by (a) the type of instruction (implicit vs. explicit), (b) the type of modality (perception vs. production), and c) the relative learning difficulty of L2 segments (more difficult vs. less difficult). The phonological target contrasts in this study are the Standard Dutch vowels /ɑ-a:/ and /ɪ-i/, as learned by native speakers of Portuguese. Their phonemically contrastive status in Dutch, but not in Portuguese, renders them problematic for L1 Portuguese learners of Dutch. These contrasts are also expected to present varying degrees of difficulty to Portuguese learners: the /ɑ-a:/ contrast involves both a difference of length and quality, whereas the /ɪ-i/ contrast is essentially one of quality and, except for the lengthening of /i/ before /r/ (Gussenhoven, 1992), both /ɪ/ and /i/ are phonetically short. The presence of additional length cues for /ɑ-a:/ may therefore make this contrast easier to perceive for Portuguese speakers than /ɪ-i/.

Two experiments will evaluate the effects of implicit and explicit instruction on the L2 acquisition of /ɑ-a:/ and /ɪ-i/ by L1 Portuguese speakers with no previous contact with Dutch. Experiment 1 consists of a phonological forced-choice identification task, where Portuguese speakers are asked to categorize Dutch sounds, mapping them onto Portuguese categories, to verify whether the assumptions of difficulty with regard to the two contrasts are correct. Experiment 2 has a pretest-posttest-delayed posttest design, assessing perceptual abilities with an AX discrimination task, and production abilities with a delayed repetition task. This experiment includes type of treatment (implicit vs. explicit) as a between-subject factor, a control group (with narrative input), and a baseline condition with L1 Dutch speakers. Implicit instruction will consist of a High Variability Phonetic Training (HVPT) regimen. Explicit instruction includes meta-linguistic information.

Specific questions for the discussants:

1. Do you believe that the implicit/explicit distinction typically used in grammar can be applied as such to phonological instruction? How might the specificities of the phonological domain impact our understanding of this distinction?
2. Working with A0 participants (complete beginners) may pose a number of challenges, for instance when developing a suitable production task. In terms of the design of the experiment, what would your recommendations be to address such challenges?
3. As for the control group, other studies reporting activities with beginner participants (Brosseau-Lapr e et al., 2013; Peltekov, 2020) mention communicative exercises without special focus on pronunciation, but no details are provided. Bearing in mind the A0 level of the participants, what might a suitable activity be for the control group?

References

- Brosseau-Lapr e, F., Rvachew, S., Clayards, M., & Dickson, D. (2013). Stimulus variability and perceptual learning of nonnative vowel categories. *Applied Psycholinguistics*, 34, 419–441.
- Gussenhoven, C. (1992). Dutch. Illustrations of the IPA. *Journal of the International Phonetic Association*, 22, 45–47.
- Peltekov, P. (2020). The effectiveness of implicit and explicit instruction on German L2 learners' pronunciation. *Die Unterrichtspraxis/teaching German*, 53(1), 1–22.

Corrective feedback provision in mobile instant messaging: Exposure conditions and timing

Bridget Murphy (University of Barcelona).

Mobile instant messaging (MIM) is underutilized yet promising as a language learning tool in the L2 classroom. The 'reply' feature of MIM applications make them great platforms through which to provide written corrective feedback (CF) to learners' messages, and the permanent and salient nature of text messages make MIM chats ideal environments for noticing and L2 acquisition to occur (Andujar, 2020; Schmidt, 1990). The purpose of the present doctoral dissertation is to better understand how written CF works in a MIM setting. WhatsApp will be used to investigate two different factors of feedback provision and their effects on English acquisition: timing (see Li, 2018) and exposure condition (see Yilmaz, 2016). To my knowledge, no studies exist that consider the variables of CF timing or exposure condition on L2 acquisition in a MIM context. Study 1 will focus on CF timing, which refers to the temporal schedule when CF is provided, and will take place during the present academic year in a research lab in dyads (researcher and student). Participants will be first-year undergraduate EFL students at a B2 (CEFR) level. One group will receive CF (corrective recasts) while completing two language-based tasks in a WhatsApp chat, and another group will receive the CF immediately after completing the two tasks. A control group will complete the tasks without any CF. Feedback will be given on students' grammatical accuracy after several expressions related to expressing regret and giving advice. A timed oral elicited imitation task will measure students' implicit knowledge of these structures, and an untimed use of English test will measure students' explicit knowledge. These tests will be given one day before, immediately after, and one week after the treatment. Study 2 will focus on exposure conditions, which refers to how students are exposed to content in a classroom either as direct receivers of content or as observers (Yilmaz, 2016). Participants will be an intact class of adult EFL (B2 level) learners enrolled in an intensive summer course at a language school (in 2024). Students will complete several communicative tasks in a WhatsApp chat over several days outside of class. Then, in class, students will complete a pretest individually correcting erroneous messages from the chat. The teacher will then prompt students to repair the erroneous messages through a whole-class post-task CF activity in WhatsApp. Individual, in-class immediate and delayed posttests (based on the same target language as the pretests) will measure students' learning (and perhaps noticing) from the CF activity. Random students will be selected to complete semi-structured interviews with stimulated recall items to provide introspective data on exposure conditions and noticing of the CF episodes. Questions: 1. What individual difference factor would make sense to study in Study 2 about exposure conditions? 2. What is the best way to measure "noticing" of CF episodes in a MIM chat in Study 2? References Andujar, A. 2020. "Mobile-mediated Dynamic Assessment: A New Perspective for Second Language Development." *ReCALL* 32 (2): 178–194. <https://doi.org/10.1017/S0958344019000247> Li, S. (2018). What is the ideal time to provide corrective feedback? Replication of Li, Zhu & Ellis (2016) and Arroyo & Yilmaz (2018). *Language Teaching*, 53(1), 96–108. <https://doi.org/10.1017/s026144481800040x> Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129–158. <https://doi.org/10.1093/applin/11.2.129> Yilmaz, Y. (2016). The role of exposure condition in the effectiveness of explicit correction. *Studies in Second Language Acquisition*, 38(1), 65-96. <https://doi.org/10.1017/S0272263115000212>

Socio-affective Factors, Retirement Resources, and Variability in L2 Trajectories of Third Age Learners

Annika Rossmann (University of Zurich).

Theoretical Framework The role of socio-affect during retirement is relevant as social and mental well-being impact this significant life event that often comes with major changes in an individual's lifestyle. As a complex and stimulating activity, second language (L2) learning has also been found to contribute to maintaining and improving social, emotional, and motivational resources in the third age (Pfenninger & Kliesch 2023). In line with the resource-based dynamic model (Wang et al., 2011), retirees performing more preparatory activities before retirement are hypothesized to possess more retirement resources, which then contribute to better well-being after retirement (Reitzes & Mutran, 2004; Scheibe & Zacher, 2013; Yeung & Zhou, 2017). These retirement resources vary greatly amongst older adults, and this may arguably modulate L2 learning in the transition from work to retirement. While factors of retirement adjustment and the association between retirement and mental and physical health are well-documented (e.g. Schmitt, 2018), there have been no studies that analyse if and how retirement impacts an individual's L2 development and use and vice versa.

Research Questions This PhD thesis is part of a large-scale longitudinal research project called VARIAGE; it addresses the following general research question: How do individual differences in social-contextual, emotional, and motivational resources relate to L2 performance during the transition from work to retirement? In other words, who profits from L2 learning across retirement, and when and why?

Methods This longitudinal study will analyse the L2 trajectories of older German-speaking adults (aged 63-65) in Switzerland who attend an intensive 2-year L2 English course, and who retire sometime in the first 18 months of the L2 training. The 40 participants are assessed at 3-week intervals on (1) a range of socio-affective measures: the Satisfaction of Life, Loneliness, Friendship, Retirement Anxiety, and Attitudes towards Retirement scales, self-efficacy, scale of Pre-retirement Planning Activities, L2 motivation, Adelaide Activities Profile, Attitude/Motivation Test Battery, (2) a L2 test battery (Kliesch & Pfenninger, 2021), and (3) demographic measures. In a mixed-methods design, combining generalized additive mixed modeling (Wood, 2006), time-series cluster analysis (Peng et al., 2022), experience sampling (Arndt et al., 2021), and qualitative thematic analyses (Braun & Clarke, 2006), this study aims to identify pre- and post-retirement activities, critical life events, and retirement resources that mediate effects of retirement on language acquisition and use and to help explain inter- and intra-individual differences.

Expected results We hypothesize that higher levels of social and mental well-being and therefore more retirement resources improve retirement adjustment and impact L2 learning. Specifically, we expect periods of significant L2 growth to correlate with higher levels of social and mental well-being. As for the potential stimuli for change perceived by the participants, this remains an open empirical question.

Innovation Because learners' needs vary, this project will offer in-depth insight into individual learner profiles and the progression of retirement resources over time. The outcomes should benefit older adults' planning for retirement, facilitate sensitivity to intervention analysis, and help tailor future language-based interventions for older individuals by making the L2 curriculum available online.

Questions for Doctoral Workshop 1. Can you think of any factors relating to social and mental well-being that I should be measuring? 2. Given the nature of the project (intensive L2 training over a long period of time), do you have ideas to keep participant attrition low? 3. Are there other analytic tools (quantitative or qualitative) that I should consider?

Lexical Stepping Stones and Stumbling Blocks in Early Foreign Language Learning – Lexical Effects on Morphosyntactic Acquisition of Primary School Students

Marie Schnieders (English Linguistics, TU Braunschweig).

At the early stages of lexical learning, foreign language (FL) instruction builds on similarities (i.e. cognate words) between the majority language and the FL. This teaching strategy is linked to psycholinguistic research which suggests that lexical access is not language specific, and can thus lead to facilitation effects in the processing of cognate words (Libben & Titone, 2009). Current language processing models show cross-linguistic lexical activation to have knock-on effects on grammatical learning (e.g. the Lexical Bottleneck Hypothesis; Hopp, 2018). Facilitation of cognate word retrieval can be a stepping stone for grammatical learning since it allows the learner to save processing resources for other aspects of language processing, e.g. morphosyntax. At the same time, word similarity promotes language-transfer at lemma-level. The mismatch between the L1 and the FL can be costly in terms of processing thereby constituting a stumbling block for grammatical learning. My PhD-project comprises 2 cross-sectional studies and a longitudinal study in which I will examine the potential and impediments of cognate effects in German primary-school students learning English as FL. In study 1, I will focus on cross-linguistic activation in terms of a cognate-advantage for morphological learning. Students' usage of subject-verb agreement (SVA; 3rd person singular -s) with cognate and noncognate verbs prior to instruction will be investigated. I predict that if lexical processing costs are reduced by the presence of cognate words, students will be more accurate in using target-like SVA than in sentences containing a noncognate word. In study 2, I will investigate the consequences of cross-linguistic lexical activation in terms of a cognate-disadvantage for learning conversion in English. In particular, I expect conversion to be constrained by L1-options (i.e. film \diamond to film/filmen vs. radio \diamond to radio/*radio-en). Difficulties with conversion can be taken as evidence for transfer of German knowledge to English at lemma-level. No difficulties with conversion could either be indicative of the students establishing an L2 lemma-entry or of their insensitivity to the syntactic context and the interpretation of these findings constitutes a point for discussion. In study 3, I will examine individual differences in cross-linguistic lexical activation in terms of examining longitudinal relations between early cognate advantages in segmentation and later learning outcomes in vocabulary and grammar. Research on infants suggests speech-segmentation abilities to be an essential skill for later lexical and grammatical acquisition outcomes (Newman et al., 2006). If this also holds for FL-learning, students with greater initial segmentation skills should have higher scores on standardized lexical and morphosyntactic tests after 1.5 years of FL-instruction. Taken together, these studies will give insight into possible synergies between vocabulary and grammar acquisition.

References Hopp, H. (2018). The bilingual mental lexicon in L2 sentence processing. *SecLang*, 17(1), 5-27. https://doi.org/10.11431/secondlanguage.17.0_5 Libben, M., & Titone, D. (2009). Bilingual lexical access in context: evidence from eye movements during reading. *ExpPsych*, 35(2), 381-390. <https://doi.org/10.1037/a0014875> Newman, R et al. (2006). Infants' early ability to segment the conversational speech signal predicts later language development: a retrospective analysis. *DevPsych*, 42(4), 643-655. <https://doi.org/10.1037/0012-1649.42.4.643>

Every puzzle has an answer: learning a language by playing video games.

Natalia Shalaeva (Swansea University).

Corpus linguistics generated clear evidence that language is highly ‘formulaic’ - made up of multiword expressions that appear to be prefabricated (f.ex. fixed phrases, collocations, phrasal verbs, etc.). In fact, formulaic sequences account for 50-80% of native speaker English and, therefore, play a key role in organising discourse (Wray, 2005). Yet, even advanced ESL learners typically have limited knowledge of formulaic sequences (Hatami, 2015) and often use them inaccurately (Wray, 2005). Pedagogic research has shown positive impacts of explicit formulaic language teaching on learners’ awareness of formulaic sequences (Hatami, 2015). However, nowadays ESL learners find traditional approaches too passive and failing to produce immediate results. Digital natives prefer video games to reading and Duolingo to traditional classroom vocabulary drills. This study aims to explore how a computer game can be adapted as a learning tool to support acquisition of formulaic sequences: 1) Does playing commercial adventure games contribute to ESL students’ acquisition of formulaic sequences? 2) Does playing commercial adventure games supplemented with gamified instructional materials contribute to ESL students’ acquisition of formulaic sequences? The participants (ESL students aged 18-28 who studied English for at least 7 years) are asked to play 2 chapters of Thimbleweed Park (story-driven adventure game) over the course of 4 consecutive days. With 16.248 recorded dialogue lines, Thimbleweed Park provides the players with an ample amount of comprehensible input and its non-linear gameplay ensures enough reoccurrences of newly encountered formulaic sequences, while simultaneously providing the players with the proper context. The participants are randomly assigned to either actively play the game or watch someone play it (passive gameplay). For some of the participants, gameplay sessions are followed by gamified vocabulary exercises. Built purposefully to teach chapter-specific formulaic sequences encountered in the in-game dialogues, these mini-games both assess participants’ noticing and transform monotonous vocabulary drills into animated game-like experiences that enhance learning motivation and engagement which should lead to higher retention rates (Papi, 2018). This is the first (to my knowledge) study that explores the possibility of using commercial adventure games for the purpose of formulaic language learning. For the EUROSLA workshop, I would like to discuss potential study design pitfalls: 1) What are the risks of mixing online and offline data? What is the best way to analyse such data? 2) Should this time-on-task variance be addressed in experiment design? And what are the possible scenarios to do so? References: - Hatami, S. (2015). Teaching Formulaic Sequences in the ESL Classroom. TESOL Journal, 6(1), 112–129. - Papi, M. (2018). Motivation as quality: Regulatory Fit Effects on Incidental Vocabulary Learning. Studies in Second Language Acquisition, 40(4), 707–730. - Wray, A. (2005). Formulaic language and the lexicon. Cambridge University Press.

The Motivation for Learning Mandarin Chinese in UK Higher Education

Siying Shen (University of Leeds).

This ongoing doctoral project is a longitudinal qualitative inquiry that tracks the motivation of six Anglophone learners of Mandarin over their first academic year at a British university. Whereas the existing literature has provided profound insights into L2 motivation at a group level, we seem to know much less about it in an individual sense, namely, the nuances of people's unique motivational dynamics in their own contexts over time, as they address the challenges of learning a new language. Responding to the need for a more 'detailed' focus on individual learners in L2 motivation research (Ushioda, 2020), this study takes a holistic relational viewpoint (Ushioda, 2009), and aims to explore students' lived experiences in their Mandarin course with a focus on their motivation-in-action, through integrating two seemingly 'contrasting' theoretical perspectives in the field: investment (Darvin and Norton, 2015) and engagement (Hiver et al., 2021).

The project aims to answer the following research questions: 1) What are the initial learning motivations of Chinese language students in the BA Chinese programme? 2) How do Chinese language students' motivation change over the first academic year in the UK? 3) What factors are associated with the changes in the Chinese students' motivation during this period? To address these questions, four data sources are used to understand motivational changes in context holistically, including: three rounds of in-depth interviews of six participants, weekly motivation tracker where students rate their motivation from 1 to 10 and add their comments, three semi-structured interviews with stakeholders, two rounds of classroom observations along with interactional data and field notes. Whereas the first two sources will be used to generate participants' narratives, the other two serve to triangulate with participants' meaning making of their own experience. Data collection is expected to be finished by June 2023.

In this workshop, the underpinning framework for the study and its methodology will be introduced, before I present my preliminary findings and how they align with and contribute to the existing knowledge in L2 motivation field. This presentation intends to generate further reflections and discussion about potential conceptual links between investment and engagement, and the data-rich approach the project has taken to capture students' motivation-in-action holistically.

References Darvin, R. and Norton, B. 2015. Identity and a model of investment in applied linguistics. *Annual Review of Applied Linguistics*. 35, 36-56. Hiver, P., Al-Hoorie, A.H. and Mercer, S. 2021. *Student Engagement in the Language Classroom*. Bristol: Multilingual Matters. Ushioda, E. 2009. A person-in-context relational view of emergent motivation, self and identity. In: Dörnyei, Z. and Ushioda, E. eds. *Motivation, language identity and the L2 self*. Bristol: Multilingual Matters, 215-228. Ushioda, E. 2020. *Language Learning Motivation: An Ethical Agenda for Research*. Oxford, United Kingdom: Oxford University Press.

The Effect of Contextual Richness on L2 Learners' Processing and Incidental Acquisition of Nonliteral English Phrasal Verbs: An Eye-tracking Study

Tianjiao Song (Faculty of Education, University of Macao) and Barry Lee Reynolds (Faculty of Education, University of Macau; Centre for Cognitive and Brain Sciences, University of Macau).

Incidental vocabulary acquisition occurs when learners “pick up” new words as by-products while being engaged in meaning-focused activities, such as reading (Hulstijn, 2003; Reynolds, 2015; Uchihara, Webb, & Yanagisawa, 2019). Some research suggested that contextual richness, the amount of information a context contains for successful inferencing of word meaning, contributed to incidental vocabulary acquisition (e.g., Webb, 2008). However, this line of studies used offline vocabulary posttests to examine incidental learning gains and limited their scope to the acquisition of single words (e.g., Mohamed, 2015; Webb, 2008). It remains unclear the role context plays in the processing and acquisition of nonliteral English phrasal verbs (i.e., PVs), a particular type of multi-word lexical item notoriously difficult for L2 learners to acquire (Gardner & Davies, 2012). To fill these gaps, this study used eye-tracking technology to gauge the effect of contextual richness on L2 learners' processing and incidental acquisition of nonliteral English PVs by answering the questions: 1. How does contextual richness affect L2 learners' processing of nonliteral PVs while reading; 2. What is the relationship between the online processing of nonliteral PVs and the acquisition of these PVs? A homogenous group of Chinese-L1 EFL postgraduates ($n = 57$) with similar L2 reading abilities and L2 vocabulary sizes were involved in the study. The participants were randomly divided into three context groups, with each group reading 8 low-frequency nonliteral English PVs that occurred 10 times in one of the three contexts: more informative, less informative, and uninformative. The participants' eye movements while reading were recorded by an EyeLink Portable Duo. Analyzing the participants' online (i.e., eye movements) and offline (i.e., vocabulary post and delayed posttests) data with mixed-effects modeling, the authors found significant interactions between contextual richness and eye movements that indicate initial lexical access (i.e., FFD: first fixation duration) and reanalysis (i.e., regressions-in) in predicting the incidental learning gains. Longer FFD on PVs form in the less informative context significantly enhanced the PVs form retention. Making regressions-in facilitated meaning recognition when the context was less informative, but promoted meaning recall when the context was more informative. These results revealed L2 learners' different cognitive processes involved in processing and learning nonliteral PVs under the different levels of context. The findings of this study further provided empirical support for the theories of language processing (e.g., TOPRA) and acquisition (e.g., the Noticing Hypothesis).

Nasal vowels, primary word stress production, and their role in foreign-accented speech of L1 Ukrainian speakers of Portuguese as an L2

Yolanda Xavier (Linguistics Research Centre of NOVA University Lisbon (CLUNL), School of Social Sciences and Humanities (NOVA FCSH)) and Susana Correia (Linguistics Research Centre of NOVA University Lisbon (CLUNL), School of Social Sciences and Humanities (NOVA FCSH)).

Most studies on linguistic stereotyping focus on revealing this phenomenon from a psychological perspective (Bazzi et al., 2022). However, no research explains which features of foreign accent contribute to linguistic prejudice. The current study is a part of a PhD thesis that aims to investigate if absent or deviant segmental and suprasegmental features can trigger linguistic stereotyping for L1 Ukrainians who are learning (European) Portuguese as an L2. Additionally, my PhD project will research if this impacts further social integration by denying access to higher-level positions. The present study will compare two separate segmental and suprasegmental features that differ in Ukrainian and Portuguese: nasal vowels ([ĩ ũ ě õ ě]), which are inexistent in Ukrainian, alongside the distribution of primary word stress, which can be misplaced due to the rare bidirectional stress iteration in Ukrainian (Łukaszewicz & Motczanow, 2018). The research has set the following questions: 1) Do L1 Ukrainians produce nasal vowels with reduced nasality? 2) Is there a misplacement of primary word stress in EP? The participants will be native speakers of Ukrainian (B1, B2, C1 levels: n=23/group). Native European Portuguese speakers from the Lisbon area will be included in the control group (n=23). The method is an audiovisual repetition task replicating Oh et al. (2011), where participants will be presented 15 coloured pictures with known objects and asked to repeat them three times each: first time by listening to an auditory cue with nasal vowels; second and third time without listening to the auditory cue. There will be a distractor introduced between the second and the third time: a participant will be asked to spot differences in two pictures. Only non-cued productions will be analysed. An acoustic analysis will be made in Praat. The results will be evaluated on nasal vowel duration, resonance and formant frequency. For primary word stress, the production task will be the same, but with different words. The analysis will be made considering intensity and duration of the syllables. We hypothesize that L1 Ukrainians will produce nasal vowels with reduced nasality and replace them with oral vowels. For primary word stress, it might be misplaced due to the different stress systems. The results will lay a foundation for the upcoming studies of my thesis and serve as a basis for pronunciation teaching for non-native learners.

Question for discussants:

How to control lexical size in a production task of spontaneous speech?

Bazzi L., Brouwer S., Planelles Almeida M., Foucart A. (2022). Would you respect a norm if it sounds foreign? Foreign-accented speech affects decision-making processes. *PLoS ONE*, 17(10), 1-21.

Oh, G. E., Guion-Anderson, S., Aoyama, K., Flege, J. E., Akahane-Yamada, R., & Yamada, T. (2011). A one-year longitudinal study of English and Japanese vowel production by Japanese adults and children in an English-speaking setting. *Journal of Phonetics*, 39(2), 156–167.

Łukaszewicz, B., & Motczanow, J. (2018). Leftward and rightward stress iteration in Ukrainian: acoustic evidence and theoretical implications. In B. Czaplicki, B. Łukaszewicz & M. Opalińska (Eds.), *Phonology, Fieldwork and Generalisations*. (pp. 259-275). Peter Lang.

Second Language Speech Development in Study Abroad context: The Role of Language Using Experience and Social Network

Tongzhou Xiao (King's College London).

Study Abroad (SA) has been considered an ideal context for second language (L2) development but existing literature on the effects of immersion abroad has drawn mixed results (Isabelli-García et al., 2018). More recently, a line of studies suggests that not only the quantity and quality of sojourners' L2 use but their social networks could also be vital in revealing their L2 immersion and engagement level and associating with the ultimate development (see Dewey, 2017; Mitchell, 2023 for reviews). The SA field has been calling for more studies to examine sojourners' L2 development by modelling and analysing the patterns of their social networks, especially those that could longitudinally collect multiple-time data because one's networks may change over time (McManus, 2019). Moreover, given that the current field primarily reports on the experiences of whom departure from the Anglophone countries (e.g., the US) and the European countries (e.g., Spain), many have also called for more voices from SA participants with different backgrounds globally to enrich the robustness (Kinginger, 2013). With attempts to fill the above gaps, this project aims to longitudinally examine the effects of language using experience and social networks on L2 speech gains (from the perspective of successful communication) in an academic study abroad context. Forty-six mainland China international students participated three times in data collection during their one-year academic sojourns in the UK. Data collection instruments include A background questionnaire and Oxford Quick Placement Test (OUP, 2001); A language contact questionnaire adapted from Freed et al. (2004); and a social network questionnaire adapted from Dewey et al. (2012). Semi-structured speech-eliciting activities designed to fit this current research context are used to collect their L2 speech samples. Moreover, semi-structured interview data are used to complement the questionnaires. Their L2 performance will be measured through global (comprehensibility & accentedness) and specific (segmental, suprasegmental/prosodic, temporal, lexical and grammatical) dimensions and linked with L1 & L2 use, social network variables and patterns.

Questions for discussion: 1. Despite some studies in the SA context, the social network variable has yet to be widely examined in the broader SLA field, to what extent the Social Network Analysis could inform the SA and wider SLA field. 2. Could other complementary or alternative measures (e.g., Functional Adequacy) apart from the ones mentioned above be applied to showcase their L2 speech gains (regarding better communication) in monologue and dialogue?

Dewey, D. P. (2017). Measuring social interaction during study abroad: Quantitative methods and challenges. *System*, 71, 49–59. <https://doi.org/10.1016/j.system.2017.09.026>

Dewey, D. P., Bown, J., & Eggett, D. (2012). Japanese Language Proficiency, Social Networking, and Language Use during Study Abroad: Learners' Perspectives. *Canadian Modern Language Review*, 68(2), 111–137.

<https://doi.org/10.3138/cmlr.68.2.111> Freed, B. F. (1990). Language learning in a study abroad context: The effects of interactive and non-interactive out-of-class contact on grammatical achievement and oral proficiency

Linguistics, language teaching and language acquisition: The interdependence of theory, practice and research, 459-477. Freed, B. F., Dewey, D. P., Segalowitz, N., & Halter, R. (2004). The Language Contact Profile. *Studies in Second Language Acquisition*, 26(02).

<https://doi.org/10.1017/s027226310426209x> Isabelli-García, C., Bown, J., Plews, J. L., & Dewey, D. P. (2018). Language learning and study abroad. *Language Teaching*, 51(4), 439–484.

<https://doi.org/10.1017/s026144481800023x> Kinginger, C. (2013). Social and cultural aspects of language learning in study abroad. John Benjamins Publishing Company.

McManus, K. (2019). Relationships between social networks and language development during study abroad. *Language, Culture and Curriculum*, 32(3), 270–284. <https://doi.org/10.1080/07908318.2019.1661683>

Mitchell, R. (2023). Documenting L2 input and interaction during study abroad: Approaches, instruments and challenges. *Second Language Research*, 39(1), 59–83. <https://doi.org/10.1177/02676583211030626> Oxford University Press. (2001). Quick Placement Test.

Retirement as a Predictor of Change in Cognitive Functioning and L2 Learning: Evidence from a Longitudinal Micro-development Study

Lukas Zbinden (University of Zurich).

Theoretical Framework

Retirement represents a significant life event that implies major changes in individuals' lifestyles and in their involvement in activities that contribute to maintaining cognitive functioning. Existing research suggests later-life cognition to be associated not only with occupational complexity and activities' intellectual demands (e.g., Carr et al., 2020), but also with L2 learning, partly because L2 learning engages an extensive neural network in the brain's fronto-temporal and parietal regions (Antoniou et al., 2013). No studies have investigated how a change in occupational status (from working to retired) may impact cognition as well as L2 performance. Because cognitive performance predicts performance differences in L2 learning and use (e.g., Kliesch et al., 2021) and because occupational status has been found to be beneficial for L2 aptitude (Roehr-Brakin et al., submitted), we expect retirement to result in periods of decrease in both cognitive functioning and L2 development, mediated by the intellectual demands of activities our participants engage in.

Research Questions

This PhD thesis is part of a large-scale longitudinal project (VARIAGE) and aims to address the following questions: (1) Does the transition from work to retirement correlate with a significant change in (neuro)cognitive functioning and L2 development? If yes, when exactly? (2) How do job complexity and pre- and post-retirement activities affect older L2 adults' (neuro)cognitive functioning and L2 learning across retirement?

Methods

The study includes 40 healthy, native German-speaking employees in Switzerland (aged 63-65) who participate in a 2-year-long intensive English course (two 90-minute sessions per week) for near-beginners and retire within the first 18 months of the training. They are assessed every three weeks on behavioural, L2, and neurophysiological parameters: • simple and complex working memory and executive functioning (see Kliesch et al., 2022), • receptive and productive L2 tests (see Kliesch and Pfenninger, 2021) • beta-band networks measured by EEG (e.g., Kliesch et al., 2022). Moreover, we implement the experience sampling method (Gregersen et al., 2020) using the app SEMA3 (<https://sema3.com/index.html>) to document (i) our participants' activities, and (ii) their daily usage of English and other languages as L2. The data is analysed using generalised additive mixed modelling (Wood, 2006) and timeseries cluster analyses (Peng et al., 2022). The research proposal will be submitted to the Ethics Board of the University of Zurich.

Expected Results

In line with the outlined literature, the following hypotheses guide this study: (a) We expect to observe periods of significant decrease in both cognitive functioning and L2 development post- rather than pre-retirement. (b) We expect job complexity and the intellectual demands of pre- and post-retirement activities to mediate the impact of occupational status on changes in (neuro)cognitive functioning and L2 development.

Innovation

L2 learning in the third age and across retirement as a meaningful and ecologically relevant pastime activity is still under-researched. Shedding light on the causes of cognitive decline and the possibilities to stave it off contributes significantly to public health in view of increased care costs associated with ageing and the trend that working lives in Europe are getting longer (e.g., Leibson et al., 2015).

Questions for the Doctoral Workshop

- Can you think of other methods (scales, etc.) to operationalise the complexity/intellectual demands of participants' (pre- and post-retirement) activities?
- Experience sampling method: At what interval can the questionnaire be administered to strike the ideal balance between data quantity/richness and manageability for the participants? (We aim to picture accurately the activities our participants engage in and to counteract the problem posed by their retrospective reporting. At the same time, however, we do not want to swamp our participants with questionnaires, as this might increase participant attrition.)

Oral presentations

Exploring the relationship between educational background, vocabulary learning strategy use, and vocabulary knowledge in immigrants learning L2 Swedish

Anders Agebjörn (Malmö University) and Lars Bokander (Jönköping University).

Research in SLA has traditionally involved mainly well-educated language learners. Therefore, scholars have recently pointed out the need to investigate more diverse populations of L2 learners (Andringa & Godfroid 2019). The present project responds to this call by examining the role of educational background in vocabulary-learning-strategy use and vocabulary knowledge in immigrants studying L2 Swedish. Research on individual differences in SLA suggests that strategy use may affect learning outcomes and that strategies might also be successfully instructed (Griffiths, 2022). Hence one can speculate that well-educated learners have generally more efficient strategies, giving them an advantage compared to less educated learners. However, findings are inconclusive with regard to the relationship between educational background and strategic behavior (LaBontee 2019). The present study addresses this knowledge gap.

In this ongoing project, a questionnaire was distributed to 42 adult immigrants studying L2 Swedish at the A2 level. The group was heterogeneous with regard to age (Mdn = 35; IQR = 25–45), length of residence in Sweden (Mdn = 4; IQR = 2–5), years of education (Mdn = 10; IQR = 9–14), and language background; the most represented L1s among the participants were Arabic, Kurdish and Farsi. Vocabulary knowledge was measured with a modified version of Bokander's (2016) SweLT test, with items selected to match the participants' level of L2 Swedish. Use of vocabulary-learning strategies was examined with a modified version of Labontee's (2019) SVLSS 2.0 test. Together, these data enable us to explore interactions between educational background, vocabulary knowledge, and strategy use, while controlling for participants' age, time of residence, and language background.

Data collected so far indicate that learners with a shorter education generally use fewer vocabulary-learning strategies, compared to learners with longer education. Importantly, learners with shorter education also appear to prefer basic strategies, like writing and rehearsing wordlists, and this preference is negatively correlated with vocabulary knowledge. In contrast, learners with longer education prefer contextual strategies, like guessing the meaning of new words, and this preference is positively correlated with vocabulary knowledge. In other words, while less-educated learners appear to focus on rehearsing words, more well-educated learners strive to learn new words in an autonomous way, which may enable them to acquire a larger vocabulary. These results suggest that research on vocabulary-learning-strategy use in well-educated L2 learners cannot be directly generalised to other populations. In addition, the findings may be helpful for L2 teachers aiming at individualising their strategy-use instruction.

References Andringa, S., & Godfroid, A. (2019). SLA for all? Reproducing second language acquisition research in non-academic samples. *Language learning*, 69, 5–10.

Bokander, L. (2016). SweLT 1.0: konstruktion och pilottestning av ett nytt svenskt frekvensbaserat ordförrådstest. *Nordand: nordisk tidskrift för andrespråksforskning*, 11(1), 39–60.

Griffiths, C. (2022). Learning Styles and Strategies. In: S. Li, P. Hiver, and M. Papi (Eds.), *The Routledge Handbook of Second Language Acquisition and Individual Differences*, 82–94. Routledge.

LaBontee, R. (2019). Strategic vocabulary learning in the Swedish second language context. Doctoral thesis. University of Gothenburg.

How Skills-based Classroom Activities Shape Learners' Foreign Language Enjoyment: A Mixed-Modelling Longitudinal Examination

Alfaf Albakistani (Birkbeck College, University of London).

Existing research on foreign language emotions has focused on the relationship between foreign language enjoyment and various individual and contextual factors (Dewaele et al., 2021). Classroom activities have been discovered to be of particular importance (Dewaele & MacIntyre, 2014). Although researchers have explored the enjoyment of classroom activities (Boudreau et al., 2018; Dao & Sato, 2021; Shirvan & Taherian, 2018, 2021; Li & Xu, 2019; Pan & Zhang, 2021; Li et al., 2020), more is needed to know about how classroom activities shape FLE in an English as a foreign language (EFL) context. This study examines how the enjoyment of certain skill-based activities changes over time and what factors contribute to its variances within and among foreign language learners. It adopts a longitudinal mixed-method approach. Over nine months, repeated surveys were employed to track the skill-related enjoyment of 160 EFL adolescent learners from three grades in a Saudi secondary school. The survey included items for rating the enjoyment of speaking, reading, listening, and writing activities, as well as items for assessing the amount of four factors linked with each activity: collaboration, control, creativity, and authenticity. Four classroom observations, eight stimulated recall interviews, and ten semi-structured interviews were conducted. A repeated analysis of variance (ANOVA) was performed to investigate the differences in skill-specific enjoyment on a particular occasion. The primary statistical analysis was conducted via linear mixed-effects models (LMMs) by constructing random intercept and slope models. The ANOVA results showed significant differences in skill-specific enjoyment on the first occasion. LMMs revealed that only speaking enjoyment increased significantly over time while reading, listening, and writing enjoyment remained stable. Interestingly, intra-individual enjoyment of the four skills increased significantly over time. Moreover, learners' initial levels of the four skill-specific enjoyment varied considerably. Hence, while learners' enjoyment of speaking and listening continued to diverge uniquely, their enjoyment of reading and writing became relatively consistent with the group patterns. At the intra-individual level, collaboration was predictive of speaking enjoyment, creativity predicted speaking and reading enjoyment, whereas control contributed to writing enjoyment. At the inter-learner level, collaboration significantly contributed to the enjoyment of speaking, listening, and writing, while creativity predicted just speaking enjoyment. The enjoyment of skill-related activities was unaffected by authenticity. Both statistical and thematic findings suggest that certain features of the skill-based activities and other individual and contextual factors positively impact learners' enjoyment. Different learners rely on distinct factors, and so does their enjoyment of a specific skill. This study adds considerably to future teaching methods in developing classroom activities with positive features that lead to activity enjoyment.

L2 Lexical Attrition in Receptive and Productive Knowledge of Collocations among Bilingual Saudi Arabic-English Returnees and Saudi Heritage Speakers in the United States

Hadil Alraddadi (Reading University) and Jeanine Treffers-Daller (University of Reading).

This study analyses the impact of the L1 on processing of L2 verb-noun collocations among returnees and heritage speakers. As noted by Schmid and Köpke (2009), the most sensitive component of the linguistic system is the lexicon. The available empirical research to date investigating attrition in lexical knowledge is scarce, particularly attrition in collocational knowledge (Schmitt, 2010). One of the few available studies (Kopotev, Kisselev, & Polinsky, 2020) suggests that heritage speakers (HSs) use transfer-based non-standard word combinations, and that analysing such combinations can throw new light on the role of input in HSs language development. Here we aim to contribute to this discussion in a study of English L2 verb-noun collocations among different groups of Arabic HSs and returnees who have received very little attention in the literature. Participants were 44 child and adolescent returnees who had lived in the US for an extended period of time and returned to their country of origin, Saudi Arabia, either in early childhood or in adolescence. They were compared to 54 US-based child and adolescent Saudi HSs and a group of 20 adult HSs, as a control group. Productive knowledge was measured with a novel online gap-filling task consisting of English verb-noun collocations which were either congruent, such as 'have experience' or incongruent between Arabic and English, for example, 'do homework'. Furthermore, an online lexical decision task (LDT) focusing on verb-noun collocations was used to measure receptive collocational knowledge, and a range of baseline tests were administered to test vocabulary and grammar knowledge. Both online tasks were designed in PsychoPy to measure accuracy and reaction time (RT). It was predicted that HSs would achieve higher accuracy scores and experience less Arabic influence on all productive and receptive tasks and would process English collocations faster than returnees, as the HSs are exposed to more and a richer input in English than the Saudi returnees. Generalized linear mixed effect modelling was used for RT and linear mixed effect modelling for accuracy among all 118 participants. This revealed that returnees obtain significantly lower scores than both groups of HSs on the LDT and the gap-filling task. In terms of RT, child and adolescent HSs perform at similar levels as adult HSs on the LDT despite being younger and less experienced. The findings indicate a strong influence of Arabic among returnees, possibly L2 attrition after return. However, accuracy data show that child and adolescent HSs score significantly lower than adult HSs. This significance seems to be the result of incomplete acquisition rather than attrition among returnees. We finish by formulating implications for the role of input and attrition in HSs and returnees.

References

- Kopotev, M., Kisselev, O., & Polinsky, M. (2020). Collocations and near-native competence: Lexical strategies of heritage speakers of Russian. *International Journal of Bilingualism*.
- Schmid, M. S., & Köpke, B. (2009). L1 attrition and the mental lexicon. In A. Pavlenko, *The bilingual mental lexicon: Interdisciplinary approaches* (pp. 209-238). Bristol, England: Multilingual Matters.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Basingstoke, UK: Palgrave Macmillan.

Bidialectalism and Executive Control in Older Adults

Kyriakos Antoniou (Cyprus University of Technology) and Kakia Petinou (Cyprus University of Technology).

Some (controversial) evidence suggests that speaking two languages (bilingualism) or two dialects of the same language (bidialectalism) enhances executive control—a set of attention and memory-related processes, such as working memory (WM) and switching [1]. Bilingualism is thought to enhance executive control (EC) because bilinguals presumably use EC during everyday communication to manage their simultaneously active languages in the mind and brain [1]. EC declines with advanced age. Importantly, the positive bilingual effects have been argued to mitigate cognitive deterioration with ageing and to delay neurodegenerative diseases, such as dementia [1]. Here, we present preliminary results from research examining bidialectalism as an experience that potentially preserves EC in older adults. Thirty-one Greek-speaking bidialectal (in Cypriot and Standard Modern Greek) older adults (>50 years old) were given the Number-Letter (NL) [e.g., 2] and Color-Shape (CS) switching tasks [e.g., 3]; and the Forward and Backward Corsi WM task [e.g., 4]. Their EC was compared to 50 multilingual (in Cypriot, Standard Modern Greek, and other languages), 78 bidialectal, and 51 monolingual (in Standard Modern Greek) young adults (<50 years old) from previous work, which showed some evidence for a multilingual and bidialectal EC benefit. From switching tasks, we used the mixing cost (reaction-time difference between repeat trials in mixed and repeat trials in pure blocks), which consistently shows lower performance in older than young adults [e.g., 5]. We formed composite scores from related variables by averaging relevant z-transformed measures (in parentheses): Mixing (reverse-scored mixing costs in NL and CS tasks) and WM (forward and backward score in Corsi). We predicted that bidialectal older adults will exhibit less EC decline compared to monolingual than compared to multilingual and bidialectal young adults. An Analysis of Variance with Group and EC as factors showed a significant Group effect ($F(3, 206)=2.75, p<.05$): older bidialectals had lower overall EC than younger multilinguals (contrast estimate=0.34, SE=0.15, $p<.05$) and bidialectals (contrast estimate=0.3, SE=0.14, $p<.05$) but not than monolinguals (estimate=0.1, SE=0.15, $p>.05$). We consider the possibility that bidialectalism mitigates ageing-related EC decline, in the context of limitations of this investigation at this stage (e.g., the absence of a monolingual older adult comparison group).

[1] Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. *Psychological bulletin*, 143, 233. [2] Karayanidis, F., Jamadar, S., Ruge, H., Phillips, N., Heathcote, A., & Forstmann, B. U. (2010). Advance preparation in task-switching: converging evidence from behavioral, brain activation, and model-based approaches. *Frontiers in Psychology*, 1, 25. [3] Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General*, 137, 201. [4] Mueller, S. T., & Piper, B. J. (2014). The Psychology Experiment Building Language (PEBL) and PEBL test battery. *Journal of Neuroscience Methods*, 222, 250-259 [5] Gajewski, P. D., Ferdinand, N. K., Kray, J., & Falkenstein, M. (2018). Understanding sources of adult age differences in task switching: Evidence from behavioral and ERP studies. *Neuroscience & Biobehavioral Reviews*, 92, 255-275.

Do captioned-videos foster the learning of noun-noun constructions in primary school EFL students?

Daniela Avello (University of Barcelona) and Carmen Muñoz (University of Barcelona).

A Usage-based perspective for L2 learning emphasizes that students are sensitive to regularities in the input (Ellis, 2002). Therefore, target language knowledge is gradually acquired in the form of linguistic constructions of different levels of abstraction and complexity where vocabulary and grammar are barely detached (e.g. Tomasello, 2003). Overall, the evidence to date suggests that learners' exposure to captioned-videos fosters the learning of L2 grammar constructions (e.g. Pattemore & Muñoz, 2020). However, learners may only benefit from this activity under sustained exposure to audiovisual input and/or intentional learning conditions (Van Lommel et al., 2006). In addition, learner-viewers' outcomes may also depend on the complexity of the target language constructions under study and learners' L2 knowledge (Cintrón-Valentín et al., 2019). Yet, in comparison with vocabulary studies, few investigations on captioned-video viewing have focused on the learning of L2 grammar constructions, and even less so have explored primary school students' outcomes. The present investigation studied the extent to which 112 EFL primary school students (9-11 years old; 4th-5th grade) learned specific noun+noun (N+N) constructions encountered in 11 captioned-videos, and the extent to which they were able to figure out and generalize the pattern to items that were not presented in the treatment. We selected English N+N constructions because of their difficulty for L1-Spanish speakers and because they are not formally taught in primary school. Additionally, we aimed at examining the effects of two types of after-viewing activity: meaning-focused (incidental learning condition) or construction-focused (intentional learning condition). Participants were randomly assigned to one of them or to a control group that did not view the videos. The comparison of a written grammaticality judgment test at pretest and posttest revealed that the two experimental groups showed greater progress than the control group, but only those who completed construction-focused activities obtained significant gains from the treatment (grades 4 and 5). In addition, within the construction-focused group, only fifth graders were able to generalize the pattern to new items. This finding may be attributed to fifth graders' higher proficiency level and possibly, to their stage of cognitive development. All in all, the results suggest that primary school learners may need even longer treatments than adults and/or the support of explicit instruction to learn target grammatical constructions.

Cintrón-Valentín, M., García-Amaya, L., & Ellis, N. (2019). Captioning and grammar learning in the L2 Spanish classroom. *Language Learning Journal*, 47(4), 1-21.

Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24, 143– 188.

Pattemore, A. & Muñoz, C. (2020). Learning L2 constructions from captioned audio-visual exposure: The effect of learner-related factors. *System*, 93.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Harvard University Press.

Van Lommel, S., Laenen, A., & d'Ydewalle, G. (2006). Foreign grammar acquisition while watching subtitled television programs. *British Journal of Educational Psychology*, 76, 243–258.

Cross-linguistic (dis)similarity and L2 acquisition as predictors of L1 morphosyntactic attrition

Lewis Baker (University of Southampton).

Grammatical attrition refers to restructuring of L1 morphosyntactic representations due to a change in a speaker's primary linguistic data (PLD) in adulthood. This study reports findings relating to two predictions derived from a recent model of grammatical attrition: the Attrition via Acquisition (AvA) Model (Hicks and Domínguez, 2020). The first prediction is that attrition is facilitated in cases of L1-L2 cross-linguistic similarity. The second is that L2 acquisition of a grammatical structure is a pre-requisite to its L1 attrition.

Previous research provides mixed results as to the role of cross-linguistic (dis)similarity in attrition, and few studies systematically test this factor through use of multiple language pairs (Schmid and De Leeuw, 2019:187). The relationship between L2 acquisition and L1 attrition is likewise not well understood as few attrition studies also test potential attriters' acquisition of the equivalent L2 grammatical structures. Some that do find, perhaps surprisingly, L1 attrition in the absence of successful L2 acquisition (Ribbert and Kuiken, 2010).

To investigate these factors, this study tested three groups of L1 German late-sequential bilinguals with minimum 15 years residence in either the Netherlands (n=29), the UK (n=33), or Spain (n=32), starting in adulthood. In addition to manipulating the holistic similarity of the L1-L2 pairs, in each pair two grammatical properties which differ in their relative structural similarity were investigated. In line with the Minimalist assumptions of the AvA model, structural similarity is formalised in terms of the degree of Feature Reassembly (Lardiere, 2009) required for the L2 acquisition and L1 attrition of a structure. Predictions regarding cross-linguistic (dis)similarity are outlined in the table below:

Properties were tested by means of bimodal Acceptability Judgement Tasks (AJTs). There were German versions and equivalents in the corresponding L2s. Potential attriters completed both L1 and L2 versions of the relevant AJTs.

Results reveal no statistically significant differences between the potential attriter groups and an L1 German control group on any of the structures. However, results from several individual participants suggest attrition of verb position and adjective agreement. Consideration of these individuals' results on the L1 and equivalent L2 tasks reveals some instances of L1 attrition in the absence of fully successful L2 acquisition. At a group level, potential attriters' L2 results differ significantly from native controls only on negation structures. The implications of these findings are considered both with reference to the AvA model and for modelling grammatical attrition more generally.

References: Hicks, G. and Domínguez, L. (2020) 'A model for grammatical attrition', Keynote Article in *Second Language Research*, 36(2), pp. 143–165. doi: 10.177/0267658319862011.

Lardiere, D. (2009) 'Some thoughts on the contrastive analysis of features in second language acquisition', *Second Language Research*, 25(2), pp. 173–227. doi: 10.1177/0267658308100283.

Ribbert, A. and Kuiken, F. (2010) 'L2-induced changes in the L1 of Germans living in the Netherlands', *Bilingualism: Language and Cognition*, 13(1), pp. 41–48. doi: 10.1017/S1366728909990320.

Schmid, M. S. and De Leeuw, E. (2019) 'Introduction to Linguistic Factors in Language Acquisition', in *The Oxford Handbook of Language Attrition*. Oxford: Oxford University Press, pp. 181–190.

Feature redeployment in L2 vowel perception: how important is the presence of an active feature in the L1?

Fernanda Barrientos (University of Konstanz).

In feature-based approaches to L2 phonology (e.g. Brown 1998), the acquisition of new segments is facilitated by the presence of the relevant phonological features in the L1, even when they are in different segments; thus, acquiring a new phonological feature should be more difficult than redeploying specified L1 features in order to create new representations. However, studies suggest that (a) new features can be learned e.g. (Escudero & Boersma 2004) and (b) that having the feature in the L1 does not necessarily lead to an advantage over learning a new feature (Barrios et al. 2016, Bohn 1995). Since Spanish does not have a [+/-tense] feature, and [+/-round] is redundant with backness (i.e. underspecified feature), while English has [+/-tense], and also a [+/-round] contrast in back vowels (specified), we hypothesize that the English group may show better sensitivity (d') and higher accuracy in identification for both contrasts than the Spanish group, which should perform equally poorly with both contrasts. We tested discrimination (AX task with minimal pairs, ISI= 1500 ms) and identification (picture selection, 2AFC) of two vowel contrasts: /i-ɪ/ (e.g. *Miete/Mitte*, where [+/-tense] is a new feature in the Spanish group) and /y-u/ (e.g. *spülen/Spulen*, where L1 feature [+/-round] is expected to redeploy to a front vowel by the L1 English group). Participants in both groups were at B2 level or above in German; instructions were given completely in German. Preliminary results showed mean d' of 4.30 for /i-ɪ/ and 3.51 for /y-u/ in the Spanish group (n=15); values for the English group (n=9) were 3.57 and 3.45, respectively. A mixed-effects model on correct/incorrect counts for the identification task, with vowel contrast and L1 group as predictor and subject as random effect ($\chi^2=14.06$, $df=2$, $p<0.001$) showed an interaction between variables. Post-hoc analyses showed that the L1 English group performs significantly better with the /i-ɪ/ contrast (Est=1.61, SE=0.329, $p<.0001$); the L1 Spanish group performed equally poorly for both contrasts ($p=.8$). These results suggest that experienced L2 learners are more successful when discriminating between sounds in a feature acquisition scenario than in redeployment, and that feature specification does not play a role. The differences in performance between tasks suggests that L2 phonological acquisition is likely to take place at a surface level and favours learning through attunement to salient acoustic cues (where duration helps with the /i-ɪ/ contrast in L1 Spanish speakers) over internal rearrangement of features.

References Barrios, S., Jiang, N., & Idsardi, W.J. (2016). Similarity in L2 phonology: Evidence from L1 Spanish late-learners' perception and lexical representation of English vowel contrasts. *Second Language Research*, 32(3), 367–395. Brown, C. (1998). The role of the L1 grammar in the L2 acquisition of segmental structure. *Second Language Research*, 14(2), 136-193. Escudero, P. & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition* 26. 551–585.

L2 Writing and NLP Tools - Raising Awareness of the Importance of Preliminary Analysis of Writing Output

Tobias Becker (RWTH Aachen University).

NLP tools (e.g. Kyle, Crossley, & Berger, 2018; Lu, 2010) have granted invaluable insights into linguistic complexity in L2 research using the various measures and indices available. In recent years, some measures have seen criticism for their misplaced application, for instance in certain target registers (Kyle & Crossley 2018). Choosing appropriate analyses thus is crucial. Nevertheless, this study proposes that another prior step be included when conducting NLP tool research on L2 data:

Learner corpora can include multifarious, unique written data, often resulting from specific language tasks. Higher proficiency writing may include source material needed to complete these tasks. German secondary school English exams are a prototypical example of such text output, generally consisting of three language tasks (1. descriptive, 2. analytic, 3. argumentative or creative) revolving around one source text (political speech, novel excerpt...). The influence of this text may have a considerable impact on the student writing, and consequently on linguistic complexity. Moreover, language mistakes (e.g. orthography) might cause variable results in automatic assessments. Focusing on lexical sophistication, this study addresses a dyadic research question: Do alterations of L2 writing regarding (a) language variation and (b) source-text impact yield divergent results when applied prior to the analysis of the indices word frequency (WF) and word range (WR)?

1979 texts resulting from tasks in hand-written curriculum-based examinations by 606 students in the final three years at three German secondary schools were collected, transcribed, and provided with xml-mark-up. This mark-up includes detailed tags for language variation (gaps and misspelled, surplus, and foreign words) and content-related features (quotes, names, dates, and references). Four versions of the above texts (original and altered, i.e. excluding the tagged words, for both language variation and content features) were assessed using the Tool for the Automatic Analysis of Lexical Sophistication (Kyle et al., 2018) and its WF and WR indices of the written sub-sections of the Corpus of Contemporary American English.

Statistical analyses using multiple regression revealed that (1) language variation did not yield significant results for the selected indices, only accounting for less than 1% of the variance in the data, yet (2) content-related features showed highly significant results ($p < 0.000$) in all indices, accounting for up to 8% of the variance in the versions (original and altered) and up to 17% when factoring in the different task types.

The implications of these findings for research on automated L2 writing assessment are discussed. Particularly, intra-corporal comparisons, comparisons across (learner) corpora and the importance of the nature of the writing task are highlighted.

Kyle, K., & Crossley, S. (2018). Measuring Syntactic Complexity in L2 Writing Using Fine-Grained Clausal and Phrasal Indices. *Modern Language Journal*, 102(2), 333–349. Kyle, K., Crossley, S., & Berger, C. (2018). The tool for the automatic analysis of lexical sophistication (TAALES): Version 2.0. *Behavior Research Methods*, 50(3), 1030-1046. Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, 15(4), 474-496.

Classification of written texts: Identifying first languages and writing conditions with machine learning algorithms

Raphael Berthele (Institut de Plurilinguisme).

Linguistic structure in first and second/foreign languages constrains positive and negative transfer (Jarvis & Pavlenko, 2007) as well as the pace of additional language learning (Schepens, 2014). A relatively recent take on language transfer is the detection-based approach that involves the use of computer classifiers to identify the first languages of the authors of second language texts (Jarvis & Crossley, 2012).

In this contribution, a total of 472 texts written by Swiss learners of English as a foreign language are analysed. All learners follow a vocational training program (age between 16 and 21) in either a French-speaking (169 texts) or in a German-speaking (303 texts) curriculum. The learners had the task to write a professional letter to a potential employer with the goal of providing some biographic information. The learners wrote two similar texts, a first one without any help from digital translation tools, and a second one either with or again without the use of tools.

The goal of this paper is to investigate to what extent it is possible to automatically categorize the texts with respect to the dominant language of their authors and with respect to the condition in which the texts were written (online/offline).

To this end, the text corpus is converted in a document term matrix and fed into machine learning algorithms. Both naive Bayesian and logarithmic models provided by a Python library are used (Pedregosa et al., 2011). The algorithm trains on 75% of the texts and is tested on the remaining 25% texts.

The performance of the classifier is assessed regarding its ability to separate the signal from the noise for both characteristics (language, online/offline condition). The results show that the dominant language can be identified with a very high probability (the AUC metric for the Bayesian model is 0.99, 1.00 being the 'perfect' performance). The performance is somewhat worse but still rather good for the online vs. offline conditions (AUC 0.90).

The linguistic features (unigrams and bigrams) that are most strongly associated with the categories are discussed. The specific use of a few cognates and false friends are the distinctive features that mark the dominant language of the authors. The online condition, on the other hand, is characterized by the use of specific low-frequency words supplied by the online translators.

Jarvis, S., & Crossley, S. A. (Eds.). (2012). *Approaching Language Transfer through Text Classification: Explorations in the Detection- based Approach*. Multilingual Matters. Jarvis, S., & Pavlenko, A. (2007). *Crosslinguistic Influence in Language and Cognition*. Lawrence Erlbaum Associates. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12(85), 2825–2830. Schepens, J. (2014). *Bridging Linguistic Gaps: The Effects of Linguistic Distance on the Adult Learnability of Dutch as an Additional Language*. Utrecht University.

Implicational patterns in the L2 acquisition of German verb inflection

Anja Binanzer (Leibniz Universität Hannover), Carolin Hagemeyer (Leibniz Universität Hannover) and Verena Wecker (Westfälische Wilhelms-Universität Münster).

German verb inflection is often divided into two inflectional subclasses: the weak (regular) and the strong (irregular) inflection class. Since not all strong verbs show strong inflectional features (mainly vowel changes) in all paradigm positions, variety within strong verb inflection may be a reason for increasing the learning task's difficulty (for English see Bybee & Slobin 1982). Bittner (1996) demonstrates that the occurrence of strong features is not arbitrary but can be predicted by an implicational pattern: imperative < present < past tense < subjunctive < past participle. If a verb's imperative shows a strong feature, all other forms will show strong features as well. If a strong feature occurs in the past tense, it will also occur in the subjunctive and the past participle, but not necessarily in the present and imperative. This implicational pattern was shown to play a role in language change (Bittner 1996) and in advanced L1 acquisition (Bittner & Köpcke 2007), but has not been examined yet regarding its impact on L2 acquisition. Hence, the presented study explores whether adult L2 learners of German do store resp. organize strong verb inflection in their mental lexicon according to Bittner's implicational pattern. We conducted two nonce-verb-experiments with learners of German as a foreign language (N=167), aged 17–40, with different L1s, who attended German language preparatory courses for university studies in Germany. We chose learners with different proficiency levels (A2 and B2 of the CEFR) to be able to detect developmental patterns and conducted the experiments with a L1 German control group (N=83) as well. All participants were presented with nonce-verb-infinitives and a second verb form (condition one: paradigm position (present/past tense/past participle), condition two: inflection class (weak/strong)) within a sentence. In experiment 1, learners had to complete a cloze test with two paradigm positions missing, whereas in experiment 2, they had to choose between a strong and a weak option for two different paradigm positions (grammaticality judgement). Results show that - participants in both experiments produced/chose strong forms significantly more often if the presented form showed strong features; - data of experiment 1 (cloze test) corresponds widely to Bittner's implicational pattern, i.e. if a strong form was given, participants created strong forms significantly more often for the paradigm position that should show strong features than for the paradigm position that does not necessarily have to show strong features; - in experiment 2 (grammaticality judgement) the effects for paradigm positions were less clear than in experiment 1.

The results will be discussed in reference to the framework of usage-based models of L2 acquisition of inflectional morphology and with regard to the task's level of metalinguistic awareness employed in the experiment.

References

Bittner, A. (1996): Starke "schwache" Verben - schwache "starke" Verben. Deutsche Verbflexion und Natürlichkeit. Tübingen: Stauffenburg. Bittner, A. & Köpcke, K.-M. (2007): Überlegungen zur Repräsentation grammatischen Wissens am Beispiel der Verbmorphologie des Deutschen. In: Di Meola, C. et al. (eds.): Perspektiven Zwei. Akten der 2. Tagung Deutsche Sprachwissenschaft in Italien (Rom, 9.-11.2.2006). Roma: Istituto Italiano di Studi Germanici, 3–15. Bybee, J. & Slobin, D. (1982): Rules and schemas in the development and use of the English past tense. *Language* 58-2, 265–289.

Learning New Collocations: The Effects of Grouping (Thematic versus Unrelated) and Language of Instruction (L2 versus L1 and L2)

Ronit Breslaw (Haifa University) and Batia Laufer (Haifa University).

Second language learners have difficulties with mastering collocations, particularly for productive purposes. Even advanced learners tend to overuse, or underuse some collocations, and misuse L1-L2 incongruent collocations (Laufer & Waldman, 2011; Paquot & Granger, 2012; Yamashita & Jiang, 2010). Research shows that practicing collocation in form-focused activities leads to better learning results than exposure in input, particularly with L1-L2 incongruent collocations (Laufer & Girsai, 2008; Nguyen & Webb, 2017; Peters, 2012, 2016). We investigate how learning new collocations is affected by different combinations of two teaching methods: Mode of grouping (in thematically related, or unrelated collocation clusters) and language of instruction (L2 only, or L1-L2 contrastive analysis). Though studies on clustering were conducted with single words (Tinkham, 1997), to our knowledge, there are no studies on collocations that investigated clustering alone, or clustering with other variables. Thirty seven EFL College learners were taught twenty new incongruent adjective-noun collocations. Ten were thematically related; ten were unrelated. Ten (five from each cluster type) were taught in English only, the other ten were taught via L1-L2 contrastive analysis. In a counterbalanced within-subject design, each student studied different collocations in four conditions: Thematic/L2, thematic/L1-L2, unrelated/L2, unrelated/L1-L2. Collocations were practiced in four form-focused activities. Learning was assessed by immediate and delayed posttests of production (recall of form for given meanings) and comprehension (recall of meaning) of target collocations. Data were analyzed by 2 (grouping) x 2 (language of instruction) Repeated Measures and post-hoc pair comparisons. Results showed effects for the two main variables. The 'unrelated/L1-L2' condition yielded the best results and 'thematic /L2' fared worst. We discuss the results in light of the interference theory in collocation clusters, L1-L2 awareness raising, and the effectiveness of form-focused instruction.

References

- Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning* 61, 647–672.
- Laufer, B., & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation. *Applied Linguistics*, 29, 694-716.
- Nguyen, T. M. H., & Webb, S. (2017). Examining second language receptive knowledge of collocation and factors that affect learning. *Language Teaching Research*, 21, 298-320.
- Paquot, M., & Granger, S. (2012). Formulaic language in learner corpora. *Annual Review of Applied Linguistics*, 32, 130-149.
- Peters, E. (2012). The differential effects of two vocabulary instruction methods on EFL word learning: A study into task effectiveness. *International Review of Applied Linguistics in Language Teaching*, 50, 213-238.
- Peters, E. (2016). The learning burden of collocations: The role of interlexical and intralexical factors. *Language Teaching Research*, 20, 113-138.
- Tinkham, T. (1997). The effects of semantic and thematic clustering on the learning of second language vocabulary. *Second Language Research*, 13, 138–163
- Yamashita, J. & Jiang, N. (2010). L1 influence on the acquisition of L2 collocations: Japanese ESL users and EFL learners acquiring English collocations. *TESOL Quarterly*, 44, 647-668.

The relationship between productive knowledge of L2 collocations and proficiency in listening, reading, speaking and writing

Dale Brown (Kanazawa University) and David Coulson (Ritsumeikan University).

Collocations, pairs of words that frequently co-occur, are ubiquitous in language use and crucial to fluency and accuracy. They are, however, “notoriously difficult for learners” (Paquot & Granger, 2012, p. 130). Intuitively, a relationship between productive collocations knowledge and overall proficiency can be anticipated; a beginner would be expected to have very little knowledge of collocations, an advanced learner to have much more, and there must be progression between these two points. As yet, however, the evidence is somewhat mixed: while Barfield (2009), González Fernández and Schmitt (2015), Nizonkiza (2017), Revier (2009) and Smith (2021) found a clear, positive relationship, Howarth (1998) and Nesselhauf (2005) did not. Moreover, there are methodological issues with the above studies pertaining to the reliable identification of collocations, knowledge of single words confounding measurement of collocation knowledge, obtaining appropriate writing samples (in learner corpus studies), and task design (in elicitation studies).

Brown (2018) developed a new tool for measuring productive knowledge of collocations, called LexCombi 2, which addresses some of the above issues. Extending work by Barfield (2009), LexCombi 2 was developed through a series of studies which trialled and evaluated a new task design, new scoring procedures and new cues. Brown found that LexCombi 2 appears to be a measure of a single construct, produces data with high reliability and can be used with learners ranging widely in ability. LexCombi 2 presents learners with high-frequency noun cues and asks for three collocates for each cue, with a 30-second time limit per cue. Learners’ responses are then scored against a list of canonical collocates for each cue.

Initial findings using LexCombi 2 with 146 undergraduates from four universities in Japan showed there to be no relationship between productive knowledge of collocations and overall L2 proficiency. It was suggested that this finding could stem from these learners’ concentration on receptive skills (and neglect of productive skills) due to the primary role of English in the Japanese education system as an object of study to be tested in high-stakes examinations rather than as a tool for communication.

This presentation will therefore describe a new study in which 32 Japanese undergraduate learners who had taken either the IELTS or TOEFL test and so were able to provide scores for reading, writing, listening and speaking skills completed LexCombi 2. It was found that the strongest correlation with these learners’ LexCombi 2 scores was for listening skills (.50), followed by reading skills (.47), speaking skills (.32) and writing skills (.26). In addition, the correlation of LexCombi 2 with overall proficiency test scores was .53. Thus, in contrast with our expectations, the correlations for receptive skills were higher than those for productive skills, while the overall LexCombi 2 – proficiency relationship was much stronger than that reported in Brown (2018). The presentation will explore possible explanations for these results.

Input quality moderates the effect of socio-economic status on young L2 learners' phonological awareness

Ann-Christin Bruhn, Alicia Strompen, Kristin Kersten

Children's linguistic and cognitive development is shaped by their previous experiences and the social environment in which they grow up. Phonological awareness (PA) at the beginning of school is an essential predictor of later literary skills (Schnitzler, 2009) for both L1 and L2 contexts, as well as academic and life success. In addition, families' socio-economic status (SES) is another predictor of language and cognition, with children from higher SES backgrounds demonstrating higher PA (McDowell et al., 2007). It is therefore important for early instructional programs to create a learning environment that promotes the development of cognitive and (multi-)linguistic skills of students from diverse social backgrounds with heterogeneous abilities.

Numerous studies have identified L2 input quality as a factor that plays an important role in language acquisition. Truscott & Sharwood Smith (2019) suggest that a comprehensive definition of input quality in SLA should not only target linguistic input but the sensory input from the entire external context, including anything that promotes the interpretation of utterances and fosters SLA. Accordingly, we operationalise L2 input in terms of 41 items referring to multisensory task characteristics, verbal and non-verbal input scaffolds and feedback in the Teacher Input Observation Scheme (TIOS, consistency: $\alpha=.905$; Kersten et al., in press). Previous studies have shown that input quality operationalised that way moderates the effect of SES on learners' linguistic skills in both L1 and L2 (Kersten et al., in press). Research on compensatory effects of input on cognitive skills, however, is scarce.

Hence, the purpose of this cross-sectional study is to investigate whether L2 teacher input also has a moderating effect on the influence of SES on the PA of young L2 learners, and if so, whether it is buffering or strengthening. PA scores were obtained from N=65 fourth graders using a standardized PA-test for the learners' L1, German (BAKO 1-4). 20 teaching videos from their respective L2 teachers were rated by two raters using the TIOS (IRR: Krippendorff's $\alpha=.793$). A parental questionnaire elicited SES (represented as HISEI).

Results from regression and moderation analyses revealed that SES significantly predicted PA and that input quality (the full score as well as the subscale 'non-verbal input') moderated the effect of SES on PA. However, in contrast to the buffering effect observed previously on linguistic skills (Kersten et al., in prep), we find a strengthening effect. This indicates that high SES learners in particular seem to benefit from high quality input. Results will be discussed with regard to learners' linguistic and social backgrounds.

Kersten, K., Blackman, A., Funke, F.L., Akram, T., Kliebisch, M., Koch, M.J. (in press). Input quality affects L1 and L2 proficiency and moderates the effect of socio-economic status. In H. Böttger, N. Schlüter (eds.), *Fortschritte im frühen Fremdsprachenlernen*. Schriftbild.

McDowell, K.D., Lonigan, C.J., Goldstein, H. (2007). Relations among socio-economic status, age, and predictors of phonological awareness. *Journal of Speech, Language, and Hearing Research*, 50(4), 1079–1092.

Schnitzler, C.D. (2008). *Phonologische Bewusstheit und Schriftspracherwerb*. Thieme.

Truscott, J., Sharwood Smith, M. (2019). The internal context of bilingual processing. John Benjamins.

Acquiring lexical patterns from the input: how quickly do learners see things in "black and white"?

Gareth Carrol (University of Birmingham), Suhad Sonbul (Umm Al-Qura University, Saudi Arabia), Dina El-Dakhs (Prince Sultan Univeristy,) and Kathy Conklin (University of Nottingham).

Usage-based accounts emphasise exposure to language patterns as a key part of both first and second language development. Amongst these, literal binomials ("X and Y" phrases with a strongly preferred order such as "brother and sister" or "salt and pepper") represent an interesting test case, since word order can (in principle) be reversed with no change in meaning (albeit with a change in emphasis and/or markedness). Amongst the range of "constraints" that have been proposed (e.g. semantic, phonological, cultural), frequency and conventionalisation seem to be key in maintaining the preferred order in language at large (Morgan & Levy, 2016), and driving the processing advantage that has been observed (e.g. Carrol & Conklin, 2020).

Here, we explore the learning of previously unseen binomials, and compare how quickly first and second language readers acquire novel patterns through reading. We build on previous work for first language speakers (Conklin & Carrol, 2020), which found that novel binomials with no established order (e.g. "wires and pipes") were read more quickly after as few as 4-5 exposures, with, crucially, a subsequent cost for encountering the reversed form (i.e. once "wires and pipes" has been seen several times, "pipes and wires" was read more slowly). In an initial replication with L2 English learners with Arabic as their L1 (Sonbul et al., 2022), we observed a general decrease in reading speeds as novel binomials were encountered more often, but no evidence that the reversed form was subsequently treated as "dispreferred".

In a second study we explore the effect for second language learners in more detail. We selected items where no existing pattern existed in either English or Arabic, and created new texts where items were shown either once or six times over the course of several short stories, followed by a reversed form. We also controlled the distribution of items such that an item seen six times was seen twice in successive sentences on two occasions, to increase likelihood of learners registering novel patterns. In addition to the reading data, we also collected offline measures (ratings of perceived reversibility). We discuss overall results in terms of their relevance for usage-based accounts of formulaic language acquisition, and vocabulary learning in general.

References Carrol, G. & Conklin, K. (2020). Is all formulaic language created equal? Unpacking the processing advantage for different types of formulaic sequences. *Language and Speech*, 63, 1, 95-122. <https://doi.org/10.1177/0023830918823230> Conklin, K. & Carrol, G. (2020). Words go together like 'bread and butter': The rapid, automatic acquisition of lexical patterns. *Applied Linguistics*, 42, 3, 492–513. <https://doi.org/10.1093/applin/amaa034> Morgan, E. & Levy, R. (2016).

knowledge versus direct experience in processing of binomial expressions. *Cognition* 157, 384–402. <https://doi.org/10.1016/j.cognition.2016.09.011> Sonbul, S., El-Dakhs, D., Conklin, K. & Carrol, G. (2022). "Bread and butter" or "butter and bread"? Nonnatives' processing of novel lexical patterns in context. *Studies in Second Language Acquisition, First View*, 1-23. <https://doi.org/10.1017/S0272263122000237>

Is it in the pausing? Effects of spoken L2 complexity, accuracy, and fluency (CAF) on communicative adequacy

Cirsten Carlson (University of Hildesheim).

A core goal of learning any L2 is communicative adequacy, which refers to how well a speaker accomplishes a task's goal (Pallotti, 2009). While complexity, accuracy, and fluency (CAF) have become viable and recognized dimensions of L2 performance, few studies have included communicative adequacy in L2 performance (e.g. Kuiken et al., 2010; Révész et al., 2016; Vasylets et al., 2020). This study takes communicative adequacy into account while assessing L2 speaking and investigates how the CAF factors influence communicative adequacy: How do complexity, accuracy, and fluency in speech production affect perceived communicative adequacy? Spoken language samples of eleven L2 English speakers of a non-English speaking country were gathered in monologic parts of semi-structured interviews (Carlson, 2020). In the study, the transcribed monologues were coded and analyzed, based on numerous CAF measures of syntactic complexity (subordination per analysis-of-speech unit (Foster et al., 2000), lexical diversity (VocD), and accuracy (error-free clauses). Fluency was measured in its sub-dimensions of speed (syllables), breakdown (number and length of pauses), and repair fluency (repetitions and self-repairs). Composite scores for each CAF dimension were calculated for each speaker. In order to rate the speakers' communicative adequacy, three L1 English speakers used a scale from 0 (unrelated response) to 7 (very successful response) on an adapted version of the communicative adequacy rating scale by Révész et al. (2016). Mean scores of communicative adequacy were calculated for each of the L2 speakers. Correlation and regression analyses of the speakers' CAF and communicative adequacy scores revealed significant correlations between several of the measured CAF scores and communicative adequacy, with breakdown fluency as a significant predictor of communicative adequacy. The results indicate that breakdown fluency may crucially influence how L2 speakers' communicative adequacy is perceived and rated by L1 speakers. The study may help to add to a more comprehensive picture of L2 performance and L2 speakers' communication demands by increasing our understanding of how each of the CAF factors influences L2 speakers' success in communicating – an understanding that could have vital implications for teaching L2 English as well.

References Carlson, C. (2020). Elementary School L2 English Teachers' Language Performance and Children's Second Language Acquisition. Hildesheim: Universitätsverlag Hildesheim. Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, 21(3), 354-375. Kuiken, F., Vedder, I., & Gilabert, R. (2010). Communicative adequacy and linguistic complexity in L2 writing. In Bartning, I., Martin, M., Vedder, I. (eds.). *Communicative proficiency and linguistic development: Intersections between SLA and language testing research*, 81-100. Pallotti, G. (2009). CAF: Defining, refining and differentiating constructs. *Applied Linguistics*, 30(4), 590-601. Révész, A., Ekiert, M., & Torgersen, E. N. (2016). The effects of complexity, accuracy, and fluency on communicative adequacy in oral task performance. *Applied Linguistics*, 37(6), 828-848. Vasylets, O., Gilabert, R., & Manchón, R. M. (2020). Task modality, communicative adequacy and CAF measures. In R. Manchón (ed.). *Writing and language learning: Advancing research agendas*. 183-205.

Definite and demonstrative descriptions in L2 acquisition of English by L1-Korean speakers: A modified replication of Ionin et al. (2012)

Vatcharit Chantajinda (University of Wisconsin-Madison), Hyun Bae (University of Wisconsin-Madison), Joseph Moran (University of Wisconsin-Madison), Eleanor Sands (University of Wisconsin-Madison) and Jacee Cho (University of Wisconsin-Madison).

Ionin et al. (2012) showed that, as Korean is article-less, L1-Korean speakers' interpretation of the L2-English definite article is influenced by transfer from the semantics of Korean demonstratives. Transfer effects in written production versus comprehension were compared. Results indicated that advanced L1-Korean L2-English speakers were targetlike in their use of the *vs.* *that* in a forced-choice elicited production task but nontargetlike in a picture comprehension task. Native speakers' interpretations of definite plurals such as *the pencils* as referring to 'the same, previously mentioned pencils' ('same') and as referring to 'all of the pencils in the picture' ('all' (maximality)) did not differ. This indicates: a) definite plurals in anaphoric contexts allow both 'same' and 'all' interpretations; and b) native speakers lack any strong preference for one interpretation over the other. L2 speakers, however, barely interpreted definite plurals maximally. They concluded that L2 speakers' nonpreference for maximality for definite plurals is due to transfer of demonstrative semantics onto the definite article. However, it is unclear whether L2 speakers simply prefer the 'same entity' interpretation over maximality in anaphoric contexts, as argued by Ionin et al., or whether L2 speakers lack the presupposition of maximality for definite plurals. To address this issue, the present study modified the picture comprehension task to test whether L2 speakers compute maximality for definite plurals. In Ionin et al. (2012), participants made drawings following instructions (e.g., Please draw triangles around the pencils), which does not disambiguate whether L2 speakers chose between two interpretations or whether they allowed only one interpretation. In this study's modified picture comprehension task, participants are shown pictures with drawings already made and asked to judge whether drawings are made accurately following given instructions on a 7-point Likert scale. We added a control condition (non-anaphoric contexts) where only the maximality interpretation is possible. Participants include native English speakers and L1-Korean L2-English speakers. Data collection is ongoing and will be completed by the conference date. Data will be analyzed using ordinal regression (cumulative link mixed effects) models. Preliminary data indicate that twelve advanced L1-Korean L2-English speakers correctly accepted drawings depicting maximality for definite plurals in non-anaphoric contexts. More important, L2 speakers incorrectly rejected drawings depicting 'all' for definite plurals in anaphoric contexts. The data suggest L2 speakers compute maximality in non-anaphoric contexts but not in anaphoric contexts, indicating the interplay between maximality and anaphoricity in L2 comprehension of English articles, especially definite plurals.

Reference Ionin, T., Baek, S., Kim, E., Ko, H., & Wexler, K. (2012). That's not so different from the: Definite and demonstrative descriptions in second language acquisition. *Second Language Research*, 28(1), 69-101.

The Foreign Language Effect (FLE) on moral judgment and the role of foreign language proficiency.

Nikki Maria Christofi, Panos Athanasopoulos and Aina Casaponsa.

Nikki-Maria Christofi, Aina Casaponsa, Panos Athanasopoulos Department of Linguistics and English Language, Lancaster University, UK.

Our choices depend on, or could be affected, by the language we use when making them. This is known as the Foreign Language Effect (the systematic influence of a foreign language on decision making; Dylman, & Champoux-Larsson, 2020). The precise reasons that the FLE occurs are still unclear. Second language proficiency, while hypothesized to underlie the processing effort account (disfluency requires more cognitive effort in processing information in a foreign language: Hayakawa et al., 2016), has so far not been rigorously measured (Circi et al., 2021). The current paper addresses this gap in our knowledge by conducting a systematic investigation of the role of L2 proficiency on the FLE, rigorously measured by means of a standard language proficiency test (CEFR level of English) and four Category and Letter fluency tasks.

107 Greek-English bilinguals were presented with four hypothetical moral dilemmas. Participants had to state (7-point scale) how possible it would be to commit the action in the dilemma, rate how emotionally upsetting they found each scenario and how well they have understood it. The dilemmas were presented in both languages in each questionnaire by using a language mode inducing task between each language, in a within-subjects design that eliminated external factors that could arise by comparing different groups of people as done in most previous research (Białek & Fugelsang, 2019)

Participants were asked to hypothetically make a utilitarian decision that will result in killing one person to save five-nine others. Moral dilemmas were divided in self-preservation SP (where participants will hypothetically save themselves along with others if they choose to commit the action) and non-self-preservation NSP (which do not involve saving one's self).

The results demonstrated a clear FLE when moral judgements were present in a FL consistent with the extensive previous research (Hayakawa et al. 2016). Greek-English bilinguals were more willing to commit the action of e.g. killing someone to save five-nine others in their FL than in their L1. But, contrary to previous research that found the FLE only in intermediate L2 users, here we show that FLE can present in both intermediate and advanced L2 users, with no differences as a function of proficiency and fluency. Such an effect has not been found before and does not support the assumption that the FLE occurs due to higher processing demands in the L2 because of low language proficiency.

References: Białek, M., Turpin, M., & Fugelsang, J. (2019). What Is the Right Question for Moral Psychology to Answer? Commentary on Bostyn, Sevenhant, and Roets (2018). *Psychological Science*, 30(9), 1383-1385. Circi, R. G. D. R. V., & Vecchi, T., (2021). The foreign language effect on decision-making: A meta-analysis. *Psychonomic Bulletin & Review*, Psychonomic bulletin & review, 2021-02-08. Dylman, A., & Champoux-Larsson, M. (2020). It's (not) all Greek to me: Boundaries of the foreign language effect. *Cognition*, 196, 104148. Hayakawa, S., Costa, A., Foucart, A., & Keysar, B. (2016). Using a Foreign Language Changes Our Choices. *Trends in Cognitive Sciences*, 20(11), 791-793.

The effects and the effectiveness of different types of instruction on the acquisition of L2 phonology: a meta-analysis

Bastien De Clercq (Vrije Universiteit Brussel), Francisco Miguel Valada (Vrije Universiteit Brussel, Universidade NOVA de Lisboa), Van Nhi Tran (Vrije Universiteit Brussel), Susana Correia (Universidade NOVA de Lisboa) and Alex Housen (Vrije Universiteit Brussel).

A substantial body of empirical studies on how (well) people learn a second/foreign language, culminating in a number of impactful meta-analyses (e.g. Spada & Tomita, 2010; Lee et al., 2015), now offers compelling evidence that instruction can have an effect on how (well) people learn a second/foreign language (L2). In an attempt to fill a number of gaps in the research literature we conducted a meta-analysis of primary studies on the effects of phonological instruction and try to elucidate outstanding issues such as the magnitude of the effects of implicit vs. explicit phonological instruction on both production and perception of phonological target features, and how these effects are mediated by the relative learning difficulty of those features, as predicted by theoretical models such as PAM-L2 (Best & Tyler, 2007).

We extracted 943 references from two databases (Web of Science and Scopus) and previous meta-analyses. 173 studies published between 1997 and 2022 were selected for more detailed coding. Using a set of five inclusion criteria targeting design and methodological characteristics (with pretest and posttest, control group, instructional treatment sessions, target segmental features, and statistical information to calculate effect sizes), 59 primary studies were included in a network meta-analysis to answer the following primary research question: to what extent does phonological instruction affect the acquisition of phonemic categories in an L2? Subsequently, we address four secondary research questions: (1) how is the impact of instruction moderated by the learning difficulty of the phonological target features?; (2) how is the impact of instruction on phonological acquisition moderated by type of instruction (implicit vs. explicit)?; (3) how does the impact of instruction depend on the type of modality (perception vs. production)?; (4) what is the interaction between type of instruction and target feature difficulty?

We discuss construct definitions, operationalisations, and coding of the independent and dependent variables, discuss the results of the meta-analysis as well as general methodological trends, strengths and shortcomings in research on phonological instruction as they emerge from our survey.

References

- Best, C.T., & Tyler, M.D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. Munro, & O.-S. Bohn (Eds.), *Second language speech learning* (pp. 13–34). Amsterdam, The Netherlands: John Benjamins.
- Lee, J., Jang, J., & Plonsky, L. (2015). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics*, 36(3), 345–366.
- Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning*, 60(2), 263–308.

Young learners' cognitive development, language preference and bilingual status in foreign language aptitude testing

M^a del Mar Suárez (Universitat de Barcelona).

Foreign language aptitude has been defined as the specific talent associated with learning a foreign or second language (Carroll, 1981; Skehan, 2002). Aptitude, specifically in young learners, might also be influenced by one's level of L1 mastery (Sparks & Ganschow, 2001), cognitive development (Suárez & Muñoz, 2011), bilingual status, as well as language learning experience (Bialystok et al., 2004). While the influence of the language-related variables mentioned could decrease if aptitude is measured using language-neutral aptitude tests (Rogers et al., 2017), it still remains to be seen if aptitude test scores are affected by the test takers' L1 preference and bilingual status when language-dependent aptitude tests are administered. Two such tests are the Modern Language Aptitude Test-Elementary in Catalan (MLAT-EC) (Suárez, 2010) and Spanish (MLAT-ES) (Stansfield et al., 2004). The objective of this study is, therefore, twofold: first, to test the validity, reliability and consistency of these two tests across bilingual and monolingual young learners and, second, to see how bilinguals compare to monolinguals in an aptitude test considering their cognitive development across ages 8-14.

629 Catalan/Spanish bilingual students from grades 3 to 7 took the MLAT-ES and the MLAT-EC for test comparison in a counterbalanced order. They were also asked which of their L1s they preferred (Catalan, Spanish or no preference of one language over the other). Likewise, their results were also compared to those of the monolingual pool whose results are reported in the MLAT-ES manual.

The results show that the bilinguals' performance on both tests presented hardly any significant differences considering students' L1 preference. In addition, these bilingual examinees outperformed the predominantly monolingual samples in the MLAT-ES norming study. However, the same score patterns related to young learner cognitive development stages were found across test versions. These results reinforce the confidence in the validity of the MLAT-E adaptations and support the hypothesis that bilingualism results in greater foreign language aptitude.

The Transferability of Processing Instruction Gains from Input to Output Tasks: the Case of L2-English Articles

Kateryna Derkach (University of Cambridge) and Theodora Alexopoulou (University of Cambridge).

Input Processing theory (VanPatten, 2015) suggests that learner difficulty with certain L2 features might originate in misprocessing them in the input (i.e. failing to create appropriate links between forms and meanings). Processing Instruction (PI) interventions, developed by VanPatten and colleagues, have been shown to be effective in improving learner accuracy on various features prone to processing issues (e.g. past tense marking, passive voice). This is achieved by modifying input to force processing the target feature for meaning (e.g. removing temporal adverbs to force interpreting “-ed” as a past tense marker).

However, most PI research shows improvement on input processing or sentence-level reconstruction tasks. Yet, it is unclear whether these gains transfer to free production. The little research addressing this issue (targeting L2-English passives) employed a passage reconstruction task (VanPatten & Uludag, 2011), justified by the potential scarcity of passives in free production. A better test case would be an abundant feature, such as articles (“a”/“the”), which could easily be assessed in any free production task.

English articles are notoriously difficult for L2-English learners, especially those whose L1s lack articles, [-art] (Murakami & Alexopoulou, 2016). We hypothesise that they should be amenable to PI, since it is possible that learners either do not process articles in the input at all or process them incorrectly associating them with, for example, referentiality instead of definiteness, as suggested by Ionin et al. (2004).

Thus, our research question is whether PI improves L2-English article accuracy not only in interpretation or sentence-level reconstruction tasks but also in free production.

The focus of this paper is on the quantitative and qualitative analysis of article use in the free production task from the pre-, post- and delayed post-tests of a three-week-long PI intervention targeting articles. The task contained a silent cartoon description and a short essay on a topical issue (e.g. Covid restrictions) and was completed by the seventy participants of the experiment (most between A2–B2 CEFR levels, half [-art], half [+art]). Preliminary analysis of a balanced subset (34 participants) indicates considerable accuracy gains in the Experimental group (ca. 10% on average). Importantly, [-art]-L1 participants appear to reduce article omission, the most pervasive type of error for such learners, although participants still struggle with choosing the right article or knowing when not to use one.

Further analysis and regression modelling is underway to confirm these preliminary findings and establish the impact of other important variables (noun countability, number, abstractness/concreteness, referentiality, presence of prenominal modifiers).

References Ionin, T., Ko, H., & Wexler, K. (2004). Article semantics in L2 acquisition: The role of specificity. *Language Acquisition*, 12(1), 3–69. Murakami, A., & Alexopoulou, T. (2016). L1 influence on the acquisition order of English grammatical morphemes: a learner corpus study. *Studies in Second Language Acquisition*, 38, 365–401. VanPatten, B. (2015). Foundations of processing instruction. In *IRAL - International Review of Applied Linguistics in Language Teaching* (Vol. 53, Issue 2, pp. 91–109). VanPatten, B., & Uludag, O. (2011). Transfer of training and processing instruction: From input to output. *System*, 39(1), 44–53.

The effect of teaching method on levels of enjoyment, boredom and anxiety of young French learners of English

Jean-Marc Dewaele (Birkbeck, University of London), Delphine Guedat-Bittighoffer (Université d'Angers), Elouise Botes (University of Vienna) and Marie-Ange Dat (Université de Nantes).

Research on foreign language learner emotions has blossomed in recent years (Dewaele et al., 2019). Researchers have sought to identify sources of variation and effects of positive and negative emotions on FL performance and progress. Variation has been linked to both learner-internal and learner-external variables. The latter included “micro-level” classroom variables such as peers and teacher behaviour. However, very little research so far has focused on “meso-level” variables such as schools and teaching methods. The current cross-sectional mixed methods study will do that exactly that. A questionnaire with closed and open questions was used to investigate the effects of a novel foreign language (FL) teaching method on learners’ enjoyment (FLE), classroom anxiety (FLCA) and boredom (FLB) in class. Participants were 181 pupils aged eleven in three secondary schools in France enrolled in beginning English Foreign Language classes. About half of the cohort in each school had standard instruction, while the other half were taught through the Neurolinguistic Approach (Germain, 2018) that emphasizes intense use of authentic communication. The questionnaire included French translations of the original instruments (Botes et al., 2021, 2022; Li et al. 2021). One-way ANOVAs revealed that pupils in NLA classes experienced significantly more FLE and less FLCA and FLB than their peers in standard teaching classes. Structural Equation Modelling revealed that NLA had the strongest (positive) effect on FLE, followed by a strong negative effect on FLB and a weaker negative effect on FLCA. FLE was found to be negatively correlated with FLCA and FLB while FLCA and FLB shared a positive correlation. An axiological analysis of the qualitative data (30,552 words) collected through two open questions (Kerbrat-Orecchioni, 2009) showed that pupils in the NLA groups reported more positive emotions and less boredom than their peers in the standard teaching groups. Pedagogical implications are presented.

References Botes, E., Dewaele, J.-M., & Greiff, S. (2021). The development of a Short-Form Foreign Language Enjoyment Scale. *The Modern Language Journal*, 105(4), 858-876. Botes, E., van der Westhuizen, L., Dewaele, J.-M., MacIntyre, P.D., & Greiff, S. (2022). Validating the Short-Form Foreign Language Classroom Anxiety Scale (S-FLCAS). *Applied Linguistics*, 43(5), 1006–1033. Dewaele, J.-M., Chen, X., Padilla, A.M., & Lake, J. (2019). The flowering of positive psychology in foreign language teaching and acquisition research. *Frontiers in Psychology. Language Sciences*, 10, 2128. Germain, C. (2018). *The Neurolinguistic Approach (NLA) for Learning and Teaching Foreign Languages – Theory and Practice*. Cambridge Scholars Publishing. Kerbrat-Orecchioni, C. (2009). *L'énonciation, de la subjectivité dans le langage*. Armand Colin. Li, C., Dewaele, J.-M. & Hu, Y. (2021) Foreign language learning boredom: Conceptualization and measurement. *Applied Linguistics Review*. <https://doi.org/10.1515/applirev-2020-0124>

Peer interaction dynamics and SLA trajectories during study abroad: Insights from longitudinal social network analysis

Dan P. Dewey (Department of Linguistics, Brigham Young University), R. Kirk Belnap (Department of Asian and Near Eastern Languages, Brigham Young University), Nicole Whitby (Department of Linguistics, Brigham Young University), Piotr Bródka (Network Science Lab, Wrocław Tech), Michał Czuba (Network Science Lab, Wrocław Tech) and Michał B. Paradowski (Institute of Applied Linguistics, University of Warsaw).

Second language acquisition during study abroad (SA) has been a burgeoning field of enquiry over the last three decades. The experience of potential immersion in a target-language (TL, L2)-speaking community coupled with the opportunity of engaging in meaningful communication are commonly believed to be conducive to accelerated progress. However, not all learners benefit equally from SA sojourns, and considerable variation has been evinced in students' linguistic attainment. One line of enquiry that has attempted to explain the varied picture has investigated students' social networks as a factor conditioning their interactions and—subsequently—progress. However, most of the extant studies have i) focused on sojourners' interactions with native speakers of the TL, passing over their communication with other interactants (despite the fact that these conversations oftentimes constitute the majority of their linguistic experience) ii) only investigating participants' egocentric networks (asking students to nominate the persons they talked to, but without gaining respective insight from the nominated alters), iii) if at all, used solely global metrics of the networks extracted (such as size, dispersion, and density), and iv) only measured presojourn proficiency and one-shot post-stay gains. This contribution analyses the longitudinal development of the social interaction network and its influence on L2 gains of 41 U.S. sojourners enrolled in a 3-month intensive study abroad Arabic program. Unlike extant research, the current study i) focuses on students' interactions with their alma mater classmates as well as other agents ii) reconstructing their complete network (in line with the novel computational social network analytic methodology laid out in Paradowski, Jarynowski, Jelińska & Czopek, 2021; Paradowski, Jarynowski, Czopek & Jelińska, 2021; Paradowski, Cierpich-Kozieł, Chen & Ochab, 2022), iii) tracing the impact of each individual student's position in the social graph using betweenness and in-/out-degree centrality metrics, and iv) includes a dynamic developmental perspective with three measurement points at 4-week intervals each, gauging the extent to which changes in the interaction networks translate to changes in both self-perceived and objectively measured progress along a range of dimensions. The learners formed mostly same-gender cliques changing minimally over time. Closeness centrality largely correlated with TL use and self-perceived gains in linguistic and cultural competence, suggesting networking with classmates might facilitate L2 use and development. The best peer-connected students tended to be highly motivated females with high starting L2 proficiency. Interaction with classmates aligned with initial Arabic proficiency and LX competence. Motivation to learn Arabic, LX competence, closeness, and self-reported progress in reading, writing, listening, vocabulary, grammar, and overall Arabic abilities were positively correlated. The strongest predictors of objective proficiency gains were LX competence and level of closeness to classmates. Paradowski, M.B., Bródka, P. & Czuba, M. (under review). Peer interaction dynamics and L2 learning trajectories during study abroad: A longitudinal investigation using dynamic computational Social Network Analysis. *Language Learning (Special Issue: Social Aspects in Language Learning: New Perspectives from Study Abroad Research)*.

Comparing explicit and implicit teaching of phoneme grapheme correspondences.

Nathalie Dherbey Chapuis (Fribourg University) and Raphaël Berthele (Fribourg University).

This study compares two teaching methods of phoneme to grapheme correspondences (=PGC) in French as a foreign language. In French, the PGC are often opaque especially when converting speech to writing (Ziegler et al., 1996) as there are numerous silent letters and multiple possible graphemes representing the same phoneme (/o/ - <o; au; eau; ot...>) (Sprenger-Charolles et al., 1998). For the case of an opaque target language, learners have been observed to have specific problems identifying PGCs and learning the target language's phonological contrasts (e.g., Bassetti, 2017; Hayes-Harb & Barrios, 2021). Learning opaque PGC with new, unlearned phonological contrasts arguably leads to problems with phonological discrimination in reception and production of phonemes, and more generally in processing of phonological lexical knowledge (Bahktiar et al., 2021; Bhide & Perfetti, 2020). However, despite their importance in foreign language learning, PGC are almost never taught in French as a foreign language (Erlar & Macaro, 2011). In this study in French as a foreign language, we compare the effects of explicit teaching (i.e., focus on forms with deductive elaboration of rules and practice) and implicit teaching (i.e., focus on meaning with reading aloud with systematic recast) on the learning of PGC for two difficult phonemes (/ʃ/ and /ʒ/) (Norris & Ortega, 2000). The effects of the two teaching methods were evaluated in a pre-, immediate post- and delayed post-test design. Participants were 127 Swiss German teenagers (mean age 12;6) learning French as a foreign language in the compulsory state school curriculum in Switzerland (level A1-A2 in French). In each of the six classes, the learners were randomly assigned to one of the two groups that followed either explicit or implicit instruction for six weeks (total lasting of instruction 4 hours). Five tasks were repeatedly administered: Receptive vocabulary (Yes/No task), pronunciation (imitation task), discrimination (AX task), maintaining in working memory (non-word repetition task) and PGC task. In the PGC task (Dherbey Chapuis & Berthele, 2019), participants had to identify the written form of dictated non-words among four written propositions: 3 lures and the correct answer; each lure corresponding to an error type either phonological, orthographic or phonotactic. For example, for /ʃ/ and the stimulus /bɔ̃tilé/, written propositions are <boutilé>-phonological lure, <bonetilé>-orthographic lure, <bonntilé>-phonotactic lure, and <bontilé>-correct answer. The response data were analysed with linear mixed models. Overall, the results indicate that participants made significant progress in all five tasks, with no significant differences between the two teaching methods except for PGC. In this latter category, on the one hand, all participants significantly reduced their error rates regardless of teaching method. On the other hand, explicit teaching proved significantly more efficient in reducing the most frequent errors made by the learners (i.e. for /ʃ/, phonological errors and for /ʒ/, orthographic errors). Explicit teaching of PGC seems thus useful for speeding up the recognition of phonemic contrasts when graphemic representation is reliable (i.e. <on> for /ʃ/), and for disambiguating graphemic representation (i.e. <j; g> for /ʒ/).

Tapping into memory systems: using a split-attention paradigm to determine memory signatures for grammar and lexicon across proficiency levels

Dagmar Divjak (University of Birmingham), Petar Milin (University of Birmingham) and Justyna MacKiewicz (University of Birmingham).

Since its proposal more than two decades ago, the declarative/procedural (DP) model has inspired an impressive amount of research into the neurocognitive processes that underly second language learning. Its main predictions, that lexical knowledge depends heavily on declarative memory while grammar depends on both memory systems, are typically tested by establishing associations between learning abilities in declarative or procedural memory and predicted learning, representation, or processing outcomes in L2. Overall, evidence from such studies supports the DP model, though some gaps and inconsistencies suggest the need for further research, using more refined methods to tap into memory systems (Morgan-Short & Ullman 2023). We propose two improvements. Rather than relying on indirect evidence as offered by associations, we introduce on a direct test which is based on the observation that, during retrieval, only declarative memory is disrupted by divided attention: implementing a dual-task paradigm, and dividing attention by loading working memory, thus provides a direct and effective method of probing the declarative and procedural underpinnings of different types of linguistic knowledge. Furthermore, a cross-sectional design spanning classroom students of proficiency levels A2 to B2 as well as real-world users at B2 or above and L1 speakers enables us to track, by linguistic type, whether and when reliance on memory systems changes across proficiency levels and whether real-world conditions of usage support transition to L1-like reliance on memory systems. We recruited a total of 48 L1 speakers of a morpho-syntactically complex language, Polish, as well as 112 L2 learners of Polish at 4 different levels of proficiency whose L1 is Chinese, a morpho-syntactically simple language. All learners were enrolled at CEFR-accredited Polish universities (tutor-assessed CEFR level: 89 in classrooms, with 21 at A2, 36 at B1, 32 at B2 and 23 above B2 studying for university degrees). Participants took a speeded grammaticality judgement task, judging 192 aurally presented sentences with a 50/50 split of correct/incorrect examples of 4 types: lexicon (collocations), morpho-syntax (case; aspect) and syntax (clause coordination and subordination). They made judgments under two conditions: a single-task, full-attention condition, in which only speeded judgments were provided, and a dual-task, divided-attention condition in which speeded judgment was paired with a digit span task. Participants' explicit learning aptitude was assessed using a grammatical inferencing task (LLAMA F) while a serial reaction time task provided an indication of their implicit learning aptitude. Following Divjak, Milin et al. (2022) we analysed three dimensions of performance, i.e., speed, accuracy and consistency or automaticity of judgment, looking at the extent to which these dimensions are affected by dual task demands, how performance changes across proficiency levels and relates to an individual's learning aptitude. While there is a reliable interaction of Condition x Type in L1 users, in L2 learners we find main effects only of Condition, Type and Level. This suggests that in learners who transition from a grammatically simple language to a complex one, no single linguistic type is handled by procedural memory even in case of immersion at levels above B2. We will discuss the theoretical implications of our findings for applications of the DP model to L2 learning.

Transfer effects in the assignment of grammatical gender in L3 vs. L4 Swedish

Kamil Długosz (Adam Mickiewicz University, Poznań).

L2 research on grammatical gender has pointed out that L2 learners typically achieve higher performance levels with nouns that share gender between L1 and L2 (“surface transfer”, e.g., Bianchi 2013). Even the mere presence of gender in L1 has been found to increase the accuracy of gender assignment in L2 (“deep transfer”, e.g., Sabourin et al. 2006). First attempts to approach transfer effects in the acquisition of gender in L3 focused on gender agreement rather than assignment (e.g., Brown 2020). What role does transfer play in the acquisition of gender assignment beyond L2? To address this question, we devised a timed gender decision task including 40 nouns with which we tested two groups of L1 Polish university students of Swedish: 65 L2 English / L3 Swedish students, and 52 L2 English / L3 German / L4 Swedish students. Adopting a pseudo-longitudinal design, we tested individuals who completed one, two, or three to four years of studies. To explore surface transfer from L1 Polish, we manipulated the gender congruency of nouns between Polish and Swedish. Since creating gender-match and mismatch conditions exclusively with German was impossible, we administered an untimed gender decision task in German to L4 Swedish students. Stimuli were controlled for length, frequency, animacy, gender transparency, and cognate status. Accuracy and response time data were submitted to Generalised Linear Models. The results showed that: (i) L4 Swedish students did not outperform L3 Swedish students in terms of accuracy in either year, (ii) 1st-year L4 Swedish students assigned gender to nouns faster than L3 Swedish students, but only if they knew the correct gender of their L3 German translations, and that (iii) 1st and 2nd-year students (both L3 and L4) achieved a higher accuracy if nouns shared gender with L1 Polish than if they did not. The study thus indicates that the mere presence of one vs. two gendered languages does not increase the accuracy of gender assignment in a novel language (no deep transfer). What does, however, is gender congruency with the native language (surface transfer), in line with previous L2 studies. Crucially, gender congruency with a non-native language seems to have a facilitative effect on the automaticity of gender assignment in the earlier stages of learning a novel language. Therefore, the study provides preliminary evidence for a selective multilingual advantage in gender assignment.

References Bianchi, G. (2013). Gender in Italian–German bilinguals: A comparison with German L2 learners of Italian. *Bilingualism: Language and Cognition*, 16(3), 538–557. Brown, M. M. (2020). Grammatical gender acquisition in sequential trilinguals: Influence of a gendered L1 vs. L2. *Proceedings of the Linguistic Society of America*, 5(1), 331–344. Sabourin, L., Stowe, L. A., & de Haan, G. J. (2006). Transfer effects in learning a second language grammatical gender system. *Second Language Research*, 22(1), 1–29.

A mixed-methods study into the role of the time variable in the construct of computer-administered C-Tests in three languages

Anastasia Drackert (g.a.s.t. e.V.; Ruhr University Bochum), Franziska Möller (g.a.s.t. e.V.) and Anna Timukova (g.a.s.t. e.V.; Ruhr University Bochum).

C-Tests as highly efficient and reliable tools for measuring global language proficiency have recently gained much attention in SLA research (e.g. Grotjahn & Drackert, 2022; Norris, 2020). In a C-Test, testees are required to fill the gaps in partially deleted words in short texts. The new time-reduced speeded C-Test (S-C-Test) appears to promise even more in terms of efficiency, but has not yet been sufficiently investigated. In particular, many questions regarding the underlying construct of a S-C-Test remain unanswered. It is hypothesized that, in contrast to a canonical C-Test with a generous time limit of five minutes primarily measuring the amount of learners' declarative and procedural knowledge, a S-C-Test additionally gauges the level of automaticity of their skills (Grotjahn, 2010) and should thus be better at predicting learners' oral skills. So far, no evidence is available to back up this claim. In the presentation, we report on the results of a pilot of a larger study that aims to investigate the role of the time variable in the construct of computer-administered C-Tests across different levels of proficiency in English, German and Russian. A total of 70 L2 learners took a canonical and a speeded C-Test along with eight instruments measuring their declarative (e.g. untimed grammaticality judgement test, orthographic awareness test) and procedural knowledge (e.g. self-paced reading test, written elicited imitation test), general oral language proficiency (oral elicited imitation test) and typing skills. Reaction times on all tests were used as a measure of automaticity. For the presentation, we will report on the results gained through reliability analyses, correlational analyses and the comparison of means for the whole sample and for each of the language groups separately. As expected, all learners' scores were lower on the S-C-Test than on the canonical C-Test but the difference was moderated by their level of proficiency. In particular, the learners of the middle proficiency group showed the largest difference between the speeded and the canonical version. Furthermore, we found higher correlations for the S-C-Test and the oral elicited imitation test than the canonical C-Test irrespective of the language, as well as slightly higher correlations with the measures of the procedural knowledge and automaticity than with the measures of declarative knowledge, which lend preliminary support to the hypothesis formulated by Grotjahn. Learners' typing skills appeared to influence performance in the speeded condition to a greater extent. Apart from the results of the pilot study, we will discuss the changes that need to be introduced into the research design for the main study, and give an overview of the further analyses we will undertake after having collected the data from a bigger sample. The study contributes not only to our understanding of the C-Test construct but also to our understanding of other instruments measuring procedural and declarative knowledge in different linguistic sub skills.

Exploring the components of vocabulary knowledge and their relationships with proficiency in listening and reading

Amber Dudley (University of York), Emma Marsden (University of York) and Giulia Bovolenta (University of York).

RATIONALE. It is generally agreed that vocabulary knowledge is a multicomponential construct (Nation, 2013) and that the individual components are highly interrelated (González-Fernández, 2022). Research further shows that vocabulary knowledge and its components predict L2 proficiency in listening and reading (Zhang & Zhang, 2022). The ability to recognise words quickly, automatically, and accurately is also important for these two skills (Schmitt, 2019). However, we have a limited understanding of the extent to which ‘fluency’ measures (e.g., speed and automaticity of word recognition) relate to ‘traditional’ measures of vocabulary knowledge (such as size or depth) and the relative impact of these two measures on listening and reading comprehension. One approach has conceptualised fluency as a component of vocabulary knowledge together with size and depth (Daller et al., 2007; Godfroid, 2019). These proposals warrant further validation, particularly among low proficiency learners when lexical representations are being established.

RESEARCH QUESTIONS. Our study explores (a) the extent to which fluency measures are psychometrically distinct from traditional measures of vocabulary knowledge, and (b) the extent to which these measures predict listening and reading comprehension.

METHOD. 113 16-year-old learners of French who had recently taken their (high-stakes) GCSE exams after approximately 400–450 hours of instruction completed three vocabulary tasks: a form–meaning recognition, a form recall (English-to-French translation), and a yes–no lexical decision, each containing the same 50 high-frequency words from their GCSE curriculum. These tasks elicited three traditional measures (recognition, recall, and yes–no accuracy scores) and two fluency measures (response times indexing lexical speed and coefficients of variation indexing automaticity). Measures of proficiency were the listening and reading components of the standardised Diplôme d’études en langue française and GCSE listening and reading scores. Data were analysed using Spearman’s correlations, confirmatory factor analyses, and structural equation modelling.

RESULTS. We found that the traditional measures were highly correlated with each other, but not with either fluency measure, and that each traditional measure significantly loaded onto a global ‘vocabulary’ latent variable, but the fluency measures did not. Furthermore, the traditional measures (individually and when loaded onto a traditional latent variable) significantly predicted L2 proficiency in listening and reading, regardless of measure, whereas the fluency measures did not. Taken together, we interpret these findings as preliminary evidence that traditional and fluency measures are psychometrically distinct and that vocabulary knowledge plays a more instrumental role in predicting L2 listening and reading comprehension than fluency, at least among these beginner-to-low intermediate learners. At this level of proficiency, vocabulary knowledge is perhaps yet to be fully automatised, weakening observed relations between fluency and knowledge or proficiency. We discuss implications for foreign language learning, teaching, and testing.

Daller, H., Milton, J., & Treffers-Daller, J. (2007). Modelling and assessing vocabulary knowledge. Godfroid, A. (2019). Sensitive measures of vocabulary knowledge and processing. González-Fernández, B. (2022). <https://doi.org/10.1017/S0272263121000930> Nation, I. S. P. (2013). Learning vocabulary in another language. Schmitt, N. (2019). <https://doi.org/10.1017/S0261444819000053> Segalowitz, N. (2010). Cognitive bases of second language fluency. Zhang, S., & Zhang, X. (2022). <https://doi.org/10.1177/1362168820913998>

The distributed practice effect with an incidental grammatical target

Amanda Edmonds (Université Côte d'Azur), Katerina Palasis (Université Côte d'Azur), Shona Whyte (Université Côte d'Azur) and Emilie Gerbier (Université Côte d'Azur).

Research on the optimal timing of repeated learning sessions shows a robust effect of spacing. This distributed practice effect finds that (a) spaced or distributed schedules result in greater learning than massed schedules and (b) the ideal spacing schedule depends on how long the newly learned material needs to be retained. Much of this research has focused on verbal skills (Wiseheart et al., 2019), involving, for example, the learning of words from a second language (L2). However, the extent to which this robust finding characterizes the learning of other skills continues to be debated, with scholars such as Ullman and Lovelett (2018) suggesting that the distributed practice effect may be specific to skills that rely on declarative memory. This begs the question whether L2 grammar learning – which may at least in part rely on procedural memory systems – is subject to the distributed practice effect. The few studies that have explored this question have focused on intentional (e.g., Sefarty & Serrano, 2022; Suzuki, 2017) or incidental (Rogers, 2015) grammar learning and have shown mixed results, highlighting the need for additional research.

We thus set out to explore the impact of spacing schedules on the learning of Subject-Verb order in embedded questions in L2 English. For this project, the grammatical target was included as an incidental learning objective. We recruited 78 Francophone participants and randomly assigned them to one of two spacing conditions (2 vs. 7 days), which determined when they completed three online learning sessions. Participants were told that they would learn new English vocabulary, although they were also exposed to the incidental grammatical target. At pre- and post-tests (immediate and delayed), learners' knowledge of word order in embedded questions was assessed using two measures, in order to tap the potential development of implicit knowledge (timed oral acceptability judgment task) and explicit knowledge (sentence completion task) (see Vafaei et al., 2017). Results were analyzed using mixed-effect regressions, showing a significant interaction between grammaticality of the target items and spacing condition. Specifically, only ungrammatical items were impacted by spacing.

Rogers, J. (2015). Learning second language syntax under massed and distributed conditions. *TESOL Quarterly*, 49(4), 857-866. Serfaty, J., & Serrano, R. (2022). Lag effects in grammar learning: A desirable difficulties perspective. *Applied Psycholinguistics*, 43, 513-550. Suzuki, Y. (2017). The optimal distribution of practice for the acquisition of L2 morphology: A conceptual replication and extension. *Language Learning*, 67(3), 512-545. Ullman, M. T., & Lovelett J. T. (2018). Implications of the declarative/procedural model for improving second language learning: The role of memory enhancement techniques. *Second Language Research*, 34(1), 39-65. Vafaei, P., Suzuki, Y., & Kachinske, I. (2017). Validating grammaticality judgment tests: Evidence from two new psycholinguistic measures. *Studies in Second Language Acquisition*, 39(1), 59-95. Wiseheart, M. et al. (2019). Enhancing the quality of student learning using distributed practice. In J. Dunlosky & K. A. Rawson (Eds.), *The Cambridge handbook of cognition and education* (pp. 550–584). Cambridge University Press.

Bilingual reading improves bilinguals' ability to integrate information from different texts

Sonja Eisenbeiß (University of Cologne), Andrea Listanti (University of Cologne), Jacopo Torregrossa (Goethe University, Frankfurt) and Christiane Bongartz (University of Cologne).

Most studies on bilinguals' processing of written texts compare bilinguals with monolinguals and involve only one language. However, bilinguals' language processing and discourse-models result from dynamic interactions between all languages of individuals' repertoires ([1],[2],[3]). Our study only involved bilinguals, but both in one-language and in two-language mode, allowing for within-participant comparisons. Specifically, we investigated effects of the two-language mode on bilinguals' abilities to comprehend texts, draw intertextual inferences, and build up discourse representations for multiple texts. This was inspired by didactic activities, where children integrate information from texts written in different languages ([4],[5]). Thirty-nine Italian-German bilingual 4th-graders of a German bilingual school played a board-game: the monkey Cheeky solved problems in 8 map-locations using superpowers. At each map-location, children found Text_A with information about 3 superfruits/superpowers. Then, children got Text_B, a story describing Cheeky's problem. Both texts contained indefinite objects in relevant sentences (Text_A: "eating Meraca lets someone bend metal pieces very easily"; Text_B: "[...] Therefore, the monkey is now stuck in a fence"). Texts appeared in one-language (Italian or German) or in two-language MODE (Italian-German/German-Italian), with three questions: (Q1) Why does Cheeky have a problem? -> COMPREHENSION of Text_B. (Correct: one point). (Q2) Which fruit should Cheeky eat? -> INFERENCE based on Text_A AND Text_B. (Correct: one point) (Q3) What can Cheeky do after eating the fruit? -> TEXT INTEGRATION (one point for referring to both Text_A AND Text_B). -> DISCOURSE BUILDING (one point for using definite/specific referential expressions)

For instance, children received two points for "Cheeky can bend THE metal [Text_A] and get out of THE fence [Text_B]."

A logistic mixed regression model showed a significant MODE-effect on TEXT INTEGRATION: In two-language mode, children integrated information from Text_A and Text_B more frequently for Q3 (Figure 1 in pdf-file). Another model showed a significant MODE-effect for DISCOURSE BUILDING: In two-language mode, children used definite/specific referential expressions more frequently for Q3 (Figure 2). However, MODE did not significantly affect COMPREHENSION and INFERENCE (models for Q1/Q2).

Overall, COMPREHENSION (Q1) and INFERENCE (Q2) patterned consistently across MODEs. For Q3, the two-language mode encouraged elaborate, information-rich responses and discourse representations integrating both texts, suggesting that reading in two-language mode allows children to build more coherent discourse representations. This finding supports the use of bilingual reading-activities in schools.

[1] DeHouwer, A. (2022). The danger of bilingual–monolingual comparisons in applied psycholinguistic research. *Applied psycholinguistics*, 1-15. [2] Rothman, J. et al. (2022). Monolingual comparative normativity in bilingualism research is out of "control": Arguments and alternatives. *Applied Psycholinguistics*, 1-14. [3] Torregrossa, J. et al. (2022). Boosting bilingual metalinguistic awareness under dual language activation: Some implications for bilingual education. *Language Learning*, 1–25. [4] Celic, C., & Seltzer, K. (2011). *Translanguaging: A CUNY-NYSIEB guide for educators*. CUNY NYSIEB. [5] Ascenzi-Moreno, L., & Espinosa, C. (2018). "Opening Up spaces for their whole selves": A case study group's exploration of translanguaging practices in writing. *NY State TESOL Journal*, 5(1), 10–29.

Input matters in L2 acquisition at the interfaces: The case of null objects in L2 European Portuguese

Alexandra Fiéis (Universidade NOVA de Lisboa), Ana Madeira (Universidade NOVA de Lisboa) and Joana Teixeira (Universidade NOVA de Lisboa).

In recent years, generative SLA research has started to examine the role of input in grammar acquisition (Rothman & Slabakova, 2017), a factor that was traditionally neglected in this field. According to some authors (e.g., Dominguez & Arche, 2014; Slabakova, 2015), properties at the syntax-discourse interface, which are proposed to be the main area of permanent difficulties in L2 acquisition by the Interface Hypothesis (Sorace & Filiaci, 2006; Sorace, 2011), only cause persistent difficulties to L2 learners when the evidence in the input is not frequent and transparent and the L1 and the L2 differ in the relevant respects. The acquisition of null objects in European Portuguese (EP) is an appropriate testing ground for this hypothesis, because: (i) they are more productive in spoken/colloquial registers than in standard EP, where clitics are the unmarked pronominal option (Rinke, 2022); (ii) the possibility of definite null objects distinguishes EP from other Romance languages, namely Spanish; and (iii) EP null objects involve the syntax-discourse interface (null objects require a salient and immediately accessible antecedent in the discourse or in the situational context), as well as the syntax-semantics interface (null objects tend to be inanimate). These properties have been extensively investigated in L1 EP (e.g., Flores, Rinke & Sopata, 2020), but not in L2.

To fill this gap, this study investigates the acquisition of clitics and null objects in L2 EP. Participants were 25 L1 EP speakers and 30 L1 Spanish-L2 EP adult learners (10 intermediates, 10 advanced, 10 near-natives). We tested clitics and null objects using two tasks that have been proposed to tap primarily into implicit knowledge (Ellis, 2005): an elicited oral production task (EOPT) and a timed acceptability judgement task (TAJT). Both tasks crossed the variables accessibility of the antecedent (immediately vs. not immediately accessible) and animacy (animate vs. inanimate).

In the EOPT, all groups produced significantly more clitics than null objects across conditions and no significant effects of animacy and accessibility were found. In the TAJT, the native and the L2 groups accepted clitics significantly more than null objects in all conditions. Regarding null objects, the native group exhibits significant effects of animacy (acceptance of null objects is higher in the inanimate condition) and accessibility (acceptance of null objects is higher in the immediately accessible condition). L2 groups do not display any animacy or accessibility effects, even at the near-native level. These results disconfirm the prediction of the Interface Hypothesis that, unlike properties at the syntax-discourse interface, grammar-internal properties are fully acquirable in an L2. Moreover, they show that, at least when the L1 and the L2 are different, learners may have permanent difficulties wrt interface properties when they are infrequent in the input. Finally, they indicate that, unlike what some studies suggest (e.g., Slabakova, 2015; Sorace, 2014), input factors do not selectively affect the syntax-discourse interface. Grammar-internal interfaces are also affected. Thus, convergence in L2 acquisition at the interfaces depends (at least in part) on the frequency and transparency of the evidence in the input.

The role of lexical fixedness in L1 and L2 processing of Italian free combinations versus collocations

Irene Fioravanti (University for Foreigners of Perugia), Marco Silvio Giuseppe Senaldi (McGill University), Alessandro Lenci (University of Pisa) and Anna Siyanova-Chanturia (Victoria University of Wellington; Ocean University of China).

A salient property used to distinguish word combinations (i.e., free combinations, collocations, and idioms) in phraseology is lexical fixedness (Howarth, 1998). Lexical fixedness refers to whether or not a word within a phrase can be substituted with a synonym. Phraseological models usually place word combinations on a continuum of lexical fixedness with free combinations and idioms at the two extremes and collocations somewhere in the middle (Howarth, 1998). Free combinations are defined as the most flexible word combinations, while idioms as the most fixed. On the contrary, collocations allow a limited substitution of their elements.

Previous studies on the processing of word combinations investigated the processing of figurative versus literal meanings of idioms (Siyanova-Chanturia et al., 2011; Carrol & Conklin, 2015), frequency effects on collocation processing (Sonbul, 2015), and the role of adjacency in the processing of collocations (Vilkaitė & Schmitt, 2017). However, no study has so far investigated the role of lexical fixedness in the processing of word combinations.

Seeking to shed light on the issue of lexical fixedness, we carried out an eye-tracking study to explore the processing of Verb+Noun free combinations and collocations by first (L1) and second language (L2) speakers of Italian.

Thirty Verb+Noun free combinations (*guardare un film*, ‘to watch a movie’) and 30 Verb+Noun collocations (*mantenere una promessa*, ‘to make a promise’) were extracted from a reference corpus of Italian and embedded in sentence contexts (original condition). In the modified condition, the verb within free combination and collocation was substituted with its close synonym. We sought to answer the following research questions: Do L1 and L2 speakers of Italian read free combinations and collocations in the two conditions (original and modified) differently? Does L2 proficiency affect the processing of target combinations?

Our results suggest that the degree of lexical fixedness does indeed affect the processing of word combinations in L1 and L2, lending support to phraseological models. Specifically, original collocations were read faster than modified ones. On the contrary, original and modified free combinations were processed in a similar way. Our results further point to the conclusion that higher proficiency L2 learners may process word combinations in a way similar to L1 speakers.

References

- Carrol, G., & Conklin, K. (2015) Cross language priming extends to formulaic units: Evidence from eye-tracking suggests that this idea “has legs”. *Bilingualism: Language and Cognition*, 20 (2), 299-317.
- Howarth, P. (1998) Phraseology and second language proficiency. *Applied Linguistics*, 19, 24-44.
- Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011) Adding more fuel to the fire: An eye-tracking study of idiom-processing by native and non-native speakers. *Second Language Research*, 1-22.
- Sonbul, S. (2015) Fatal mistakes, awful mistakes, or extreme mistakes? Frequency effects on off-line/on-line collocational processing. *Bilingualism: Language and Cognition*, 18, 419-437.
- Vilkaitė, L., & Schmitt, N. (2017) Reading Collocations in an L2: Do Collocation Processing Benefits Extend to Non-Adjacent Collocations?. *Applied Linguistics*, 1-27.

Receptive grammar acquisition from a processability perspective: application to the L3 English context

Helen Forsyth (Free University of Bozen-Bolzano).

The proposed presentation examines Third Language Acquisition (L3 English) within the context of Italian-medium primary schools in bilingual (German-Italian) South Tyrol, Italy. It takes learner early-stage receptive competence in morphosyntactic structures as a focus. Existing research indicates a qualitative difference between Second Language Learning and Third Language Acquisition, and certain psycholinguistic and developmental aspects to multilingual learners have been little explored in the existing literature. The theoretical framework applied in the study is Processability Theory (PT) (Pienemann, 1998, 2005), which examines learner interlanguage from a processing perspective, and proposes that learners proceed through a series of stages in an incremental manner on their route towards acquisition of procedural knowledge of an L2. The study recruited two groups of participants (n=126) and utilized a picture selection task with 6 items to target each of the 9 morphological structures under observation (the Early Language and Intercultural Acquisition Studies Grammar Test (Kersten et al., 2010)), as an instrument for data collection. No control group was used to observe the same patterns in production, as PT is considered an established model of language development for production (Jordan, 2004; Di Biase et al., 2015). Implicational scaling methods (Rickford, 2002) were used to allow quantitative analysis of the results generated. The results were typically in line with the tenets of PT, in that learners exhibited the predicted systematicity in their receptive competence of early-stage morphosyntactic structures. This presentation aims to address the role that processing approaches to language acquisition proposing constraints on developmental readiness may play for the emergence of receptive competence in a bilingual Italian-German population. It concludes by offering implications for classrooms in the multilingual context. The study makes a contribution to the field of processability approaches to language acquisition by exploring their application to L3 development, as well as offering an innovative perspective on learner receptive competence, at both psycholinguistic and applied levels.

Exploring the psycholinguistic reality of L2 corpus-based phraseological complexity: an eye-tracking study across proficiency levels

Luciana Forti (University for Foreigners of Perugia), Irene Fioravanti (University for Foreigners of Perugia), Maria Roccaforte (Sapienza University of Rome), Veronica D'Alesio (Sapienza University of Rome), Valentino Santucci (University for Foreigners of Perugia) and Francesca Malagnini (University for Foreigners of Perugia).

Word combinations (e.g., collocations, lexical bundles, formulae, etc.) are widely recognised as key components in second language competence development, as they aid language fluency, processing and use (Bybee, 2006; Ellis, 2012). Recent linguistic complexity research has thus introduced the construct of phraseological complexity (Paquot, 2019b). Several studies have subsequently shed light on important features of phraseological complexity (Paquot, 2019a; Paquot et al., 2022; Rubin et al., 2021; Vandeweerd et al., 2021), while still, however, relying on corpus data only.

In this paper, we report on a study that seeks to explore the psycholinguistic reality of corpus-based L2 phraseological complexity. Forty-five verb + noun (object) collocations were selected from each proficiency-level subcorpus (B1, B2, C1 and C2) contained in the CELI corpus (Spina et al., 2022), an Italian L2 corpus. Measures of diversity and sophistication were computed for each of the selected collocations, which were then used to build four eye-tracking experiments, one per level. A total of eighty participants (i.e., twenty L2 Italian speakers per level) were involved in the study.

The analysis considered both early and late eye-tracking measures, with respect to the whole collocation. Mixed-effects models show that early measures are influenced by sophistication while late stages are influenced by diversity with both measures interacting significantly with proficiency. Specifically, learners at all four proficiency levels show an increase in reading times as sophistication of collocations increases in the early stages. Similarly, learners at all four levels process collocations with a high degree of diversity more slowly than collocations with a low degree of diversity in the final stages of reading. However, this increase is less evident in advanced learners than in intermediate learners. We will discuss the implications of the study in relation to theory, methodology and pedagogical applications.

Bybee, J. L. (2006). From Usage to Grammar: The Mind's Response to Repetition. *Language*, 82(4), 711–733.

Ellis, N. C. (2012). Formulaic Language and Second Language Acquisition: Zipf and the Phrasal Teddy Bear. *Annual Review of Applied Linguistics*, 32, 17–44.

Paquot, M. (2019a). Using syntactic co-occurrences to trace phraseological complexity development in learner writing: Verb + object structures in LONGDALE. In Le Bruyn, N. and Paquot, M. (Eds.), *Second language acquisition and corpora* (pp. 122- 147). Cambridge University Press.

Paquot, M. (2019b). The phraseological dimension in interlanguage complexity research. *Second Language Research*, 35(1), 121–145.

Paquot, M., Gablasova, D., Brezina, V., & Naets, H. (2022). Phraseological complexity in EFL learners' spoken production across proficiency levels. In A. Leńko-Szymańska & S. Götz (Eds.), *Complexity, Accuracy & Fluency in Learner Corpus Research* (pp. 115–136). Benjamins.

Rubin, R., Housen, A., & Paquot, M. (2021). Phraseological complexity as an index of L2 Dutch writing proficiency: A partial replication study. In S. Granger (Ed.), *Perspectives on the L2 phrasicon: The view from learner corpora* (pp. 101–125). Multilingual Matters.

Spina, S., Fioravanti, I., Forti, L., Santucci, V., Scerra, A., & Zanda, F. (2022). Il Corpus CELI: Una nuova risorsa per studiare l'acquisizione dell'italiano L2. *Italiano LinguaDue*, 14(1), 116–138.

Vandeweerd, N., Housen, A., & Paquot, M. (2021). Applying phraseological complexity measures to L2 French. A partial replication study. *International Journal of Learner Corpus Research*, 7(2), 197–229.

Developmentally Moderated Factors Influencing the L2 Production of English Dative Construction: A Learner Corpus Study

Junya Fukuta (Chuo University), Akira Murakami (University of Birmingham), Masato Terai (Nagoya University) and Yu Tamura (Kansai University).

Previous studies suggested that the choice of competing alternative construction is probabilistic and influenced by various factors. Bresnan et al. (2007) showed that native speakers' choice of dative structures (e.g., John gave Mary a book. vs John gave a book to Mary.) were strongly influenced by such variables as the length of the noun phrase, (in)definiteness, (in)animacy, and other discourse-related variables. Gries and Deshors (2020) and Bernaisch et al. (2014) also took a similar approach and described the different choices of dative structure in second language (L2) use. Previous studies have examined L2 learners' acquisition of multiple constraints related to the English dative alternation phenomenon. However, we still know little about which rules and constraints are acquired at what point in the developmental process. The current corpus-based study describes how various factors influence the L2 choice of dative construction at different proficiency levels.

From EF-Cambridge Open Language Database Cleaned Subcorpus (Shatz, 2020) including approximately 723,000 writings in L2 English on a variety of topics, we extracted 5,785 occurrences of double-object (DO; e.g., I'll give you a call) and prepositional object (PO; e.g., Give your money to charity) constructions involving 23 verbs (e.g., give, offer, show) across CEFR A1-C1 levels and 10 L1 groups. Drawing on the previous literature, we then coded each occurrence in terms of the length, pronominality, and animacy of the theme and recipient. As a verb-level variable, we further added the ratio that DO occupies out of all the occurrences of DO and PO (i.e., $DO / (DO + PO)$) in the Corpus of Contemporary American English. This variable represents the statistical preemption (Goldberg, 2011) of PO in the sense that the higher the value, the more frequently a learner is likely to have been exposed to DO when both DO and PO are communicatively appropriate.

We hypothesized that (i) the length would affect construction selection regardless of proficiency because it involves processing strategies independent of individual languages, (ii) the pronominality would affect construction selection from the intermediate level as reported by Marefat (2005), and that (iii) only high proficiency learners would be sensitive to animacy because its use requires the understanding of meaning of the noun phrase and its relationship to both constructions. Regarding preemption, Goldberg (2019) predicts that learning from such statistical information is extremely difficult for L2 learners, while Xiang and Chang (2022) reported that preemption influences the choice of dative construction even for L2 learners.

A mixed-effects regression model predicting the learner's choice between DO and PO showed that all the coded variables influence the choice of dative construction and that the learners' proficiency interacts with the pronominality of the theme as well as the animacy of the theme and recipient, thereby largely confirming the hypotheses (i) - (iii) above. Interestingly, statistical preemption is associated with the choice of dative construction regardless of the proficiency levels, which points to the need for a multidimensional analysis of statistical preemption, incorporating further covariates and moderators to examine its effects.

Learning L3 words in class: do similarity effects accumulate across languages?

Malgorzata Foryś-Nogala (University of Economics and Human Sciences in Warsaw), Olga Broniś (Cardinal Stefan Wyszyński University in Warsaw), Agata Ambroziak (Institute of English Studies, University of Warsaw), Breno Barreto Silva (University of Warsaw), Aleksandra Janczarska (Institute of Specialised and Intercultural Communication, University of Warsaw), Borys Jastrzębski (University of Warsaw) and Agnieszka Otwinowska-Kasztelanic (Institute of English Studies, University of Warsaw).

L3 lexical acquisition is still under-researched. Although some studies show a processing advantage of triple (L1-L2-L3) cognates over double (L1-L3 and L2-L3) cognates (e.g., Poarch & van Hell, 2014), there is little evidence whether L1-L2-L3 similarity accumulates to enhance L3 lexical acquisition. Bartolotti and Marian (2017) suggest that similarity to just one of the known languages suffices to facilitate L3 learning and that cumulative similarity effects do not arise. Still, their study used short, tailor-made nonwords rather than real L3 words, so their results may not hold for language classrooms. Also, in classroom settings, it is unclear whether learners need to be trained to capitalize on cross-linguistic similarity in learning L3 vocabulary. To fill these gaps, we designed a longitudinal classroom experiment on L3 acquisition testing whether L1-L2-L3 cognates are advantaged in learning compared to non-cognates and L2-L3 cognates. By implementing an experimental manipulation, we also investigated if raising learners' awareness of L2-L3 similarity influences the acquisition of cognates. The stimuli comprised 120 L3-Italian target words: 40 non-cognates (*gorzki/bitter/amaro*), 40 L2-L3 cognates (*skromny/modest/modesto*), and 40 L1-L2-L3 cognates (*stabilny/stable/stabile*), controlled for the part of speech (adjectives or nouns), frequency, length, concreteness, and the orthographic overlap. All words were taught in the classroom through exercises controlled for the number of occurrences of the keywords and task Involvement Load (Laufer & Hulstijn, 2001). The knowledge of the target words was tested twice: in a pre-test and a post-test via a modified VKS scale (Wesche & Paribakht, 1996), which elicited productive knowledge of the L3-Italian keywords and was designed to activate the L1-Polish and L2-English. The participants of our study were L1-Polish university students with upper-intermediate L2-English (pre-test: 67 students; post-test: 41 due to COVID-19; $M_{age}=19.63$, $SD=1.6$). They studied L3-Italian at beginner level and were followed from the first days of the academic year for a period of 10 weeks. We controlled for participants' proficiency in L2 English and L3 Italian (DIALANG), associative memory (LLAMA B), and working memory (a Reading Span task). Crucially, the participants were randomly assigned to experimental and control groups. The experimental group completed six obligatory 5-minute online tutorials concerning L2-L3 (English-Italian) morphological similarity patterns, while the control group did dummy online training. Importantly, the trainings were delivered outside of the vocabulary classes, thus the teacher was blind to the group assignment of the individual participants. The data were analysed via a mixed-effects logistic regression model. The results showed that the instruction benefited the learning of all word types in both experimental and control groups. In the control group, the L1-L2-L3 and L2-L3 cognates were learnt better than non-cognates but there was no difference between both types of cognates. Crucially, in the experimental group, the awareness-raising manipulation advantaged the learning of L1-L2-L3 cognates, as compared to L2-L3 cognates. To conclude, the results suggest that the advantage of the cumulative L1-L2-L3 similarity emerged due to raising the awareness of lexical cross-linguistic similarities.

Factors determining the receptive vocabulary size of school-aged immigrants in their second language

Melanie Fuchs (University of Cologne)

The importance of vocabulary skills particularly for second language learners is well-known (Lervåg and Aukrust, 2010), but most research has focused on young children acquiring a second language. The few available studies that have also focused on older school-aged immigrants indicate that it is easier for children who arrive at a younger age to meet their age-appropriate norms than it is for older children (Paradis et al., 2020; Thordardottir and Juliusdottir, 2013), highlighting age of arrival as an important factor.

The present research focuses on the receptive vocabulary size of school-aged children who recently moved to Germany and have at least spent one year in a German school. The first goal is to investigate the difference between their receptive vocabulary size in German in comparison to that of their peers. The second goal is to assess different factors (age of arrival, length of schooling, German use outside of school) that correlate with larger vocabulary sizes. Since frequency of word exposure affects incidental learning (Rott, 1999), the receptive vocabulary size was assessed using a frequency-based vocabulary test. The test is based on Nation's (1983) Vocabulary Levels Test and consists of five frequency levels containing 30 target words each, which were taken from the 1000, 2000, 3000, 4000 and 5000 most frequent German words.

In a pilot study, 12 newly immigrated students (average age: 17,67) who had spent between 0;3 and 2;01 years in a German school knew more frequent than less frequent words, but their overall results did not correlate with the amount of time they had spent in a German school. In an ongoing study, 110 students (including 17 immigrants) attending German secondary schools have already completed the vocabulary test, a standardized reading comprehension test and a questionnaire. To assess the group difference and the factors determining larger vocabularies, the results of binomial mixed effects models will be reported.

Lervåg A and Aukrust VG (2010) Vocabulary knowledge is a critical determinant of the difference in reading comprehension growth between first and second language learners: Growth in L1 and L2 reading comprehension. *Journal of Child Psychology and Psychiatry* 51(5): 612-620. DOI: 10.1111/j.1469-7610.2009.02185.x.

Nation ISP (1983) Testing and teaching vocabulary. *Guidelines* 5(1): 12-25.

Paradis J, Soto-Corominas A, Chen X, et al. (2020) The home language environment and English second language learning of Syrian refugee children: Comparisons with other newcomer children in Canadian schools. In: Korntheuer A, Pritchard P, Maehler DB, et al. (eds) *Refugees in Canada and Germany: From Research to Policies and Practice*. *GESIS-Schriftenreihe* 25. Köln: GESIS - Leibniz-Institut für Sozialwissenschaften: 143-155.

Rott S (1999) The effect of exposure frequency on intermediate language learners' incidental vocabulary acquisition and retention through reading. *Studies in Second Language Acquisition* 21(4): 589-619. DOI: 10.1017/S0272263199004039.

Thordardottir ET and Juliusdottir AG (2013) Icelandic as a second language: a longitudinal study of language knowledge and processing by school-age children. *International Journal of Bilingual Education and Bilingualism* 16(4): 411-435. DOI: 10.1080/13670050.2012.693062. / /

Enhancing English pronunciation awareness: a longitudinal study of self-perception and self-assessment of own speech

Cristina Aliaga Garcia (University of Barcelona) and Mireia Ortega Duran (University of Barcelona).

Although EFL learners often express a strong desire to attain native-like pronunciation in the L2, research evinces that the nativelikeness principle is far from being a realistic goal. The comprehensibility principle is a more realistic goal in L2 speech learning [1]. Previous studies have widely examined the dimensions that underlie accentedness and comprehensibility through native listeners' ratings [2]; fewer studies have done so through L2 learners' self-assessments [3]; even fewer have analysed these dimensions longitudinally [4]. Emerging work on pronunciation training through self-assessments has shown that such training can have a positive impact on learners' comprehensibility [5]. However, self-perception of pronunciation is a crucial dimension to understand learners' awareness of their pronunciation and learning progress which is still under-researched. The current study explores longitudinal quantitative and qualitative data on learners' self-assessments of comprehensibility and accentedness, and on their self-perception of own speech. We examine (1) the extent to which learners' self-assessments of comprehensibility and accentedness match the ratings of native listeners over time, and (2) the extent to which learners' self-perceptions change after a three-month training period on L2 segmentals. Catalan-Spanish EFL learners (N=48) performed a picture-based story-telling task at pre- and post-test, which they later self-assessed for comprehensibility and accentedness on a 9-point scale. Seven native English listeners also evaluated learners' speech samples for these dimensions. A questionnaire with closed and open-ended questions was administered, prompting learners to reflect on their self-perceived L2-English pronunciation (i.e. satisfaction, awareness of difficulty, self-confidence, speaking anxiety, fear of errors and fear of others' judgements). Findings point to the relevance of training to better align learners' self-assessment with their actual performance, and to minimize the mismatch between self-assessments and listeners' ratings for comprehensibility and accentedness. However, no effects of training were found on pronunciation self-perception, suggesting that self-image of one's accent is rather stable, and that a short training per se cannot guarantee higher self-confidence in pronunciation. This longitudinal study shows a moderate level of anxiety towards mistakes and other people's judgements, which urges some kind of treatment. Future studies could therefore explore the extent to which training to build positive pronunciation self-perception could play a crucial role in L2 pronunciation improvement.

[1] Saito, K., Trofimovich, P. & Isaacs, T. (2017). Using listener judgments to investigate linguistic influences on L2 comprehensibility and accentedness: A validation and generalization study. *Applied Linguistics*, 38, 439–462. [2] Trofimovich, P., and Isaacs, T. (2012). Disentangling accent from comprehensibility. *Bilingualism: Language and Cognition*, 15, 905-916. [3] Trofimovich, P., Isaacs, T., Kennedy, S., Kazuya, S., & Crowther, D. (2016). "Flawed self-assessment: investigating self-and other-perception of second language speech." *Bilingualism*, 19(1), 122–140. [4] Saito, K., Trofimovich, P., Abe, M. & In'nami, Y. (2020). Dunning-kruger effects in second language speech learning: How does self-perception calibrate with other perception over time? *Individual Differences and Learning*. 79, 101849. [5] Tsunemoto, A., Trofimovich, P., Blanchet, J., Bertrand, J. & Kennedy, S. (2021). Effects of benchmarking and peer-assessment on French learners' self-assessments of accentedness, comprehensibility, and fluency. *Foreign Language Annals*. Published online 21/09/2021.

Limitations of the cognate effect: How L2 proficiency and stimulus frequency modulate adolescent second language learners' word recognition

Freya Gastmann (LMU Munich; University of Groningen), Sarah Schimke (Ludwig-Maximilians-Universität München) and Greg Poarch (University of Groningen).

Cross-language activation in bilinguals is a widely researched topic. A word type frequently used to assess cross-linguistic influences is cognates, i.e., translation equivalents that share the same or similar form and meaning across languages (e.g., German-English *Tomate*-*tomato*). Cognates have been found to be processed faster and more accurately by bilinguals than noncognates, i.e., translation equivalents without such overlap (e.g., German-English *Kartoffel*-*potato*), which is considered evidence for co-activation of languages. Although there is ample evidence of such language co-activation during bilingual word recognition in both adults (Dijkstra et al., 2010) and children (Gastmann & Poarch, 2022), research with adolescent second language learners has yielded fewer such findings. To gain further insight into bilingual language comprehension and cognate processing in adolescents, the present study explored L2 word recognition in L1 German low-intermediate learners of L2 English (N = 58; mean age = 13.4, SD = 0.7). Participants performed an English Lexical Decision Task on cognate and noncognate words. Contrary to our predictions, (generalized) linear mixed effects analyses failed to replicate previous findings by yielding no cognate facilitation in either accuracy or reaction time (RT) data. However, post-hoc analyses revealed that cognate processing was modulated by both L2 English proficiency and word frequency. Linear mixed models yielded a three-way interaction ($p = .048$) of cognate status, L2 proficiency (measured by the standardized PPVT; Dunn & Dunn, 1997), and word frequency (SUBTLEX-Lg10; Brysbaert & New, 2009), showing a cognate facilitation effect, with less-frequent items inducing a cognate effect in RTs in low(er)-proficient learners. Generalized linear mixed models on accuracy yielded marginally significant interactions between cognate status and word frequency ($p = .074$) and between cognate status and L2 proficiency ($p = .06$), with cognate effects being restricted to lower-proficient learners and less frequent items. In sum, we did not observe cognate effects for more frequent items and relatively higher-proficient learners, possibly due to ceiling effects. These results expand previous findings of limited cognate facilitation (e.g., Bultena et al., 2014) to a younger learner population. In our presentation, we will discuss limitations of the current study and possible directions for future research.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior research methods*, 41(4), 977-990. <https://doi.org/10.3758/BRM.41.4.977> Bultena, S., Dijkstra, T., & Van Hell, J.G. (2014). Cognate effects in sentence context depend on word class, L2 proficiency, and task. *Quarterly Journal of Experimental Psychology*, 67(6), 1214-1241. <https://doi.org/10.1080/17470218.2013.853090> Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language*, 62(3), 284-301. <https://doi.org/10.1016/j.jml.2009.12.003> Dunn, L.M., & Dunn, L.M. (1997). *Peabody picture vocabulary test-III*. American Guidance Service. Gastmann, F., & Poarch, G.J. (2022). Cross-language activation during word recognition in child second-language learners and the role of executive function. *Journal of Experimental Child Psychology*, 221, 105443. <https://doi.org/10.1016/j.jecp.2022.105443>

EFL learners' receptive and productive vocabulary sizes and the CEFR: An exploratory study

Ferran Gesa (Universitat de Barcelona) and Rosa María Jiménez Catalán (Universidad de La Rioja).

Vocabulary size (VS) has often been linked to different language skills and overall proficiency level (e.g., Miralpeix & Muñoz, 2018). However, there is very scarce research linking VS and the Common European Framework of Reference for Languages (CEFR; Council of Europe, 2020). The limited existing research on the subject matter has shown that learners' receptive VS increases as so does their CEFR level (e.g., Milton & Alexiou, 2009). However, such research was limited to the 1K-5K word range, leaving larger and productive VSs virtually unexplored.

To fill the existing gaps, 295 adult English as a foreign language (EFL) learners participated in this exploratory study. First, they took the grammar and listening parts of the Oxford Placement Test (OPT; Allan, 2004), providing their CEFR level. In the following session, they completed a randomised extended version of the Productive Vocabulary Levels Test (Laufer & Nation, 1999), assessing their productive VS up to 10,000 words. Last, to measure their receptive VS, they took a randomised version of the Vocabulary Size Test (Nation & Beglar, 2007), evaluating the vocabulary included in the 1K-14K word range.

Descriptive statistics indicate that receptive VS ranges from just over 5,450 words at the A1 level to almost 11,200 words at the C2 level. In comparison, productive VS varies between 1,110 words at the A1 level and 6,865 words at the C2 level. Moreover, inferential statistics reveal that OPT scores are correlated with both receptive ($\rho=.644$, $p<.001$) and productive VSs ($\rho=.845$, $p<.001$). They further show that EFL learners' receptive VS significantly increases as they become more proficient [$H(5)=121.276$, $p<.001$], and so does their productive VS [$H(5)=204.558$, $p<.001$]. However, post-hoc comparisons indicate that this growth is not statistically significant between all CEFR levels, especially regarding receptive VS. Findings provide insight into how L2 vocabulary breadth develops in EFL contexts and help to link CEFR descriptors to receptive and productive VS estimates.

References

Allan, D. (2004). Oxford Placement Test 1. Oxford University Press.

Council of Europe (2020). Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume. Council of Europe Publishing. Retrieved from: <https://www.coe.int/en/web/common-european-framework-reference-languages>. Accessed May 3, 2023

Laufer, B., & Nation, P. (1999). A vocabulary size test of controlled productive ability. *Language Testing*, 16(1), 33–51. <https://doi.org/10.1177/026553229901600103>

Milton, J., & Alexiou, T. (2009). Vocabulary size and the Common European Framework of Reference for languages. In B. Richards, M. H. Daller, D. D. Malvern, P. Meara, J. Milton & J. Treffers-Daller (Eds.), *Vocabulary Studies in First and Second Language Acquisition: The Interface Between Theory and Application* (pp. 194–211). Palgrave Macmillan.

Miralpeix, I., & Muñoz, C. (2018). Receptive vocabulary size and its relationship to EFL language skills. *IRAL - International Review of Applied Linguistics in Language Teaching*, 56(1), 1–24. <https://doi.org/10.1515/iral-2017-0016>

Nation, P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9–13. Retrieved from: https://jalt-publications.org/tlt/issues/2007-07_31.7. Accessed May 3, 2023

Examining the relationship between teacher characteristics and learning gains in adult L2 learners with diverse educational backgrounds.

Katrijn Gijswijt (Ghent University), Marieke Vanbuel (Ghent University) and Bart Deygers (Ghent University).

To date, instructed SLA research on L2 gains has overwhelmingly focused on highly educated learners (Andringa & Godfroid, 2020). The few studies that exist on the effects of L2 instruction in learners with limited alphabetic print literacy and/or reduced formal education (LESLLA learners) show that (a) these learners make fewer gains than their more highly educated peers (Deygers et al., 2022) and that (b) L2 learning gains in both populations are significantly and substantially mediated by what happens in the instruction context. This finding corroborates K12 education effectiveness research, which links teacher effectiveness to the teacher's background and (professional) development (Muijs et al., 2014). Similar effectiveness studies are scarce in L2 adult education research, but recent work by Yin et al (2022) in adult basic education confirmed a correlation between teachers' educational background and learning gains. All this indicates that instructor individual differences may impact didactic practices as well as learning gains of students, both for LESLLA learners and highly educated L2 learners.

Because the effectiveness of in-class learning is largely determined by didactic practices, it is important to understand how teacher variables are related to these didactics (Gurzynski-Weiss, 2017). As such, this study investigates the impact of teacher background variables on L2 gains in listening and reading comprehension made by adult L2 learners with various educational backgrounds. We used a questionnaire (N=100) to gauge teacher variables (teacher training, age, experience, etc.) and teaching practices (operationalized by four of Long's Methodological Principles (2014).

By means of two-level regression models (principles within teachers) we examine to what extent these factors influence classroom practice. Secondly, using multilevel modeling and path analyses, we explore to what extent teacher individual differences and didactic practices are associated with learners gains (N=550) in reading and listening skills over a period of four months. Based on the limited research available on non-WEIRD adult SLA learners, we expect to find systematic differences in educational practices among teachers with different profiles. Additionally, we expect differential effects of these principles on SLA development for learners with diverging educational backgrounds.

Andringa, S., & Godfroid, A. (2020). Sampling Bias and the Problem of Generalizability in Applied Linguistics. *Annual Review of Applied Linguistics*, 40, 134–142. Deygers, B., Vanbuel, M., & Knoch, U. (2022). Can L2 course duration compensate for the impact of demographic and educational background variables on second language writing development? *System*, 109, 102864. Gurzynski-Weiss, L. (2017). L2 Instructor Individual Characteristics. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 451–467). Routledge. Long, M. H. (2014). *Second language acquisition and task-based language teaching* (First Edition). Wiley-Blackwell. Muijs, D., Kyriakides, L., van der Werf, G., Creemers, B., Timperley, H., & Earl, L. (2014). State of the art – teacher effectiveness and professional learning. *School Effectiveness and School Improvement*, 25(2), 231–256. Yin, M., Cronen, S., Condelli, L., & Ogut, B. (2022). Teacher Effectiveness in Adult Education: The Relationship Between Teacher Characteristics and Student Test Gains and Transitions Into Postsecondary Education. *Adult Education Quarterly*, 72(3), 262–283.

A methodological synthesis of reporting practices in eye-tracking research

Aline Godfroid (Michigan State University).

Eye-tracking research has taken hold in Second Language Acquisition (SLA) and bilingualism, with over 130 journal articles published to date. Together with the increase in empirical research, the field has seen a growing concern with the methodological aspects of eye tracking (Conklin et al., 2018; Godfroid, 2020), yet a comprehensive picture of methodological practices in second-language (L2) eye-tracking research is still lacking. To help fill this gap, here I will present the findings of a methodological synthesis of the L2 and bilingual eye-tracking literature (k = 145), building on a similar initiative for eye tracking in judgment and decision making research (Fiedler, Schulte-Mecklenbeck, Renkewitz, & Orquin, 2019).

An online search of academic databases and 16 SLA and bilingualism journals resulted in the identification of 81 eye-tracking studies with text and 49 visual world studies. Each unique study was coded for 69 substantive and methodological features in the areas of study rationale, equipment, materials design, participants and item characteristics, measures, and analysis. I will focus on reporting practices as a window into eye-tracking study quality (compare Plonsky, 2013, 2014). The results indicate that reporting standards in eye-movement research differ widely. Methodological details that are essential for reproducing a study are generally provided in 30% to 70% of the studies, but range from virtually absent in all studies (e.g., checks on data quality, recording accuracy and precision, specific screen information) to perfect reporting (e.g., eye tracker model and description of written stimuli).

By taking stock of reporting practices, I aim to advance a Minimal Reporting Standard for eye tracking in SLA and bilingualism and raise awareness of critical methodological features in well-designed eye-tracking studies. Such common standards are essential to enhance replication and reproducibility of eye-tracking research as a part of the Open Science movement.

References Conklin, K., Pellicer-Sanchez, A., & Carrol, G. (2018). *Eye-tracking. A guide for applied linguistics research*. Cambridge University Press. Fiedler, S., Schulte-Mecklenbeck, M., Renkewitz, F., & Orquin, J. L. (2019). Increasing reproducibility of eye-tracking studies: The EyeGuidelines. In M. Schulte-Mecklenbeck, A. Kühberger & R. Ranyard (Eds.), *A Handbook of Process Tracing Methods for Decision Research. A Critical Review and User's Guide*. Psychology Press. Godfroid, A. (2020). *Eye tracking in second language acquisition and bilingualism: A research synthesis and methodological guide*. Routledge. Plonsky, L. (2013). Study quality in SLA: An assessment of designs, analyses, and reporting practices in quantitative L2 research. *Studies in Second Language Acquisition*, 35(4), 655-687. Plonsky, L. (2014). Study quality in quantitative L2 research (1990-2010): A methodological synthesis and call for reform. *The Modern Language Journal*, 98(1), 450-470.

Comparing the L2 proficiency and lexical development of app-based vs. classroom-based EFL learners

Beatriz González-Fernández (The University of Sheffield).

Online language learning applications (apps) have become increasingly popular among second/foreign language (L2) learners for their convenience, affordability, flexibility, and personalisation of the learning process (Loewen et al., 2020; Kukulska-Hulme & Viberg, 2018). Recent studies suggest that app-based language learning can be effective for developing L2 competence (Loewen et al., 2020; Rachels & Rockinson-Szapkiw, 2018), although the evidence is stronger for receptive than productive knowledge (Loewen et al., 2019; Lin & Lin, 2019). However, very few empirical studies have compared the effectiveness of app-based language learning and face-to-face classroom-based instruction in facilitating L2 development (Jiang, et al., 2021), leaving many SLA scholars and practitioners sceptical about apps' potential to be an alternative to traditional classroom instruction (Lord, 2015; Loewen et al., 2020). In addition, research has not typically examined how the mode of language study (app-based vs. classroom-based) influences learners' motivation and engagement with the course and language (He & Loewen, 2022).

This quasi-experimental pretest-posttest study explores the L2 linguistic development and motivation of 300 adult L1-Spanish learners enrolled exclusively in app-based (i.e., treatment, n =150) or classroom-based (i.e., control, n = 150) L2-English instruction over the course of an academic semester (16 weeks). Participants had just completed an A1-English course either in the app Duolingo or in a public language institute in Spain, and were starting an A2 course (CEFR, Council of Europe, 2017). Participants' general L2 proficiency and receptive and productive vocabulary knowledge were measured with the Oxford Placement Test and the Updated and Productive Vocabulary Levels Test, respectively (Webb et al., 2017; Laufer & Nation, 1999), at the start of the A2-English course and after 1 semester. In addition, a questionnaire assessing motivation and engagement with the language course was also administered before and after the 16-week learning period. To ensure a comparable amount of time studying English across both groups, the treatment group had a study goal of 3-4 hours per week, equivalent to the instruction time of the control group.

Results from descriptive statistics and linear mixed-effect regression analyses showed that app-based learners generally spent less time studying English than the classroom-based learners, despite their weekly study goal. Yet, their performance on the proficiency and vocabulary tests did not differ substantially between groups. Learner engagement and motivation with the course evolved during the instruction period between and within groups, being generally higher and sustained to a greater extent by the classroom-based learners than app-based learners. Findings based on qualitative thematic analysis of learner motivation and engagement supported the importance and beneficial effect of apps' interactivity and flexibility for the app-based learners, and of the teacher and peer support for the classroom-based group. Practical implications for research and pedagogy will be discussed.

On phonological and visual processing in L2 spelling

Cecilia Gunnarsson-Largy (LNPL Université Toulouse - Jean Jaurès).

This presentation reports on the last study in a series of experimental studies on the processing of spelling in L2 French writing. We had observed that L2 writers are quite good spellers and in homophonous contexts sometimes even better than L1 writers (Author, 2013). We have since set up different studies to explore why. The first studies seem to indicate that the better spelling in homophonous contexts in L2 learners than L1 writers, is not due to more control (Author, 2016), and seems to be related to the retrieval and maintaining of different word forms (verbal vs. visual) in L1 and L2 writers (Author, 2019). These results made us continue exploring the process of retrieving the words from long term memory and maintaining them in Working Memory (Kellogg, 1996) in L1 and L2 writing. Adult L1 writers retrieve and then maintain a phonological form of the word corresponding to the intended sense of the phonological form. Having developed automaticity they will transcribe the most frequent corresponding orthographic form (Largy et al., 1996). This sometimes leads to spelling errors (e.g., of homophones). As L2 writers, even deprived of control, make less errors, our hypothesis is that they directly retrieve and then maintain the exact visuo-orthographic form (van Berkel, 2004) of the intended word. The verbal vs. visual working memory (WM) (Baddeley, 1986; 2000) should therefore be differently impacted in L1 and L2 writers. We tested this hypothesis in a double-task experiment with three conditions: 1) no double-task (no extra cognitive load); 2) a visual double-task inducing cognitive load on visual WM; 3) a verbal double-task inducing cognitive load on verbal WM. The main task concerned the grammatical spelling of dictated sentences with silent (non-audible) verbal morphology and agreement by proximity (Fayol & Got, 1991), i.e., the verb is agreed to the nearest noun: *Le chien des voisins mange* vs. **Le chien des voisins mangent*. The protocol had control sentences, with audible agreement (i.e., *Le chien des voisins boit* vs. *Les chiens du voisin boivent*), and filler sentences. In total the participants had to write 24 sentences per condition: 6 with non-audible morphology (3 singular and 3 plural), 6 with audible morphology (3 singular and 3 plural) and 12 filler sentences. While writing the sentences in the double-task conditions the participants had to maintain either three non-words or three (not-namable) figures. We compared 24 adult L1 French writers with 24 adult L2 French writers of different L1s at advanced levels (B2/ C1). The participants passed all three conditions in counter-balanced order to avoid an order effect on the results. The first results seem to show that L1 writers make more errors in the homophone (non-audible agreement) context and rely more on phonological WM. L2 writers seem less impacted by the non-audibility. As to the reliance on WM, both visual and verbal WM seem to be implicated in the L2 writers.

Crossing research paradigms: the case of formulaic expressions in longitudinal learner corpus data

Thomas Hammond (University of Sheffield) and Kook-Hee Gil (University of Sheffield).

Aims: This study adopts concepts from two competing approaches to second language acquisition (SLA) (usage based vs. generative) to analyse the effect of formulaic expressions (FEs) on learners' syntactic development. We examine how these approaches interplay in accounting for the development of syntactic knowledge associated with FEs. **Background:** The role of FEs in SLA is traditionally associated with usage-based approaches which posit an extraction of their utterance schemas as a gateway to novel L2 creativity (Eskilsden 2015). For example, through frequent exposure to and usage of the prototypical fixed wh-expression 'where do you live', learners can derive the utterance schemas [where do]/[WH + AUXDO] + X and [do you]/[AUXDO+ PRN] + X, facilitating production of other interrogatives such as 'where do you go', 'do you like' etc. Generative approaches claim the creative language process develops independently of FE analysis (e.g., Krashen & Scarcella 1978). We take this to mean that related computational derivations (e.g., wh-movement, do-support) develop independently of FE usage. **Analysis:** Our data comes from the spoken transcripts of the Barcelona English Language Corpus. Nine bilingual Spanish/Catalan learners of English participated in naturalistic interview tasks across four rounds, split into beginner rounds (ages 10&12) and final rounds (ages 16&17). We measure both the use of four learned formulaic wh-expressions (FwhEs) and the evidence of related computational derivations outside their use. **Results and Discussion:** We find that 53% (20/38) of learners' interrogatives (wh- and yes/no) across the course of the data collection period embody the same utterance schemas of previously used FwhEs. We also find the development of computational derivations, independent of FwhE schematic extraction. In these seemingly independent processes, an earlier and more frequent FwhE use appears to correlate with an overall better knowledge of the expressions' associated syntactic derivations more generally: wh-movement (.596), T-C movement (.548), A-movement (.731) and do-support (.600). We argue that a unified account for learners' observed development based on FwhE use is achieved by combining concepts from both usage-based and generative approaches. Over half of learners' L2 interrogatives embody the same utterance schemas as previously used FwhEs, suggesting that the extraction and generalisation of these schemas facilitates the production of similar complex structures. An increased production of these complex surface forms can help to better 'activate' the more general L2 feature specifications on the functional categories for which they exemplify. This explains the correlations between FwhE use and knowledge of their associated syntactic derivations more generally. We discuss how this finding can be explained using the concept of 'activation' under the MOGUL framework (Truscott & Sharwood-Smith 2004).

References Eskilsden, S. (2015). What counts as a developmental sequence? Exemplar-based L2 learning of English questions. *Language Learning*, 65(1), 33-62. Krashen, S. & Scarcella, R (1978). On routines and patterns in language acquisition and performance. *Language Learning*, 28(2), 283-300 Truscott, J. & Sharwood Smith, M.S. (2004). Acquisition by processing: A modular perspective on language development. *Bilingualism Language and Cognition*, 7, 1-20

An investigation of changes in Japanese EFL learners' anxiety, enjoyment, and confidence while performing three similar task-based activities.

Masashi Haneo (Kansai University).

This study investigated the emotional states of second language learners who experienced task-based activities. Recent research on learner emotions in L2 learning using tasks (e.g., Xu & Qiu, 2021; Van den Branden et al., 2007) has shown that emotions like anxiety, confidence, and enjoyment are influential in L2 learning. These feelings are also found to be dynamic and flexible within a task-based activity (Xu & Qiu, 2021). Yashima (2019) insists paying attention to learners' emotions, especially anxiety, is important in classrooms where students frequently perform communicative activities; it is especially true in Japanese English classrooms. However, the research focusing on students' emotions in task-based learning is still limited. Thus, this current study investigated if and how the emotional states of Japanese learners of English changed when they participated in three task-based activities. A total of 86 learners of English as a foreign language from 4 classes at a Japanese university completed three similar tasks (task repetition) in pairs. They also answered questionnaires in which they rated their emotions on a six-point scale at four different points in time: 1) before beginning a series of task activities, 2) after task 1, 3) after task 2, and 4) after task 3. After the 3rd task, they were also asked to create an emotion graph based on their ratings and write reports of the emotional changes in their own words. The findings showed that learners' enjoyment and confidence generally increased with time, whereas their anxiety reduced. Many students' anxiety and enjoyment scale scores changed by 2 points or more at Time 4 (i.e., after the third task) in comparison to Time 1. Their reports indicated that task repetition contributed significantly to changes in their anxiety and enjoyment. It was found in the reports that the students became familiar with the tasks when their anxiety decreased. They also recognized that their English had improved. When the students' enjoyment in the emotion graph increased, many students reported they felt their communication skills and English skills improved. The familiarity with their partners also contributed to the increase in enjoyment. On the contrary, confidence did not change as much as anxiety or enjoyment. Students' reports revealed that they hardly felt they were competent in speaking in English even after performing the three tasks. There were some unique cases: Some students' anxiety and enjoyment levels fluctuated due to the difficulty and content of tasks. In addition, differences among learners were also noteworthy: while some students showed an increase in anxiety and a decrease in enjoyment, others showed no changes in the target emotions throughout the tasks. In sum, the students in this study showed positive emotional changes after repeated tasks: they became less anxious, had more enjoyment, and were relatively confident in their learning. Furthermore, some students exhibited strong individualities. The findings suggest that teachers using task-based language instruction need to be conscious of the task features they utilize in the classrooms and the dynamic emotional changes of their students.

Predictive processing in the heritage language?: The case of Mandarin classifiers

Jiuzhou Hao (UiT The Arctic University of Norway), Maki Kubota (UiT The Arctic University of Norway), Fatih Bayram (UiT The Arctic University of Norway), Jorge González Alonso (UiT, The Arctic University of Norway), Theres Gruter (University of Hawai'i at Mānoa) and Jason Rothman (UiT the Arctic University of Norway).

Background: Classifiers are morphological units that mark the class of nouns. In Mandarin, classifiers are obligatory when a counting of the noun is present (example 1). Importantly, Mandarin classifiers encode both lexical-semantic and lexicalized grammatical information, allowing us to examine the relative weighting of different linguistic informational cues in sentence processing. For example, the sortal classifier, *tiao*, semantically associates with nouns that conform to the description “slender, long-shape, often flexible”, e.g., *shengzi* “rope”. On the other hand, *gou* “dog” co-occurs with the *tiao* classifier but is clearly not a long, slender, flexible object, suggesting the lexicalized grammatical association between the classifier and the noun. Moreover, there are nouns that cannot co-occur with the *tiao* classifier, but nevertheless possesses semantic properties associated with it, e.g., *shoubiao* “wristwatch. Recent literature [1] using eye-tracking visual world paradigm showed that L1 Mandarin speakers rely more on the lexicalization (grammatical cue) of classifiers, while L2 speakers rely more on semantic information. This current study examines how Mandarin-English heritage bilingual speaker (HSs)—native speakers of Mandarin who are, however, dominant in a different (their societal) language—process Mandarin classifiers.

Methods: We adopt the visual word eye-tracking task used in [1], targeting three Mandarin sortal classifiers, i.e., *tiao*, “~long, flexible”; *zhi*, “~stick-like, long”; and *zhang*, “~flat, spread open”. Participants are shown visual stimuli with three objects (Figure 1): the target, the competitor, and a distractor, while listening to sentences (audio stimuli, example 2). For the target, we selected the noun, e.g., *gou* “dog”, that can co-occur with the classifier, e.g., *tiao*, (lexicalized grammatical match; G+), but does not belong to the prototypical semantic class the classifier is associated with, (semantic mismatch; S-). For the competitors, we manipulated the grammatical and semantic (mis)matching between the nouns and the classifiers, i.e., G-S- (*pingguo* “apple”), G-S+ (*shoubiao* “wristwatch”), and G+S+ (*shengzi* “rope”). A total of 12 experimental trials, with 4 items in each of the 3 conditions were made and rotated across three lists. Audio stimuli were constructed by concatenating extracted tokens of *na yi* “which one”, the classifier, *shi* “is”, and the target noun.

Results and discussions: Data from 55 Mandarin-English HSs were included in the analysis. Figure 2 illustrates the eye-gaze patterns across conditions, showing a descriptive difference of eye-gaze between G-S+ and G-S- as well as between G+S+ and G-S+. Statistical analyses using logistic mixed-effect regressions confirmed the across-condition differences between G+S+ and G-S+ but a marginal difference between G-S+ and G-S-. These results suggest that HSs used both grammatical and semantic features encoded in the classifiers to predict upcoming nouns but to different degrees. As such, HSs, at least at the aggregated group level, differ from both L1 and L2 speakers in their relative weighting of grammatical and semantic cues. Further inspection of individual HSs showed individual differences in their relative reliance on different cues, calling for future research to understand individual differences.

References: [1] T. Grüter, E. Lau, W. Ling, *Lang. Cogn. Neurosci.* 35 (2020) 221–234.

Constructing a new test for metalinguistic awareness: The MetaLearn test and its baseline results

Kaja Haugen (Western Norway University of Applied Sciences), Christine Möller-Omrani (Western Norway University of Applied Sciences), Monika Bader (Western Norway University of Applied Sciences) and Cecilie Hamnes Carlsen (Western Norway University of Applied Sciences).

The present study is part of the larger research project Education for plurilingualism: Metalinguistic awareness in early instructed language learning (MetaLearn; <https://www.hvl.no/en/research/prosjekt/metalearn/>). MetaLearn centers on the investigation of metalinguistic awareness (MLA) as a key element in promoting plurilingualism and language learning (cf. Beacco et al. 2016; Candelier et al. 2012). MetaLearn investigates, from a Norwegian perspective, the state and development of MLA of elementary school pupils from early third grade (age 7 to 8), through the end of third grade (age 8 to 9), till the end of fourth grade (age 9 to 10). For this purpose, we developed a new MLA test, the MetaLearn test, which systematically draws on principles from measurement theory and language testing. To the best of our knowledge, this is a unique approach in the study of MLA.

The MetaLearn test was pre-tested in fall 2021 with third (N= 176; age 8-9) and fourth graders (N= 144; age 9-10). Based on pre-test analyses, two parallel test versions were constructed, which are of equal difficulty (p-value, mean item difficulty set 1: .63; set 2: .64) and discriminate well between pupils with different degrees of MLA (mean biserial coefficient, set 1: .40; set 2: .42). The two versions are similar in construct, content, structure, and task formats, which makes it possible to measure the development of MLA in reliable, valid, and practical ways.

In fall 2021, we conducted the baseline assessment in third grade (N=177; mean age 8). The third graders obtained, on average, 61% of the total possible score (mean score 22 out of 36). We also found clear individual differences—evident, for example, from the scoring range (min. 5 to max. 34 points). Additionally, our results confirm differences in development regarding the subdimensions of MLA: As early work on MLA already suggested, e.g. Gombert 1992, metaphonological awareness seems to develop earlier, for example, than metagrammatical awareness.

In the present paper, we will first discuss our conceptualization of MLA (which draws on Bialystok 2001) and how it was operationalized for the test. We will then give a brief overview of the steps we took to develop the MetaLearn test. Finally, we will discuss some of the results obtained in the baseline assessment in more detail.

References Beacco, J.-C., Byram, M., Cavalli, M., Coste, D., Cuenat, M.E., Goullier, F. & Panthier, J. 2016. Guide for the development and implementation of curricula for plurilingual and intercultural education. Strassburg: Council of Europe.

Bialystok, E. 2001. *Bilingualism in Development*. Cambridge: Cambridge University Press.

Candelier, M., Daryai-Hansen, P., & Schröder-Sura, A. 2012. The Framework of Reference for Pluralistic Approaches to Languages and Cultures: A Complement to the CEFR to Develop Plurilingual and Intercultural Competence. *Innovation in Language Learning and Teaching*, 3, 243-257.

Gombert, J. E. 1992. *Metalinguistic development*. Chicago: University of Chicago Press.

Procedural memory and lexical constraints on L2 implicit grammatical learning

Holger Hopp (TU Braunschweig), Jana Reifegerste (Georgetown University) and Michael T. Ullman (Georgetown University).

Syntactic adaptation constitutes one form of implicit grammatical learning: language users tune their grammatical knowledge to input statistics (Kaan & Chun, 2018). Compared to native speakers, L2 learners seem show less syntactic adaptation (e.g. Hopp, 2020; Kaan et al., 2019). Here we examine linguistic and cognitive constraints on L2 syntactic adaptation.

In a pre-post test design with an exposure phase, we investigated whether advanced German-English learners adapt their interpretation preferences for ambiguous relative clauses (RCs; (1)) to the target low/N2 attachment preference in English (\diamond sister). Following the pretest, they listened to 48 exposure sentences with non-ambiguous RCs (plus 72 fillers). Of the 48 RCs, 42 were disambiguated by number agreement on the verb to low/N2 attachment (2). To test lexical effects on syntactic adaptation, there were two proficiency-matched exposure groups ($n = 48$ each). In the BE-group, exposure sentences were disambiguated by the verb “be” (is; (2)). In the LEXICAL-group, disambiguation was by a lexical verb (e.g., advises; (2)).

(1) Bruce liked the uncle of the sister who lived in London. (2) Polly mistrusted the consultants of the manager who is/advises at the board meeting.

All learners were tested on L2 lexical knowledge and processing (LexTALE, lexical decision tasks) and on learning in declarative memory (object recognition memory test, following incidental encoding) and procedural memory (Serial Reaction Time task) to assess potentially relevant individual differences in lexical and learning abilities. According to the Lexical Bottleneck Hypothesis (LBH; Hopp, 2018), more effortful lexical processing restricts syntactic structure building and learning in sentence processing. The LBH predicts more learning overall in the BE-group due to easier lexical processing of the disambiguating verb “be” than in the LEXICAL-group, within which individual differences in lexical knowledge and processing should predict the amount of learning. According to the Declarative/Procedural model (DP model; Ullman, 2020), fast feedback (e.g., disambiguation) favors learning in procedural (versus declarative) memory. Because “be” should be processed faster than the lexical verbs, the BE-group should therefore show more grammatical/procedural learning, with individual differences in procedural learning ability predicting amount of learning.

Both groups displayed no RC attachment preference at pretest (= chance), and only the BE-group showed adaptation to N2/low attachment from exposure in the posttest, with posttest preferences differing from chance (Fig. 1). In regression analyses, only procedural learning ability (Serial Reaction Time) predicted learning in the BE-group (Fig 2a), and only vocabulary size (LexTALE) predicted learning in the LEXICAL-group (Fig. 2b). These findings are in line with both the LBH and the DP model and suggest that individual differences in both linguistic (lexical) and cognitive (procedural) factors systematically constrain implicit grammatical learning in adult L2 learners.

References Hopp (2018). Bilingual mental lexicon in L2 sentence processing. *Second Language* Hopp (2020). Morphosyntactic adaptation in adult L2 processing. *Applied Psycholinguistics* Kaan & Chun (2018). Syntactic adaptation. *Psychology of Learning and Motivation*. Kaan et al. (2019). Adaptation to syntactic structures in native and nonnative sentence comprehension. *Applied Psycholinguistics*. Ullman (2020). The Declarative/Procedural model. *Theories in second language acquisition*.

Presentation Formats and Attention: Evidence from Eye Movements When Learning L2 Chinese Vocabulary

Xuehong Stella He (Swansea University).

Instructed second language (L2) research has long focused on enhancing the salience of input so that learner attention can be efficiently directed to stimulate L2 development (Robinson et al., 2012). Besides input salience, the presentation format of L2 input can be another important factor of attention according to the Cognitive Load Theory (Sweller et al., 1998, 2019). Specifically, when multiple sources of information are presented separately, split-attention occurs, and learners have to assign attentional resources to mentally integrate the separated information before they can concentrate on learning the information itself. Lee and Kalyuga (2011) studied split-attention in learning L2 Chinese words and found better learning outcomes when the Chinese characters, pinyin (sound script), and English translation were presented vertically (from top to bottom) rather than horizontally (from left to right). They speculated that the vertical format reduced split-attention and that an adjacent format (pinyin below and English translation to the right of the characters) might reduce split-attention even further. However, their study did not employ online measures of attention to test this hypothesis.

This study adopted eye-tracking technology (EyeLink 1000) to investigate how these three presentation formats affect learner attention and learning outcomes of L2 Chinese words. Sixty-nine English native speakers' eye movements were recorded when learning 30 Chinese words. These words were presented in three groups of ten, with each group arranged in one of the three formats. Vocabulary learning was measured in a pre/post-test design with bilingual recall and recognition tasks (Laufer & Goldstein, 2004). Learner perceptions of the learning process were collected with surveys and interviews for data triangulation. L2 Chinese proficiency was also assessed as an individual difference.

Results from mixed effects modeling showed that the adjacent format generally led to higher vocabulary learning gains than the horizontal and vertical formats, and that L2 proficiency also contributed positively. Across the three presentation formats, learner attention as measured by fixation durations and fixation counts, affected vocabulary learning in a positive or negative way. Specifically, longer fixation durations and more fixation counts on characters led to better learning outcomes, whereas longer fixation durations and more fixation counts on pinyin and meaning had a negative impact on learning outcomes. Theoretical and pedagogical implications will be discussed.

References Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning*, 54(3), 399-436. Lee, C. H., & Kalyuga, S. (2011). Effectiveness of different pinyin presentation formats in learning Chinese characters: A cognitive load perspective. *Language Learning*, 61(4), 1099-1118. Robinson, P., Mackey, A., Gass, S., & Schmidt, R. (2012). Attention and awareness in second language acquisition. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 247-267). New York: Routledge. Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3), 251-296. Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, 31(2), 261-292.

Intensity matters in CLIL: Evidence from primary school learners' receptive skills

Maria Angeles Hidalgo (Public University of Navarre) and Izaskun Villarreal (Public University of Navarre).

Content and language integrated learning (CLIL) methodologies are becoming commonplace in primary schools all over the world (Enever, 2018, García Mayo, 2021). These methodologies improve learners' linguistic competence (Ruiz de Zarobe, 2011), particularly receptive skills (Ruiz de Zarobe & Jiménez Catalán, 2009). Still, great achievement differences can be found within these programs. These differences have been linked to the varying number of hours of exposure to the target language learners have (Pladeval-Ballester & Vallbona, 2016). This number may fluctuate greatly from programme to programme, even in the same city. Nevertheless, few studies have addressed linguistic achievement in varying-intensity CLIL programs. In order to shed some light into this research gap, we have addressed the two ends of a continuum by exploring data gathered from two state schools in northern Spain: A low-intensity (LoI) CLIL primary school where learners receive 7 hours of exposure to English (a subject course taught through CLIL and English language classes) and a high-intensity (HiI) school which teaches 15 hours in English (three subject courses taught through CLIL and English language classes). The participants were 114 (LoI: n= 72; HiI: n= 42) young learners (YIs) in their 6th year of primary school (mean age: 11.42) with an A2 level of command of English on the Common European Framework of Reference for Languages (CEFR). Both schools follow CLIL programs which are mandatory for all pupils. Specifically, we have compared the effect of the number of hours of exposure on YIs' proficiency in receptive skills (listening and reading). The participants took a series of short listening and reading comprehension tests following the A2 Flyers exams from the Cambridge English Qualifications Papers for Young Learners series. Results show that the learners who receive the most intense exposure to the target language obtained results that were significantly better in both tests than those with less intense exposure (15 vs 7 hours per week). With the current study, we hope to inform teaching practices by exploring the role of the intensity of exposure and contribute to the development of common guidelines for the design of more effective CLIL programs.

Enever, J. (2018). Policy and politics in global primary English. Oxford University Press. García Mayo, M. P. (2021). Editorial: Research on EFL learning by young children in Spain. *Language Teaching for Young Learners*, 3 (2), 181–188. Pladevall-Ballester, E., & Vallbona, A. (2016). CLIL in minimal input contexts: A longitudinal study of primary school learners' receptive skills. *System*, 58, 37–48. Ruiz de Zarobe, Y. (2011). Which language competencies benefit from CLIL? an insight into applied linguistics research. In Y. Ruiz de Zarobe, J. M. Sierra, & F. Gallardo de Puerto (Eds.), *Content and foreign language integrated learning: Contributions to multilingualism in European contexts* (pp. 129-153). Frankfurt-am-Main: Peter Lang. Ruiz de Zarobe, Y., & Jimenez-Catalán, R. M. (2009). In *Content and language integrated learning: Evidence from research in Europe*. Bristol: Multilingual Matters.

Lexical complexity and assessment of EFL writing: a study of the assessment of English vocabulary in the Swedish national tests

Christian Holmberg Sjöling (Luleå University of Technology).

Vocabulary is a core component of language proficiency (Nation, 2020) and texts containing an appropriate and precise vocabulary are considered to be of high quality by assessors (e.g., Garner et al., 2019; Vögelin et al., 2019). Despite regularly being claimed as an explicit grading criterion in language tests, what constitutes sufficient/good vocabulary is often expressed in vague terms in guidelines to assessors (e.g., ETS, 2022). Every year in Sweden, upper secondary school students are required to take the national tests of English to ensure that their proficiency is on par with the level at which they study (Olsson, 2018). During the tests, students are required to write texts on a specific topic and these are then assessed by teachers with instructions of assessment created by a group of experts on behalf of the Swedish National Agency for Education. These instructions specify vocabulary as a grading criterion and indicate that there should be a progression in terms of lexical complexity between the lowest and highest grade. Furthermore, to ensure fair and equal assessment, the instructions provide graded example texts to assist teachers' assessment. This paper aims to investigate the assessment of written productive vocabulary by analysing these graded example texts and texts graded by teachers during the exams. The material consists of a corpus of 142 graded example texts and 190 teacher graded texts from two courses written between 2011 and 2022. A range of measures of lexical sophistication (e.g., frequency, range, n-grams) were employed utilising different written subsections of the Corpus of Contemporary American English (COCA) as reference corpora. Also, one measure of lexical diversity was used (Moving Average Type-Token Ratio). The results indicate that there is no structured difference in terms of lexical complexity between texts awarded the lowest and highest grade, suggesting that written vocabulary proficiency was largely overlooked in the assessment. In addition, the findings call into question the construction and validity of the Swedish national tests of English since productive vocabulary, although an essential part of overall language competence and explicitly mentioned a grading criterion, does not seem to have been taken into consideration when the example texts were graded.

References Educational Testing Service (ETS). (2022). Writing Scoring Guides (Rubrics). TOEFL iBT® Writing Scoring Guide Flyer (ets.org) Garner, J., Crossley, S., & Kyle, K. (2019). N-gram measures and L2 writing proficiency. *System*, 80, 176–187. <https://doi.org/10.1016/j.system.2018.12.001> Nation, I. S. (2022). *Learning vocabulary in another language*. Cambridge University Press. Olsson, E. (2018). Vokabulär och bedömning av skriftlig förmåga. In G. Erickson (Ed.), *Att bedöma språklig kompetens* (pp. 139–167). University of Gothenburg. Vögelin, C., Jansen, T., Keller, S. D., Machts, N., & Möller, J. (2019). The influence of lexical features on teacher judgements of ESL argumentative essays. *Assessing Writing*, 39, 50–63. <https://doi.org/10.1016/j.asw.2018.12.003>

L2 perception training and its effects on perception and word recognition in the Dutch classroom

Marlisa Hommel (Utrecht University).

Previous L2 perception training studies have shown that perception training improves perception and production (e.g. Akahane-Yamada et al., 1996; Rochet, 1995; Wong, 2013). However, the effect of perception training on word recognition has not been investigated. Two theoretical models that support the claim that improved L2 segmental perception improves L2 word recognition are Shortlist B (Norris and McQueen, 2008), a Bayesian model, and the revised L2LP model (Van Leussen and Escudero, 2015). Accurate L2 word recognition can fail at the segmental level, where misheard phonemes can cause the merge of minimal word pairs or spurious lexical activation (e.g. Broersma, 2002; Broersma and Cutler, 2011). Improved segmental perception mitigates segmental confusion. Thus, speech perception training could improve L2 segmental perception, in turn facilitating L2 word recognition. Sixty L1 Dutch students were tested (test group N = 30 and control group N = 30) in a classroom setting (rather than in a laboratory). Both groups took a pretest and posttest, and the test group completed five training sessions. Each training session consisted of the same thirty-eight L2 English phonemes that are (relatively) difficult for Dutch learners, embedded in word contexts. Results indicate that participants in the test group improved in their perception but not in their word recognition. Improvement in participants' speech perception in the classroom indicates that perception training is ecologically valid, and that it could therefore be a useful tool in the classroom to improve Dutch learners' English speech perception. The lack of improved word recognition could be due to various reasons. One reason is the nature of the task; ecological validity (Goldinger, 1996) or the ease of the task could have been an issue. Another reason is L1 interference; perhaps this interference is too robust. Finally, perhaps the distance in conceptual space is too large, i.e., word recognition is too distant from segmental perception. Many steps are involved in accurate word recognition, namely from an acoustic sound to a phonetic output to a phonological form to a lexical item (Van Leussen and Escudero, 2015).

From reluctance to confidence: A longitudinal study of Japanese students' foreign language speaking fluctuations in Europe

Simon Humphries (Kansai University).

In Japan, classroom observations indicate that students often remain silent during English as a Foreign Language (EFL) classes (King, 2013). Potential reasons for this reticence include demotivating instruction (Kikuchi, 2009), examination-focused study (Sakui, 2007), teacher-led norms (Nishino & Watanabe, 2008), hypersensitivity to others (King, 2013), linguistic problems (Harumi, 2011), and uninteresting topics (Yashima, Ikeda & Nakahira, 2016). Although this image prevails of silent and undermotivated students in English classes, a different picture may emerge among Japanese learners of languages other than English (LOTE). Unlike “Global English,” which can be considered a fundamental skill (Graddol, 2006) or even an economic commodity (Ushioda, 2017), LOTE-learning may have personal motivations tied to specific communities that own the language (Dörnyei and Al-Hoorie 2017). To date, Japanese education and research focus on L2 English study. This study explored patterns and factors influencing Japanese English majors' capacity to speak (CTS) in L2 English and L3 LOTE before and during a dual-language study abroad (SA) program in France and Germany. The dual language SA was an academic year consisting of L3 LOTE language skills classes and undergraduate courses conducted using English as a medium of instruction (EMI). Five female Japanese undergraduates—three studying L3 German and two studying L3 French—volunteered. They wrote monthly “focused essay” reflective journals (MacIntyre et al., 2011) that described successful or unsuccessful speaking experiences, and they sketched retrospective timeline graphs (Falout, 2016) that indicated fluctuations in their foreign languages' CTS. The researcher used the iterative-inductive constant comparison memo-writing approach (Corbin & Strauss, 2008) to analyse emerging data patterns and to guide pre- and mid-SA interviews. Results indicated fluctuations in their motivation and speaking attitudes. Before SA, the participants had a mixture of vague instrumental and integrative orientations to study and travel in Europe, influenced by their parents who had experienced France or Germany. During SA, they struggled to speak their L3 with local people in spontaneous daily interactions; however, their L3 classes and dormitories exposed them to international students who used English as a lingua franca to communicate. The participants also tended to make friends with local French and German students who were English majors with shared topics of interest. Therefore, although they persevered and improved in their L3 LOTE, English became their social foundation. The participants displayed growth mindsets (Dweck, 1999) to develop strategies to overcome situations with low CTS. In Japan, their sensitivity to others reduced their CTS due to the fear of making mistakes or offending others; however, in Europe, the participants discovered they could take the risk of speaking their foreign languages without jeopardising their relationships. The presenter will discuss the implications of these findings for supporting East Asian students in adapting to European SA programs.

The role of linguistic distance and language use on L2/L3 English proficiency

Nils Jaekel (University of Copenhagen / University of Oulu), Eliane Lorenz (Justus Liebig University Gießen) and Markus Ritter (Ruhr-University Bochum).

Successful acquisition and high command of English as a Foreign Language play a crucial role almost everywhere in the world. Zooming in on Germany, English is typically the first foreign language studied, introduced in primary school, and enjoys extremely high prestige (e.g., Siemund, 2018). Due to increasing globalization, migration, and mobility, (foreign) language classrooms are characterized by heterogeneity and include both monolingually as well as bi/multilingually raised children (e.g., Cenoz & Gorter, 2015; Gogolin, 2021). Moreover, social and individual factors such as socio-economic status or cognitive ability are known to influence the acquisition of foreign languages (Diehl et al. 2016; Berthele & Udry, 2021). However, there is limited knowledge of how language use outside of school impacts the acquisition of English. The current study addresses this research gap by investigating language use and linguistic distance (LD) of the languages spoken by 2,019 secondary school students in Germany. The students come from a linguistically heterogeneous population (from German as their only home language to multiple home languages) and attended grade five at the time of testing. While controlling for biological sex, country of birth, cognitive ability, as well as cultural capital of the family, the study assesses students' performance in English reading and listening tests. Mixed-effects models indicate that LD – operationalized as a continuous measure, based on cognate similarity (Wichmann et al., 2020; see also Schepens et al., 2013) – to English of the home language(s) as well as of the language(s) spoken to friends outside of school were negatively associated with the English listening comprehension scores, but it did not reach statistical significance for the English reading task. These results are critically discussed in light of curricula goals for language learning while considering the tension between approaches valuing diversity, such as pedagogical translanguaging, and the prevalent practiced monolingual habitus in foreign language classrooms in Germany.

Berthele, R. & Udry, I. (eds.). (2021). Individual differences in early instructed language learning: The role of language aptitude, cognition, and motivation. Berlin: Language Science Press. Cenoz, J. & Gorter, D. (eds.). (2015). Multilingual education. Cambridge: Cambridge University Press. Diehl, C., Hunkler, C., & Kristen, C. (2016). Ethnische Ungleichheiten im Bildungsverlauf. Eine Einführung. In *Ethnische Ungleichheiten im Bildungsverlauf. Mechanismen, Befunde, Debatten*, C. Diehl, C. Hunkler & C. Kristen (eds.), 3–31. Wiesbaden: Springer. Gogolin, I. (2021). Multilingualism: A threat to public education or a resource in public education? – European histories and realities. *European Educational Research Journal*, 20, 297–310. Schepens, J., van der Slik, F., & van Hout, R. (2013). Learning complex features: A morphological account of L2 learnability. *Language Dynamics and Change*, 3, 218–244. Siemund, P. (2018). Modeling World Englishes from a cross-linguistic perspective. In *Modeling World Englishes: Assessing the interplay of emancipation and globalization of ESL varieties*, S. C. Deshors (ed), 133–162. Amsterdam: Benjamins. Wichmann, S., Brown, C. H., & Holman, E. W. (eds.). (2020). The ASJP Database. Max Planck Institute for the Science of Human History. <https://asjp.cld.org/>

Third language acquisition in high intercomprehension settings: The case of displaced Ukrainians learning Polish

Andrzej Jarynowski (Freie Universität Berlin), Karolina Czopek (Institute of English Studies, University of Warsaw), Michał B. Paradowski (Institute of Applied Linguistics, University of Warsaw) and Andrea Palmi (Institut für Veterinär-Epidemiologie und Biometrie, Freie Universität Berlin).

Since the beginning of the Russian invasion of Ukraine on 24 February 2022, 11.9m refugees crossed the country's border into Poland. We investigate peer learner networks of 251 participants in an intensive course of the Polish language dedicated to the newcomer population. The participants came mainly from eastern, southern and northern Ukraine (the further east the origins, the higher the proportion of declared use of Russian). Apart from the special situational context, together with the close typological similarity between the languages spoken and being acquired, the students present a unique language constellation profile, with most being functionally bilingual in Ukrainian and Russian, but with different degrees of dominance in each language and complicated attitudes to the latter. Using a pre-/post-test design and a specially devised survey, we apply the tools of computational social network analysis to find out whether and how patterns of out-of-class communication within and beyond the cohort affected linguistic gains. The refugee students are most satisfied with their communication in Polish with neighbours and volunteers, less so in service encounters and the workplace, and the least in the administrative sphere. Polish language use was highest in text messages. The students self-rated their progress best in vocabulary and lowest in grammar. While dominance in Ukrainian vs Russian did not affect progress in the TL, speaking Ukrainian correlated with centrality in the contact network. Russian speakers often concealed their L1 use (62% of users of Russian in the private sphere declared Ukrainian as their L1). A reconstruction of the student networks shows higher weighted degree centrality among students declaring Ukrainian as their L1, while L1 Russian speakers are at the network periphery, suggesting linguistic segregation with symptoms of marginalisation. The most influential significant predictors of self-perceived progress overall and in grammar were level of course enjoyment and two social network measures: the degree of being indicated as interlocutors by well-connected students (pagerank) and degree of interaction with Russian-speaking friends. Objectively measured progress, however, instead hugely negatively correlated with length of residence in Poland.

Paradowski, M.B., Czopek, K., Palmi, A. & Jarynowski, A. (under review). Language learning in forced immersion: Proximal sociolinguistic context, peer networks, and progress in Ukrainian refugee students' acquisition of L3 Polish. *Language Learning* (Special Issue: Social Aspects in Language Learning: New Perspectives from Study Abroad Research).

Supporting Second Language Learners in Science Education: Insights from a Science and Literacy Teaching Project

Sofie Johansson (Dept. of Swedish, Multilingualism, Language Technology, University of Gothenburg) and Clas Olander (Dept. of Natural Science, Mathematics and Society, Malmö university).

Due to the recent surge of immigration to Sweden, it is estimated that approximately 25% of students aged 7-19 are second language learners. This heterogeneous group of students poses a challenge for teachers across all subjects, due to diverse individual differences, varying levels of language proficiency and the unique linguistic characteristics of each subject.

Enhancing science literacy can be facilitated by incorporating a focus on comprehending words in context, as proposed by Logan and Kieffer (2021). Furthermore, functional science literacy, as defined by Norris and Phillips (2003), encompasses not only an understanding of science terms, but also the ability to effectively communicate, read, and write using these terms in non-technical settings.

The project, Science and Literacy Teaching (SALT), funded by the Swedish Research Council, attempted to identify the best strategies for supporting science teachers in working with second language learners.

A model of word types was constructed, drawing upon the categorization of words into everyday common words, academic words, and scientific words, as proposed by scholars such as Nation (2013), Hulstijn (2012) among others.

In order to assess the vocabulary difficulties faced by both first language (L1) and second language (L2) students, a sample of 232 participants was administered vocabulary tests that consisted of seven types of words. Three of these word types were neutral in terms of subject matter, while the other four were subject-specific and drawn from textbook texts in the fields of biology, physics, and chemistry. The tests were designed to examine students' receptive knowledge of the words, including their forms, meanings, and usage in authentic contexts. The tests involved matching test words with their correct meanings or synonyms in order to evaluate the depth of students' knowledge.

We found that there are four types of word groups that require special attention from teachers. In summary, L2-students had particular challenges with general language comprehension, such as academic words and homonymous words with subject-specific meanings. These types of words were not difficult for L1 students. Finally, all students had difficulty with homonymous words with academic and subject-specific meanings, as well as general subject-related words that were new to all students.

The research sheds light on the difficulties faced by second language learners in science education and informs strategies for their support, contributing to the field of multilingualism.

Hulstijn, J.H. (2012). The construct of language proficiency in the study of bilingualism from a cognitive perspective. *Bilingualism*, 15(2), Cambridge. Logan, J. K., & Kieffer, M. J. (2021). Investigating the longitudinal and concurrent relationships between polysemous word knowledge and reading comprehension among Spanish-English bilingual middle school students. *Reading and Writing*, 34, 301-335. Nation I.S.P. (2013), *Learning Vocabulary in Another Language*. Cambridge. Applied Linguistics. Cambridge University Press. Norris, S. P., & Phillips, L. M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Science education*, 87(2), 224-240.

Synchronization of tone and intonation in the production of Hungarian learners of Mandarin Chinese

Kornélia Juhász (Eötvös Loránd University, Hungarian Research Centre for Linguistics) and Huba Bartos (Hungarian Research Centre for Linguistics).

In this acoustic analysis we aim to compare Mandarin Chinese (MC) syntactically marked (trisyllabic) and unmarked (disyllabic) interrogative and declarative intonation patterns in the production of two L2 learners' groups with different levels of language experience, whose L1 is atonal Hungarian. Two types of tone sequences were analysed: 'level Tone 1 + rising Tone 2' and 'level Tone 1 + falling Tone 4' utterances, complemented by a toneless question particle *ma* in trisyllabic marked questions. While both L1 and L2 declaratives feature a descending pattern [1, 2], MC interrogatives are characterized by an elevated f_0 compared to declaratives throughout the whole utterance [1], but in Hungarian, questions feature a rising-falling contour [2]. We hypothesized that due to the different interrogative character contours in L1 and L2, and the absence of tones in L1, the synchronization of tone and intonation in contrasting the two sentence types would pose problems for L2 learners. In particular, we expected that if the analysed L2 tonal patterns sharply differ (i.e., are inversely related) to the native L1 pattern, then L2 learners favour the L1 interrogative character contour due to direct L1 transfer [3]. We recorded short dialogues of interrogative and declarative sentences, and the extracted f_0 contours were compared by GAMMs [4]. Our results showed that L2 tone sequences containing T1 + T2 posed more difficulties for L2 learners than those with T1 + T4. Our hypothesis regarding the presupposed difficulties caused by L1 transfer effects were partially confirmed, although L2 learner groups with different levels of L2 experience differed in producing L2 interrogative patterns: Beginners produced patterns similar to the L1 interrogative character contour in both unmarked and marked questions. Advanced learners, on the other hand, only produced L1-like patterns in the case of unmarked T1 + T2 questions, while in all other cases they managed to approximate the native L2 contour, except that they failed to produce the elevated f_0 pattern in each type of interrogatives. Our results contribute to the deeper understanding of the acquisition of a tonal L2.

[1] Shen, X. 1990. *The Prosody of Mandarin Chinese*. University of California Press. [2] Varga, L. 2016. The Intonation of Topic and Comment in the Hungarian Declarative Sentence. *Finno-Ugric Languages and Linguistics* 5/2, 46–77. [3] Viger, T. 2007. *Fundamental Frequency in Mandarin Chinese and English: Implications for Second Language Speakers*. PhD dissertation. The City University of New York. [4] Wood, S. N. 2017. *Generalized Additive Models: An Introduction with R*. Chapman and Hall.

Types of L2 input and instructional techniques predict novice and more advanced L2 learners' proficiency differentially

Kristin Kersten (Mannheim University).

While input quality is a core construct in Instructed Second Language Acquisition (ISLA), its operationalization is challenging. Truscott & Sharwood Smith (2019) define input as all sensory stimuli in the environment that induce intake, processing and storing, and thus contribute to language comprehension. Based on this comprehensive definition, the Teacher Input Observation Scheme (TIOS, Kersten et al., 2018) operationalizes 'input' through 41 instructional techniques in four scales, 'Cognitively Stimulating Tasks', 'Verbal Input', 'Nonverbal Input', 'Support of Learners' Output'. Techniques are derived from the Cognitive-Interactionist Framework (Long, 2015), and describe a distinct, observable communicative behavior or activity in the classroom in contact with the learner/s. Previously established interrater reliability of two raters and 18 L2-teaching videos was high (Krippendorff's $\alpha=.88^*$). In a pilot study with N=169 German primary school students, 'Total Score', 'Task Characteristics' and 'Verbal Input' correlated significantly with learners' L2 lexical and grammar comprehension; the 'Total Score' predicted 21% of variance in the data (Kersten, 2021).

The TIOS was constructed based on theoretically derived categories. However, it does not account for teachers' differential use of techniques for novice versus advanced L2 learners (van de Pol et al., 2010, Tedick & Lyster, 2020).

The current study addresses this gap. It investigates N=94 teachers' L2 instructional techniques, focusing on their differential predictions on the L2-English proficiency of their learners' (N=1935, grades 4-6) in German classrooms. Techniques were operationalized using TIOS items, learners' L2 proficiency using descriptors from the CEFR and Lingualevel (9 levels, A1.1-C1, M=4.38; SD=2.11).

Data were elicited online via SoSciSurvey from the teachers. They reported on their use of TIOS-techniques in a particular English class, and on their students' L2 proficiency in this class, their gender, multilingual background, and average class social status. A principal-axis factor analysis was carried out to identify related TIOS-techniques, using oblimin rotation because factors were expected to correlate.

The analysis yielded five factors (36.36% of variance). Two factors were hypothesized to be used with novice learners ('Input Scaffolding', 'Speech Scaffolding & Appreciation'), two with advanced learners ('Cognitive Stimulation', 'Focus on Meaning'). For the factor 'Comprehension & Expression Scaffolding', no hypothesis was formulated. Due to the lack of directionality in cross-sectional designs, negative effects in regression analyses suggest an association with learners' low L2-proficiency, positive effects with high L2-proficiency.

Means over each factor's items were entered into regression analyses (dependent variable: learners' L2-proficiency), controlling for grade, gender, multilingualism, and social status ($R^2=.168$). As hypothesized, results showed that the two "novice" factors showed negative predictions, the two "advanced" factors positive ones, as did social status and grade (all $p<.001$). 'Comprehension & Expression Scaffolding', multilingualism and gender did not have significant effects.

Negative effects suggest that, in early stages of SLA, teachers predominantly use scaffolding techniques which render language and content comprehensible, while techniques activating cognitive involvement and authentic interactions foster SLA in later stages (positive effects). Results support concepts of "fading support" and "transfer of responsibility" in van de Pol et al.'s (2010) Conceptual Model of Scaffolding and Tedick & Lyster's (2020) CAPA Model.

Out of sight, out of mind: Investigating the role of salience in the initial processing of morphology in SLA

Georgia Knell (Vrije Universiteit Brussel) and Saioa Cipitria (Vrije Universiteit Brussel).

Despite growing theoretical support for the importance of salience of L2 grammatical forms (i.e., the extent to which they stand out from their linguistic environment) on learners' acquisition thereof, there remains "a lack of systematic...research comparing acquisition difficulty for morphemes with different degrees of salience" (DeKeyser, 2015:15). To address this gap, this study is the first in a series of experiments that directly investigate different salience manifestations (categorized as perceptual, psycholinguistic, and experiential; Ellis, 2016) to measure their causal contribution to learners' attention, awareness, and consequent acquisition of L2 morphological forms. Sixty L1 Dutch participants conducted a reading task in an English-based semi-artificial language called Englishti, which includes in its grammar two novel target morphemes, perceptually low-salient -o and high-salient -ulp according to orthographic length (Simoens et al., 2018). The task consists of two phases: during the learning phase, participants receive input flooding of the target morphemes while reading short Englishti texts, after which they are asked a content question related to the story. During the testing phase, participants are asked to make grammaticality judgments about individual Englishti sentences, some familiar from the learning phase and some new. Eye movement data will be analyzed in terms of target morpheme fixation time in both phases to measure level of attention, as well as comparisons of fixations, first run dwell time, total time and rereading time of grammatical versus ungrammatical trials in the test phase using a grammatical sensitivity index (Granena, 2013) to measure implicit learning of the rule. A retrospective interview determines participants' level of awareness of the form and its meaning. We further analyze the interaction between salience effects and individual learner variables, including working memory and implicit learning ability. Finally, we also consider the influence of incidental versus intentional learning contexts, with the intentional group receiving instructions guiding participants toward attention to the grammar of Englishti while the incidental group does not. We hypothesize that participants will exhibit greater degrees of attention, awareness, and acquisition of -ulp compared to -o, after controlling for the effects of individual learner variables and learning context. We further expect better overall learning outcomes for both morphemes among the intentional learner group than the incidental learner group. The results of this study will act as the first step in a larger project comprising a systematic empirical investigation of the effect of various types of linguistic salience on SLA.

References DeKeyser, R. (2015). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in Second Language Acquisition. An introduction* (pp. 94-112). London: Routledge. Ellis, N. (2016). Salience, cognition, language complexity, and complex adaptive systems. *Studies in Second Language Acquisition*, 38(2), 341–351. Granena, G. (2013). Individual differences in sequence learning ability and second language acquisition in early childhood and adulthood. *Language Learning*, 63(4), 665–703. Simoens, H., Housen, A., & De Cuypere, L. (2018). The effect of perceptual salience on processing L2 inflectional morphology. In S. M. Gass, P. Spinner, & J. Behney (Eds.), *Salience in Second Language Acquisition* (pp. 107–130). New York: Routledge.

Evaluating a Malay-English and Mandarin-English cognate list and the cognate effect through bidirectional translation tasks

Ying Xian Koh (University of Nottingham, Malaysia), Csaba Szabo (University of Nottingham, Malaysia) and Jessica Price (University of Nottingham, Malaysia).

Evidence from pedagogical-oriented cognate studies suggests using cognates can facilitate second language acquisition (e.g., García et al., 2020; Urdaniz & Skoufaki, 2019; Vidal, 2011). Understanding how cognates are processed by L2 English speakers provides a better position for educators and researchers to examine the effectiveness of translanguaging across the curriculum in bi- and multilingual contexts (e.g., Cenoz et al., 2021; Ticheloven et al., 2021). Given the lack of research investigating the cross-linguistic similarity at the lexical level between English and Malaysian vernaculars, a cognate list between English-Malay and English-Mandarin was developed for the 10,000 most frequent English words as defined by the BNC/COCA word list (Nation, 2012). Based on the proportion and distribution of cognates between English-Malay and English-Mandarin, this study explored the extent to which cognates were produced as translations and whether cognates were translated more accurately than non-cognates. In addition, the impact of other item-related and participant-related factors on translation accuracy were examined. Proficient Malay-English bilinguals (N = 52) and Mandarin-English bilinguals (N = 50) completed a forward and backward translation task with 400 Malay words and 170 Chinese words, which controlled for cognateness, word length, and word frequency. The results indicated that English-Malay cognates were translated more accurately and were less translation-ambiguous than non-cognates in both translation directions. For English-Mandarin cognates, the cognate advantage in translation accuracy was detected only in the backward translation direction. Multiple regression analyses indicated cognateness and word frequency to be significant predictors for translation accuracy. Cognateness was found to explain additional variance for both directions in the Malay translation task, and only for the backward direction in the Mandarin translation task. This study provided evidence that cognates are processed more easily than non-cognates in translation tasks, regardless of the word length. This accentuates the importance of the cognate list developed and the possibility to extend its usefulness in further empirical investigation to develop more effective teaching practices and material design for the ESL context.

Integrating into the Host Country: The Role of Initial Proficiency and Individual Differences

Zeynep Köylü (University of Basel).

The discussion that linguistic, personal, and sociocultural gains after studying abroad are largely influenced by several individual difference (ID) variables (Hessel, 2017), such as cognition (Segalowitz & Freed, 2010), motivation (Hernández, 2010), and personality (Baker-Smemoe et al., 2014) is well-documented in the literature. Some program-related variables, such as length of stay, type of accommodation, and social-networking styles have also been investigated in terms of language gains. These studies focus on single variables and their development (e.g., Tracy-Ventura et al., 2016), or their particular influence on gains, while only a few considers the predictive power of IDs on a sojourn experience (Baker-Smemoe et al., 2014). Yet, we need further research to explore the combined effects of such ID variables on increased interaction and consequently linguistic and intercultural gains of sojourners. This study aims to explore the role of initial proficiency in English, and several ID variables, such as personality, intercultural awareness, and mental and emotional wellbeing on language and intercultural gains and sociocultural adaptation after a semester abroad. It also investigates the complex relationship among these variables and their predictive power on language and intercultural gains. As part of a larger project, this mixed-methods study analyzes data from a group of tertiary level sojourners (n = 300) who participated in the ERASMUS program for a semester. The data were elicited via an online version of the Oxford Quick Placement Test, the short form of the Multicultural Personality Questionnaire (MPQ, Van Oudenhoven & Van der Zee, 2013), the short Big Five Personality Inventory (Rammstedt & John, 2007), the Warwick-Edinburgh Mental Well-being Scale (Tennant et al., 2007) and the Sociocultural Adaptation Scale (Ward & Kennedy, 1999) twice at pre-departure and immediately post-sojourn. A subgroup of participants with varied pre-departure proficiency, MPQ, and personality types were also interviewed following a semi-structured interview pertained to their experiences abroad in a way to triangulate our quantitative findings. The data from the ID variables were analyzed through a series of linear mixed effect models to understand how fixed (Time) and random (participants) effects interact to predict participants' language and intercultural gains and sociocultural adaptation abroad. The preliminary results suggest that an intermediate level of proficiency in English is significant to initiate interactions, and adapt the host environment, as well as a basic level of intercultural awareness in the form of openness to diversity. The results are anticipated to bring further insights into how future sojourners could make the most of their stays abroad for linguistic, personal, social, and cultural gains. Baker-Smemoe, W., Dewey, D. P., Bown, J., & Martinsen, R. A. (2014). Variables affecting L2 gains during study abroad. *Foreign Language Annals*, 47(3), 464-486. Hernández, T. A. (2010). The relationship among motivation, interaction, and the development of second language oral proficiency in a study-abroad context. *The Modern Language Journal*, 94(4), 600-617. Hessel, G. (2017). A new take on individual differences in L2 proficiency gain during study abroad. *System*, 100(66), 39-55.

Rammstedt, B. & John, O. P. (2007). Measuring personality in one minute or less: A 10 item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, 41, 203-212. Segalowitz, N., & Freed, B. F. (2004). Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in second language acquisition*, 26(2), 173-199. Van der Zee, K., Van Oudenhoven, J. P., Ponterotto, J. G., & Fietzer, A. W. (2013). Multicultural Personality Questionnaire: Development of a short form. *Journal of personality assessment*, 95(1), 118-124. Ward, C., & Kennedy, A. (1999). The measurement of sociocultural adaptation. *International journal of intercultural relations*, 23(4), 659-677.

Insights from multimodal analysis on longitudinal interview data: the case of I don't know

Pascale Leclercq (Université Paul Valéry Montpellier 3), Amanda Edmonds (Université Côte d'Azur), Elisa Sneed German (Université Paul Valéry Montpellier 3) and Pauline Beaupoil-Hourdel (Sorbonne Université, INSPE de Paris).

Discourse markers (DM) are usually defined as expressions with little referential meaning but multiple pragmatic functions (Pichler, 2009, p. 561). One such DM is the expression I don't know (IDK), which is widely used by native speakers and learners alike (Baumgarten & House, 2010; Bybee & Scheibman, 1999). However, there is currently no consensus on the different functions that this DM can take in interactional discourse. Proposed functions include floor holding, topic closing, prefacing, opting out of the interaction, hedging, etc. (e.g., Pekarek Doehler, 2016), although the comparability and reliability of coding schemes for such functions is unclear. This may be in part because existing studies have largely relied on audio files or transcriptions. We hypothesize that coding may be rendered more reliable if multimodal sources of information are used to disambiguate pragmatic functions (Mondada, 2017). In particular, shrugs are frequently associated with the use of IDK (Beaupoil-Hourdel & Morgenstern, 2021, Debras, 2017), while gaze provides interesting clues as to how speakers and listeners manage interactions (Rossano 2012). In this paper, we used a case-study approach to investigate the evolution of the use of IDK by five L2 learners of English over a year abroad, a learning context demonstrated to increase DM use (Tavakoli, 2018). Five filmed semi-guided interviews were conducted over the course of 12 months. The resulting 12 hours of recordings were transcribed, and 424 instances of IDK were identified. We coded them for pragmatic function, drawing on both linguistic and multimodal cues, taking into consideration the speaker's body posture, gestures (shoulder-lifts, palm-ups associated with lifted eyebrows and/or a mouth shrug), gaze and the prosodic contours of their vocal productions. The coding process revealed that multimodal evidence was essential in identifying pragmatic functions in more than half of all instances. On the basis of these findings, we argue for the importance of using a variety of evidence sources for pragmatic coding.

Baumgarten, N., & House, J. (2010). I think and I don't know in English as lingua franca and native English discourse. *Journal of Pragmatics*, 42, 1184-1200.

Beaupoil-Hourdel, P., & Morgenstern, A. (2021). French and British children's shrugs: A cross-linguistic developmental case-study of a recurrent gesture. *Gesture*, 20(2), 180-218.

Bybee, J., & Scheibman, J. (1999). The effect of usage on degrees of constituency: The reduction of don't in English. *Linguistics*, 37(4), 575-596.

Debras, C. (2017). The shrug. *Gesture*, 16(1), 1-34.

Mondada, L. (2017). Le défi de la multimodalité en interaction. *Revue française de linguistique appliquée*, 2, 71-87.

Pekarek Doehler, S. (2016). More than an epistemic hedge: French *je sais pas* 'I don't know' as a resource for the sequential organization of turns and actions. *Journal of Pragmatics*, 106, 148-162.

Pichler, H. (2009). The functional and social reality of discourse variants in a northern English dialect: I DON'T KNOW and I DON'T THINK compared. *Intercultural Pragmatics*, 6(4), 561-596.

Rossano, F. (2012). Gaze behavior in face-to-face interaction. Doctoral Thesis, Radboud University Nijmegen.

Tavakoli, P. (2018). L2 development in an intensive study abroad EAP context. *System*, 72, 62-74.

'Good-enough' processing and L2 comprehension: Evidence from Korean

Chanyoung Lee (Yonsei University), Gyu-Ho Shin (Palacký University Olomouc) and Boo Kyung Jung (University of Pittsburgh).

The 'good-enough' processing account argues that the processor favors less-costly and more-accessible options from memory (*heuristics*) over structure-based, detailed computations (*algorithms*) in coping with incoming/upcoming information in real-time (Christianson,2016; Ferreira,2003). L2 studies join this account (Lim&Christianson,2015; Tan&Foltz,2020), pointing to the 'noisier' nature of L2 knowledge due to between-language competition and increased cognitive load in executing L2 activities at hand (Futrell&Gibson,2017; Pozzan&Trueswell,2016; Tachihara&Goldberg,2020). With that, we ask how L1-Czech (SVO but free word-order; case via inflection; active agreement) L2-Korean (SOV; case-marking; suffixation; scrambling/omission of sentential components) learners, a unique L1-L2 pair, achieve the 'good-enough' comprehension in Korean, focusing on morphological causative (MC) and suffixal passive (SP) constructions (Table.1). They are distinctive regarding the alignment between thematic roles and case-marking (MC: typical/frequent; SP: atypical/infrequent) and verbal morphology (MC: valency-increasing; SP: valency-decreasing). The two patterns in each construction differ concerning word order (Verb-Final: canonical/frequent; Verb-Initial: non-canonical/infrequent) and verbal morphology (Verb-Final: late-arriving cue[®] checking (MC) or revising (SP) the initial interpretation; Verb-Initial: early-arriving cue[®] guiding the following interpretation) (Table.2).

28 L1-Czech learners (CZH; $M_{age}=24.1$, $SD=2.8$) and 40 native speakers (NSK; $M_{age}=23.6$; $SD=4.1$) of Korean joined two tasks: self-paced reading (SPRT; involving a moment-by-moment, sequential and cumulative interpretation of the incoming items with time constraints) and acceptability judgment (AJT; involving both partial and holistic considerations of a sentence, with less time constraints, to arrive at its complete interpretation). Thirty-two sentences (eight instances per condition; Table.3) and fillers were split into four sub-lists and were randomly assigned to participants. CZH's proficiency in Korean was measured separately (Lee-Ellis,2009); they passed the threshold score (63/188). The trimmed data from each task were submitted to the respective linear mixed-effects models (fixed effects: *Group*, *Construction*, *Canonicity*; random effects: *Participant*, *Item* [AJT] or *Word* [SPRT]).

Results/AJT (Figure.1) CZH rated all conditions less acceptable than NSK ($Group^{***}$) but dispreferred Verb-Initial relative to Verb-Final (like NSK) ($Canonicity^{***}$). Results/SPRT (Figure.2) Global ($\alpha=.05$): CZH spent more time reading each region than NSK ($Group^{***}$). We also found a two-way interaction ($Group*Canonicity$) at $R2^{***}$, a three-way interaction ($Group*Construction*Canonicity$) at $R3^*$, and a two-way interaction ($Group*Canonicity$) at $R5^{***}$. Post-hoc analysis ($\alpha=.025$): NSK spent more time reading Verb-Initial than Verb-Final at $R3^*/R5^{**}$ only in SP. Whereas CZH spent more time reading Verb-Initial than Verb-Final at $R2^*$, they took more time reading Verb-Final than Verb-Initial at $R5^{**}$, both only in SP.

NSK's reading-time patterns show their primary reliance on heuristics (word-order typicality; frequent alignment between thematic roles and case-marking) than algorithms (expected processing benefit from the early-arriving verbal morphology). The finding that CZH's by-construction reading-time patterns considerably differed from NSK's implies that they may have conducted algorithmic parsing (driven by verbal morphology) to conduct necessary treatments involving this algorithmic cue in the course of moment-by-moment interpretation. Together, our findings suggest that (i) L2 processing can be 'good-enough', as shown by the CZH's acceptability ratings for Verb-Final (subject-first & predicate-final; consistent with canonical/frequent word order) than Verb-Initial, irrespective of construction type, and (ii) various factors can influence the degree to which the L2 processor conducts 'good-enough' comprehension. (word count: 499 / 500)

Table 1. Morphological causative & suffixal passive (canonical pattern) in Korean

Construction	Morphological causative (MC)	Suffixal passive (SP)
Example (canonical pattern)	Cinwu-ka Minci-hanthey Cinwu-NOM Minci-DAT umsik-ul mek-i-ess-ta. food-ACC eat-CST-PST-SE 'Cinwu made Minci eat food.'	Cinwu-ka Minci-hanthey Cinwu-NOM Minci-DAT cap-hi-ess-ta. catch-PSV-PST-SE 'Cinwu was caught by Minci.'
Word order	Agent-Recipient-Theme	Theme-Agent
Case-marking	Typical (NOM=agent; DAT=recipient)	Atypical (NOM=theme; DAT=agent)
Verbal morphology	-i/hi/li/ki/wu/kwu/chwu- (under allomorphic distribution)	-i/hi/li/ki- (under allomorphic distribution)

Note. The dative marker *-hanthey* is used more frequently than *-eykey* in colloquial contexts.

Table 2. Verb-final/initial patterns of two construction types used in this study

Cx	Pattern	Scheme	Word order	Case-marking facts	Verbal morphology
MC	Verb-Final	N-NOM + N-DAT + V-CST	Frequent	Typical	Late-arriving
	Verb-Initial	V-CST + N-NOM + N-DAT	Infrequent	Typical	Early-arriving
SP	Verb-Final	N-NOM + N-DAT + V-PSV	Frequent	Atypical	Late-arriving
	Verb-Initial	V-PSV + N-NOM + N-DAT	Infrequent	Atypical	Early-arriving

Table 3. Scheme of test sentences: SPRT (R2-R4: critical; R5: spill-over)

Cx	Pattern	R1	R2	R3	R4	R5	R6
MC	Verb-Final	<i>that N3-ACC,</i>	N1-NOM	N2-DAT	V-CST	<i>yesterday</i>	<i>night</i>
	Verb-Initial		V-CST	N1-NOM	N2-DAT		
SP	Verb-Final	<i>I heard that</i>	N1-NOM	N2-DAT	V-PSV		
	Verb-Initial		V-PSV	N1-NOM	N2-DAT		

Note. To precisely conduct region-by-region comparisons across the conditions, we topicalized the direct object of the MC instances as in R1. Creation of AJT items: R2+R3+R4 for SP; R2+R3+N3+R4 (Verb-Final) or R2+R3+R4+N3 (Verb-Initial) for MC.

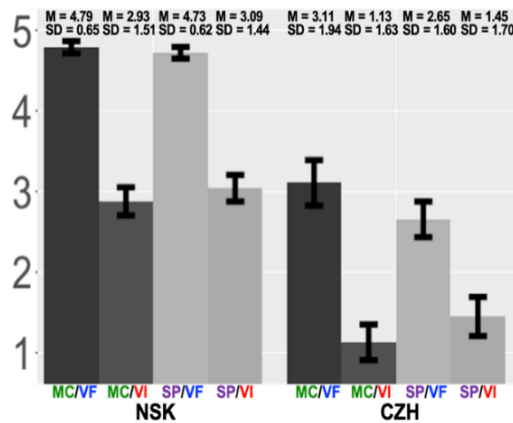


Figure 1. Results: AJT. X-axis = group/condition; Y-axis = rating (trimmed raw score; 6-point Likert scale). Error bar = 95% CI. Data trimming: any response with response times below 1000 ms or above 10000 ms was excluded.

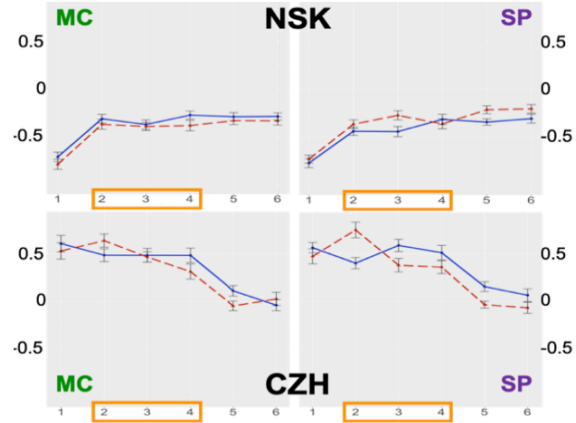


Figure 2. Results: SPRT. Blue = Verb-Final; Red = Verb-Initial; X-axis = region; Y-axis = reading time (log-transformed & residualized). Error bar = 95% CI. Data trimming: Data failing to pass comprehension questions & whose reading times below or above the 3SD were excluded.

Abbreviations. ACC = accusative case maker; CST = causative suffix; DAT = dative marker; N = noun; NOM = nominative case marker; PST = past tense marker; PSV = passive suffix; SE = sentence ender; V = verb.

The role of differential cross-linguistic influence and other constraints in L2 predictive gender processing

Tekabe Legesse Feleke (Arctic University of Norway).

Previous studies on linguistic prediction in a second language have shown differences between native and nonnative prediction. For example, L2 studies on the use of morphosyntactic gender cues for linguistic prediction show that nonnative speakers usually exploit grammatical gender information differently from native speakers, and that their ability to use gender information essentially depends on factors such as L1-L2 similarity, L2 fluency, knowledge of lexical gender and other constraints. The present study investigated the role of three factors: 'differential cross-linguistic influences', i.e., influences of dialects of L1, knowledge of L2 lexical gender and L2 fluency based on data obtained from 120 L1 Oromo late L2 speakers of Amharic and 60 monolingual Amharic speakers, using visual world eye-tracking and of-line experiments. Half (60) of the L2ers were native speakers of the Eastern Oromo dialect, and the remaining half (60) were native speakers of the Western Oromo dialect. The two dialects differ in their grammatical gender; the Eastern dialect has highly productive agreement system that distinguishes between masculine and feminine gender, like the L2 Amharic. Due to history of contact with other languages, the distinction between masculine and feminine gender has been neutralised in the Western dialect. The study investigated the influence of these dialects, focusing on three Amharic gender agreement conditions: noun-attributive adjective, noun-demonstrative pronoun and noun-interrogative pronoun gender agreement. Between-group and across condition comparisons based on the proportion of eye fixation prior to the target nouns show that the monolingual and L1 Eastern L2 Amharic speakers employ gender cues to anticipate the upcoming nouns, but only advanced L1 Western L2 Amharic speakers could predict. Knowledge of lexical gender and L2 fluency moderate L2 predictive gender processing only if gender is not instantiated in the first language. The study has ecological relevance as it presents empirical data from underrepresented languages.

Dynamical Acquisition Theory: The emergence of intentionality and the simulation of SLA dynamics

Anke Lenzing (Innsbruck University), Manfred Pienemann (Paderborn University) and Howard Nicholas (LaTrobe University).

In this paper we present the key tenets of Dynamical Acquisition Theory (DAT), a novel theory of SLA that is based on the application of central mathematical concepts of Dynamical Systems Theory (DST) to the field of SLA. The application of the mathematical core of DST to SLA is distinct from existing DST approaches that are based on DST-inspired metaphors. We show that DAT enables a) predictions in specific domains of second language development and b) the modelling of specific aspects of the dynamics of the L2 acquisition process.

A precondition for applying DST to SLA is to clarify whether mental phenomena and physical phenomena are fundamentally different or whether the dynamics driving the two types of phenomena are sufficiently similar to permit DST-based mathematics to be applied to the modeling of (some aspects of) mental processes in SLA. In the philosophy of mind this issue is discussed under the heading “intentionality” (Bechtel, 1988).

In DAT, we align ourselves with Tschacher’s (2014) naturalization approach that focuses on explaining the emergence of intentional states out of the dynamics of the natural world. We argue that this understanding of intentionality imposes limitations on the domains of SLA that can be addressed using a mathematically-based approach to DST. These limitations constrain the potential applications of DAT to lower order mental states. In the case of SLA, these applications are restricted to selected procedural aspects of language processing.

We empirically demonstrate the potential of DAT by focusing on L2 simplification in equational sentences (Pienemann, Lanze, Lenzing & Nicholas 2022). We present an agent-based model of L2 simplification programmed in NetLogo (Wilenski 1999), which simulates the linguistic dynamics in the acquisition of equational sentences over time. The simulations generated by the model are then compared to data from six longitudinal studies of child and adult L2 learners. The data span observational periods between eight and eighteen months and include learners of English and German as L2 with different L1s (Japanese, Turkish, English, Italian). The data include both learners who progress over time and learners who stabilize in their acquisition process. Overall, the simulations generated by our model closely match the developmental trajectories found in the longitudinal studies, which suggests that the restrictions and specification built into DAT underpin a viable means of predicting and modelling aspects of SLA.

Bechtel, W. (1988). *Philosophy of mind. An overview for cognitive science.* Lawrence Erlbaum.

Pienemann, M., Lanze, F., Nicholas, H. & Lenzing, A. (2022). Stabilization. A dynamic account. In A. Benati & J. Schwieter (Eds.), *Second language acquisition as shaped by the scholarly legacy of Michael Long* (pp. 29-76). Benjamins.

Tschacher, W. (2014). Intentionality: Steps towards naturalization on the basis of complex dynamical systems. In R. A. Meyers (Ed.), *Encyclopedia of complexity and systems science* (pp. 1-15). Springer.

Wilensky, U. (1999). NetLogo. <http://ccl.northwestern.edu/netlogo/>. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.

Reflections on transdisciplinarity and diversity in SLA studies

Alex Ho-Cheong Leung (Northumbria University, UK).

This presentation aims to extend discussions around the numerous “turns” in applied linguistics and second language acquisition, e.g. social- (Block 2003; Firth and Wagner 2007), multilingual- (Conteh and Meier 2014; May 2014; Meier 2017), transdisciplinary- (Douglas Fir Group 2016; Duff and Byrne 2019), and racial- turn (Hudley, Mallinson, and Bucholtz 2020; Kubota 2020). At the same time, it reflects on the implications of the multi-/trans- disciplinary interests in second language acquisition (SLA), which stimulate the departure from a monolithic conceptualisation of language (Hall and Wicaksono 2019), the re-examination of the monolingual basis (or bias) (Dewaele 2018; Genesee 2022) against which development and success are measured (Sakai 2018), and perhaps more fundamentally the interrogation of the nature of multilingual competence (Cook and Li 2016; Bassetti and Filipović 2022), among many of the recent developments. This presentation explores how, enabled by methodological and technological advances, researchers are encouraged to look for synergy across disciplines to generate insights that pertain to the nature of linguistic representation, processing, development, language use, and pedagogy (Atkinson 2011; Hulstijn et al. 2014; Leung and Young-Scholten 2013).

Subsequent to identifying the strengths and celebrating the diversity in field of SLA, the second half of this presentation underscores the need to enhance the empirical base of our field by including evidence from languages other than English (LOTE) (cf. Ushioda and Dörnyei 2017), utilising longitudinal designs, investigating diverse research population including younger children (Pinter 2014) and older adults (Nilsson, Berggren, Garzón, Lebedev, and Lövdén 2021) as well as participants from non-WEIRD (Western, Educated, Industrialised, Rich, Democratic) contexts (Andringa and Godfroid 2020; Henrich, Heine, Norenzayan 2010). Moreover, more efforts should be directed at and merits given to research replication (Porte 2012; Porte and McManus 2019), in line with recent calls of the Open Science movement to enhance research transparency, rigour, and accessibility (Marsden and Plonsky 2018; Marsden 2019; Marsden, Trofimovich, and Ellis 2019). This presentation argues that it is only by attending to the above are we able to continue to advance our field.

As the field of SLA sits on and benefits from intersections (e.g. between languages, cultures, experiences), the presentation concludes by calling for closer attention to be paid to ethical issues at all stages of SLA research (cf. De Costa, Sterling, Lee, Li, and Rawal 2021; Pinter 2022; Thomas 2009) and encouraging fellow researchers to contemplate the potential our field has in contributing to social justice (Hudley et al. 2020; Ortega 2005; 2019).

Using multiple-choice exercises for collocation learning: Do distractors linger in memory?

Mengxue Li (Western University) and Frank Boers (Western University).

Analyses of contemporary EFL textbooks have shown that collocations are often introduced through selected-response exercise formats—essentially multiple-choice tests—where the learners are required to choose from two or more options to form an appropriate word partnership (Boers et al., 2017; Strong & Boers, 2019). These exercises supposedly pique learners' curiosity, after which they will learn from the corrective or affirmative feedback on their responses. However, studies examining the effectiveness of such exercises have revealed that, when learners are tested on the same collocations after a week or so, they sometimes repeat the mistakes they made in the exercise or produce one of the other distractors they were presented with in the exercise (Boers et al., 2014; Boers et al., 2017; Stengers & Boers, 2016). These findings suggest that the distractors that learners are invited to consider may linger in memory and may hinder recall of the correct response later. While this seems plausible, it is also conceivable that learners may not even remember the exercise response options by the time they take the post-test, and the emergence of the same responses in the post-test is coincidental rather than a product of false memories. To find out if post-test errors can truly be attributed to the learners' consideration and recollection of distractors in the multiple-choice exercises, the present mixed methods study invited 20 advanced learners of L2 English to verbalize their thoughts (i.e., to think aloud) as they tackled such exercises and took a one-week delayed post-test. The targets for learning were 20 verb-noun collocations (e.g., run a business, break the silence, and carry a tune). The multiple-choice exercises were presented to them again after the post-test as a prompt for stimulated recall, to see if the learners had any episodic memories of how they had tackled certain exercise items and of the feedback received. Wrong post-test responses were then traced back not only to distractors in the exercise that were selected by the learners but also distractors that were not eventually selected but which were considered by the learners (according to the think-aloud and stimulated-recall data). The odds of wrong post-test responses under these different scenarios help to evaluate the merits of multiple-choice exercises for collocation learning and may inform ways in which the effectiveness of collocation exercises could be improved.

References

- Boers, F., Dang, T. C. T., & Strong, B. (2017). Comparing the effectiveness of phrase-focused exercises. A partial replication of Boers, Demecheleer, Coxhead, and Webb (2014). *Language Teaching Research*, 21(3), 362–280. <https://doi.org/10.1177/1362168816651464>
- Boers, F., Demecheleer, M., Coxhead, A., & Webb, S. (2014). Gauging the effects of exercises on verb-noun collocations. *Language Teaching Research*, 18(1), 54–74. <https://doi.org/10.1177/1362168813505389>
- Stengers, H., & Boers, F. (2015). Exercises on collocations: A comparison of trial-and-error and exemplar-guided procedures. *Journal of Spanish Language Teaching*, 2(2), 152–164. <https://doi.org/10.1080/23247797.2015.1104030>
- Strong, B., & Boers, F. (2019). The error in trial and error: Exercises on phrasal verbs. *TESOL Quarterly*, 53(2), 289–319. <https://doi.org/10.1002/tesq.478>

The processing and production of L2 tense-aspect by Chinese and Arabic learners of English

Yu Liu (University of York) and Leah Roberts (University of York).

The acquisition of L2 morphosyntax is a challenge for L2 learners both in comprehension and production, particularly when grammatical structures differ between a learner's L1 and the target language (DeKeyser, 2005). Studies on L2 acquisition of grammar have found varying performance depending on modality and also on what type of knowledge is being tapped into. For instance, learners may display metalinguistic/explicit knowledge but are less able to apply it in real-time comprehension (Roberts & Liszka, 2013). Some learners might be sensitive to grammatical distinctions during on-line comprehension, but persistent difficulties are found in their production (Trenkic et al., 2014). This study investigates the explicit knowledge (comprehension: AJT; production: gap-filling), online comprehension (visual world eye-tracking) and production (sentence imitation), focusing on tense-aspect in L2 English, of upper-intermediate L1 Chinese (24) and Arabic (24) learners, compared with 24 native English controls.

Like in the other studies, all the participants patterned together in the judgment task, demonstrating similar explicit knowledge of the present progressive, present perfect and past simple items. The results of the eye-tracking task from cluster-based permutation analyses found that both L2 groups predicted the correct construction (looked to the pictures containing ongoing events when hearing present progressive sentences, and to the pictures depicting completed events hearing present perfect sentences, e.g., The girl is planting/has planted flowers in the garden), although this prediction emerged later than the English group. Cross-linguistic influences were suggested: the L1 Chinese group failed to show prediction when hearing simple past sentences, and their divergency points of looking at the target picture appeared significantly later than the Arabic group (95% CI = [-202, -750]ms, $p = .003$). In contrast, L1 Chinese showed predication 100ms earlier than L1 Arabic group upon hearing present perfect sentences, but the difference was not significant.

Similar to the above, all participants scored highly in the gap-filling task (Chinese: 88%; Arabic: 86%; English: 92%), in contrast to the task tapping more implicit processes. The results of the online production task (elicited imitation) found that both L2 groups performed less well in comparison to the AJT and the comprehension task. The two learner groups performed similarly numerically (Chinese - corrected 34%; Arabic - 37%), and much worse than the controls (81%). The Chinese group tended to omit verb inflections and the Arabic learners mainly used simple past structures in present perfect contexts, which could be attributed to L1 influence.

The initial findings provided positive evidence for successful L2 predictive processing in tense-aspect and the potential influence of L1. It also suggested that while learners may be able to utilise morphosyntactic information during comprehension, their production knowledge often falls behind.

References: DeKeyser, R. M. (2005). What makes learning second-language grammar difficult?. *Language Learning*, 55, 1–25. Roberts, L., & Liszka, S. A. (2013). Processing tense/aspect-agreement violations online in the second language. *Second Language Research*, 29(4), 413–439. Trenkic, D., Mirkovic, J., & Altmann, G. T. M. (2014). Real-time grammar processing by native and non-native speakers. *Bilingualism*, 17(2), 237–257.

On the coupling between perception and production of second-language sounds in tasks involving lexical processing

Miquel Llompart (Universitat Pompeu Fabra).

The extent to which speech perception and production develop together throughout the acquisition of a second-language (L2) phonological system has been a topic of interest for decades. A common procedure for its investigation has been to examine the perception and production of particular L2 sounds in separate tasks and assess whether a relationship between individual performances can be found. This approach has rendered mixed findings, and it has been suggested that methodological differences across studies have likely contributed to this outcome [1]. One of the key parameters appears to be the level(s) of processing required by the experimental paradigm, with a recent study [2] showing that perception and production tasks that similarly tapped into prelexical processing only elicited a tight relationship between the two modalities, but that this was not the case when one, but not the other, additionally involved lexical processing. Building on this, the present study asks whether a relationship between perception and production is also apparent when they both unambiguously rely on lexical processing.

Thirty-four intermediate-to-advanced German learners of English completed an auditory lexical decision task, a word reading task and an imitation task focusing on the English /ɛ/-/æ/ contrast, which is known to be challenging for this population (e.g., [3]). In the lexical decision task (LDT), participants were presented with real English words and nonwords containing phonological substitutions (e.g., h[æ]ppy vs. *dr[ɛ]gon). In the word reading task, they read a set of English words aloud and, in the imitation task, they imitated the steps of an /ɛ/-/æ/ continuum. Following [2], we tested whether the accurate rejection of /ɛ/- and /æ/- nonwords in the LDT could be predicted by the production of /ɛ/ and /æ/ in the word reading task and/or in the imitation task. Results of mixed effects logistic regression modeling revealed that accuracy in perception was only predicted by the former.

These results indicate that, when properly accounting for level(s) of processing, a robust perception-production link can be observed even in tasks that do require lexical access and do not have an obvious segmental scope, but that this is not the case when there is a mismatch between modalities in terms of processing demands. The present results thus extend previous findings for prelexical perception and production to situations requiring higher-order processing, highlight the impact of crucial methodological decisions on the degree of coupling between the two modalities, and may have profound pedagogical implications for the design of L2 pronunciation training regimes.

References

- [1] Flege, J. E., Bohn, O. -S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, 25(4), 437-470.
- [2] Melnik-Leroy, G. A., Turnbull, R., & Peperkamp, S. (2022). On the relationship between perception and production of L2 sounds: Evidence from Anglophones' processing of the French/u-/y/contrast. *Second Language Research*, 38(3), 581-605.
- [3] Nagle, C. L., & Baese-Berk, M. M. (2022). Advancing the state of the art in L2 speech perception-production research: Revisiting theoretical assumptions and methodological practices. *Studies in Second Language Acquisition*, 44(2), 580-605.

Decoding the functions of gestures in adult learners of English as a foreign language: Preliminary results from a 3 year long multi-participant study

Renia Lopez-Ozieblo (The Hong Kong Polytechnic University).

Gestures are a –mostly ignored– window onto second language learners’ thinking processes (Stam, 2015). In second and foreign language acquisition, studies of gestures are crucial because learners use gestures for communicative and psychological functions related to language learning. Gestures can be an indication of cognitive load and of thinking in the foreign language (Thinking for Speaking, Slobin, 1987), yet gestures are under-represented in this field and in linguistics in general. Although a number of scholars have identified this gap (Littlemore, 2009, Gullberg, 2023), the field still lacks a comprehensive theoretical foundation linking speech and gesture in adult second language learners. This paper presents the preliminary results of a long term, multi-participant, longitudinal gesture vs. proficiency study. Data has been collected over a period of three years from over 100 Hong Kong learners of English as a foreign language. During this time, participants have performed similar video-narrating tasks up to four times (based on Tweety and Sylvester’s videos, often used in gesture studies). The gestures co-occurring during the narrations have been transcribed, together with the speech, and categorized as either representational or pragmatic (with sub-categories for meta-discursive functions, Lopez-Ozieblo,). Our results will present the frequency of gestures co-occurring with speech at each session for each participant and discuss whether changes in proficiency level have resulted in gesture frequency differences. Gesture frequency has been measured per gestures, per word as well as per clause to highlight how analytical practices can result in different conclusions. Specific examples will be presented to articulate the relationship gesture-speech in understanding the difficulties second language learners might have when expressing thoughts that have been fully conceptualized but cannot be expressed due to linguistic limitations, thus highlighting the importance of gesture evaluation during oral communicative exercises. We will argue that when learners have a clear mental model of the concept to be externalized their use of representational gestures decreases with proficiency while pragmatic meta-discursive gestures develop to present a story richer in logico-argumentative relationships.

References Lopez-Ozieblo , Gullberg, M. (2023). The integration of speech and gestures. In A. Godfroid, & H. Hopp, (Eds.). (2022). *The Routledge handbook of second language acquisition and psycholinguistics* (137-149) Taylor & Francis. Littlemore, J. (2009). *Applying cognitive linguistics to second language learning and teaching*. Springer. Slobin, D. I. (1987, September). Thinking for speaking. In *Annual Meeting of the Berkeley Linguistics Society*, 13, 435-445. Stam, G. (2015). Changes in Thinking for Speaking: a longitudinal case study. *The Modern Language Journal*, 15, 83-99.

How language and emotional intelligence shape judgements of real-life moral transgressions of first and foreign language users of English

Pernelle Lorette (Univeristät Mannheim), Irini Mavrou (Universidad Nebrija), Andreas Kyriakou (Universidad Nebrija) and Jean-Marc Dewaele (Birbeck, University of London).

It has been argued that bilinguals are more likely to evaluate moral transgressions and murder scenarios less harshly in their LX than in their L1 due to differences in the emotionality of L1 and LX (e.g. Hayakawa et al., 2017). However, the mediating role of emotion in the relationship between moral judgements and language may have been assumed but has not yet been empirically demonstrated. Moreover, previous research almost exclusively used hypothetical scenarios that were exclusively presented in the written mode and that are unlikely to occur in real life (e.g. Costa et al., 2014; Driver, 2022). The current study is the first one to employ real-life moral transgressions in an authentic audiovisual format to investigate how 209 bilingual Greek-English and 187 Hungarian-English speakers judge the severity of different types of moral transgressions (mild, medium, severe) in their L1 (Greek/Hungarian) versus LX (English), the degree to which these moral transgressions elicit emotions in these bilinguals, and the punishment they proposed to the perpetrators. Moreover, this study takes the role of individual differences into consideration by focusing on emotional intelligence (EI). To this end, the Greek-English and Hungarian-English bilinguals, as well as 256 British L1 users, were presented with six authentic TV news bulletins presenting moral transgressions of mild, medium and extreme severity. The results revealed strong correlations between offensiveness, emotionality ratings and proposed punishments for the medium and extreme severity videos in both L1 and LX. However, the Greek and Hungarian participants who watched the videos in their LX English reported lower offensiveness and emotionality ratings and opted for less harsh punishments for the perpetrators than in their L1. Moreover, the role of language in moral decision-making was mediated by emotional intensity, but only for the extreme severity stimuli. The results also suggest that higher levels of emotional intelligence are linked with tougher judgments of offence seriousness and stronger emotions while watching real moral transgressions. These findings contribute to a better understanding of the moral foreign language effect in a number of realistic moral transgressions. They highlight the implications that SLA have beyond language development itself and add to our understanding of the lived experience of bilingualism.

References

Costa, A., Foucart, A., Hayakawa, S., Aparici, M., Apesteguia, J., Heafner, J., & Keysar, B. (2014). Your morals depend on language. *PLoS ONE*, 9(4), Article e94842.

<https://doi.org/10.1371/journal.pone.0094842>

Driver, M. Y. (2022). Switching codes and shifting morals: How code-switching and emotion affect moral judgment. *International Journal of Bilingual Education and Bilingualism*, 25(3), 905–921.

<https://doi.org/10.1080/13670050.2020.1730763>

Hayakawa, S., Tannenbaum, D., Costa, A., Corey, J. D., & Keysar, B. (2017). Thinking more or feeling less? Explaining the foreign-language effect on moral judgment. *Psychological Science*, 28(10), 1387–1397.

<https://doi.org/10.1177/0956797617720944>

How working memory relates to reading comprehension between screen and print among Chinese and English university students

Handan Lu (University of Leeds).

Previous research (Singer & Alexander, 2017) has shown that university-level students' reading comprehension is more effective when they read in print than on screen, however, limited studies have examined this issue with L1- and L2-English-speaking university students. Furthermore, in these studies, reading is generally assumed to be a global construct rather than being comprised of a more complex architecture. This study is novel in that it first examines the differences that might exist in print and screen reading comprehension in L1- and L2-English-speaking university students while taking into account two dimensions of reading, namely, literal and inferential (Kintsch, Rawson, Snowling, & Hulme, 2005). Second, this study explores whether and how working memory capacity could account for the differences in literal and inferential comprehension given that it plays an important role in individual differences in reading comprehension ability (Baddeley, 2012).

Two groups of first-year undergraduate students (112 in total) from three UK universities read an English text. Each group consisted of 28 L1-English-speaking students and 28 L1-Chinese-speaking students who use English as a second language. One group read a 2500-word text on global marketing on paper and the other read the same text on a laptop screen. Participants then completed a researcher-designed comprehension test with forty-two open-ended questions (21 literal and 21 inferential). Tests of simple verbal memory (forward digit recall), simple visuospatial memory (Corsi block-tapping), and complex working memory (backward digit recall) were also administered to assess students' memory storage and processing performance. The data from the reading comprehension test were analyzed using two-way ANOVA and the data from the working memory tests were analyzed by regression statistical techniques. Results suggested that the medium had a significant effect on reading comprehension performance, with participants achieving higher scores after reading on paper than reading on screen. However, whilst higher scores were obtained in the print condition than in the screen condition for the inferential questions, no such difference was apparent for the literal comprehension questions. This pattern was found to be consistent with both L1-English and L1-Chinese speakers. Further, the capacity of working memory measured by the composite score of both complex and simple span tasks had a significant role in reading comprehension across two medium. In particular, the working memory capacity measured by complex span tasks had a meaningful contribution to inferential level comprehension across two medium; but at literal level such connection was only found with screen. This study has made important methodological contributions to the reading research comparing screen versus print and working memory research. Pedagogical and technological implications for onscreen academic reading activities at higher education were discussed.

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annual review of psychology*, 63, 1-29. Kintsch, W., Rawson, K., Snowling, M., & Hulme, C. (2005). The science of reading: A handbook. The science of reading: A handbook. Singer, L. M., & Alexander, P. A. (2017). Reading across mediums: Effects of reading digital and print texts on comprehension and calibration. *The Journal of Experimental Education*, 85(1), 155-172.

When cross-linguistic influence intersects with deixis: Can web-based contrastive instruction improve L2 spoken accuracy?

Matt Lucas (Kansai University).

One strand of instructed second-language acquisition research seeks to understand how properties of the native tongue (L1) can be harnessed for the purposes of learning a second or additional language (L2). A common means of doing so is by utilizing explicit information (EI) to raise awareness of problematic linguistic features in conjunction with task-essential practice. Awareness can be raised through the L1 (McManus & Marsden, 2018, 2019), through the L2 (Henry et al., 2017), or through contrastive instruction using the L1 and L2 (Lucas & Yiakoumetti, 2017; Spada et al., 2005). Despite these developments, however, deeper investigations into contrastive instruction remain scarce. Furthermore, related studies involving awareness-raising techniques are limited in four main respects. Firstly, they tend to focus on a single grammar point at the expense of the bigger picture. For instance, the concept of countability can adversely affect English plural-form accuracy among learner groups whose L1 lacks such a system (Shintani & Ellis, 2010), yet its influence can extend to a global level, particularly when surrounding sentences are deictically referenced (e.g., that/those; it is/they are). Moreover, noun phrases (NPs) can be omitted or repeated, resulting in further inaccuracies (Tomioka, 2003; Yoshioka, 2011). Secondly, instructional techniques that incorporate EI typically raise awareness through the input, leaving room to explore the additional role of the output (Swain, 2005). Thirdly, prior enquiries are seemingly limited in terms of their scope of outcome measures, specifically within the spoken domain (Shintani, 2015). Finally, previous associated research has typically been classroom/laboratory-based, highlighting the need to extend practice beyond such constraints through web-based instruction, particularly as language teaching transitions into a post-pandemic era. Therefore, this exploratory study (N = 81) enquired whether a synthesis of input/output-based contrastive instruction delivered through an online medium would improve spoken deictic referencing accuracy among intermediate-level Japanese university EFL learners. The design involved an experimental group being exposed to four weekly online lessons that raised awareness of L1-L2 differences concerning three target features: deictic plural forms, NP repetitions, and NP omissions. This was provided first in the input (via form-meaning connections between pictures and sentences; VanPatten, 2004), followed by task-essential practice (L1-L2 translation) and EI (L1-L2 contrasts regarding the target features), before focusing attention on the output (via L2 shadowing) with further task-essential practice (voice-recording of the target sentences). A control group received four equivalent online lessons following similar principles but with no contrastive components. Accuracy was tested at pre/post/delayed intervals using an online controlled speaking task (eliciting the target deictic features through pictures and gap-fill prompts) according to three principles of Levelt's (1989) Speech Production Model (conceptualization, formulation, articulation). ANOVA analyses revealed that for deictic plural forms, no significant differences were observed between the experimental and control groups. However, for NP omissions and NP repetitions, the experimental group displayed significant within-group improvements, suggesting partial effectiveness of the treatment. These findings reiterate the resistant nature of linguistic features influenced by the underlying concept of countability, yet highlight the rich potential that web-based contrastive instruction may provide. Pedagogical implications are discussed.

Examining Gender Agreement Processing in Spanish as a Heritage Language: An EEG Study.

Alicia Luque, Eleonora Rossi, Maki Kubota, Megan Nakamura, César Rosales, Cristina López-Rojas, Yulia Rodina and Jason Rothman

Heritage speakers (HSs) bilinguals are native speakers of their heritage language (HL) (cf. Rothman & Treffers-Daller, 2014), which they typically acquire early and naturalistically but in a context of reduced input or opportunities over the lifespan to use the language in a comparatively similar manner as homeland native speakers. As a result, research in the context of the United States (US) has shown differences between HSs and homeland native speakers across various grammatical domains (Montrul, 2016, Polinsky & Scontras, 2020). However, it is unclear whether this holds true for linguistic processing, as most of this research has relied on offline behavioral measures rather than online ones. The present study aims to fill this gap by examining grammatical gender processing in Spanish in HSs using neurolinguistic methods, which have been used often to investigate language processing with homeland native speakers and second language learners, but, to our knowledge, no neurolinguistic studies exist yet with HSs. Building on previous studies that have demonstrated a robust P600 brain response, typically associated with morphosyntactic processing (see Morgan-Short, 2014), to gender agreement anomalies in Spanish homeland native speakers and second language learners, in the present study, we aim to provide a basis of comparison to explore heritage language processing in an otherwise well-studied domain of grammar.

A total of 44 US Spanish-English HSs participated in our study and were asked to complete a behavioral gender assessment task and a grammaticality judgment task (GJT) while their brain activity was recorded using EEG. The behavioral gender assignment task was aimed to examine each participant's assigned gender value for the complete set of nouns used in the EEG/ERP experiment. The EEG/ERP GJT task comprised grammatical and ungrammatical sentences that included gender violations, where morphological transparency was manipulated. Our results indicated that, at the neural level, gender violations elicited a significant P600 effect ($p < .001^{**}$) across all relevant conditions and, crucially, that transparency matters during online gender processing, with transparent nouns eliciting a more positive response than opaque ones (see Figure 1). Our results provide converging behavioral and brain evidence suggesting that HSs' grammatical representations and processing strategies for gender are similar to those of other Spanish native speakers, contradicting previous behavioral findings for HSs' grammatical gender processing thus far. Notably, the results of our study highlight the importance of incorporating online processing methods, such as EEG, to examine both the universality and diversity in HL processing in a more fine-grained and ecologically valid manner, not only for gender but across all domains of HL grammar(s).

Which aspects of visual motivation aid the implicit learning of signs at first exposure?

Chloe Marshall (University College London), Julia Hofweber (University College London), Lizzy Aumonier (University of Kent), Vikki Janke (University of Kent) and Marianne Gullberg (Lund University).

The largely arbitrary nature of form-meaning mappings is well-established. Nevertheless, many studies now propose that motivated form-meaning mappings are more prevalent in language and more relevant to its acquisition than previously acknowledged (e.g., Nielsen & Dingemans, 2020). There is particular interest in sign languages, where the visual modality offers greater affordance for motivated mappings between form and meaning, and which are therefore rich in iconicity. Although iconicity appears to support lexical sign learning in instructed settings, it is not known whether learning is supported in implicit contexts where the learner must infer sign meaning from only the form and from conceptual knowledge of what that form might refer to.

We addressed this research gap by exploring three facets of visual motivation, namely iconicity (the recognised resemblance between a form and a given meaning), transparency (the ability to infer the meaning of a lexical item based on the form without prior knowledge of the meaning) and gesture similarity (the resemblance between a silently gestured form and a lexical sign).

Adapting an earlier paradigm (Gullberg et al., 2010, 2012) to sign, we investigated whether sign-naïve learners can infer and learn the meaning of signs after minimal exposure to continuous, naturalistic input in the form of a weather forecast in Swedish Sign Language. Participants were L1-English adults. Two experimental groups watched the forecast once (N=40) or twice (N=42); a control group did not (N=42). Participants were then asked to assign meaning to 22 target signs. We explored four predictors of meaning assignment: item occurrence frequency, iconicity, transparency, and gesture similarity.

Not unexpectedly, this was a difficult task, with mean accuracy levels around 12%. Meaning assignment was enhanced by exposure and item frequency, thereby providing evidence for implicit language learning in a new modality, even under challenging conditions. Accuracy was also contingent upon iconicity and transparency, but not upon gesture similarity. Transparency was the most influential factor. Meaning assignment at first exposure to sign is thus visually-motivated, although the overall low accuracy rates and further qualitative analyses suggest that visually-motivated meaning assignment is not always successful.

The findings of the present study indicate that learning is possible from naturalistic input on first exposure to sign (see also Hofweber et al., 2022). These discoveries complement work on spoken languages (Gullberg et al., 2010, 2012) revealing that adult language-learning mechanisms are able to operate across modalities.

References Gullberg, M., et al. (2012). What word-level knowledge can adult learners acquire after minimal exposure to a new language? *International Review of Applied Linguistics in Language Teaching*, 50, 239-276. Gullberg, M., et al. (2010). Adult language learning after minimal exposure to an unknown natural language. *Language Learning*, 60, 5-24. Hofweber, J., et al. (2022). Breaking into language in a new modality: The role of input and individual differences in recognizing signs. *Frontiers in Psychology*. 13:895880. Nielsen, A., & Dingemans, M. (2020). Iconicity in word learning and beyond: A critical review. *Language and Speech*, 64, 52-72.

L1 attrition in instructed and immersed bilinguals: evidence from the production, interpretation, and processing of subject referring expressions

Fernando Martín-Villena (Universidad de Granada), Cristóbal Lozano (Universidad de Granada) and Antonella Sorace (University of Edinburgh).

The interaction of two languages in the bilingual brain generally leads to bidirectional influence and thus, differential L1 and L2 outcomes. While most research to date has investigated bilinguals' L2 outcomes, research on the bilinguals' L1 within L1 attrition has received considerably less attention, although it has certainly been gaining ground (Schmid & Köpke, 2019). L1 attrition has primarily investigated long-immersed bilinguals who have spent a considerable time in an L2 environment and bilinguals in an L1-dominant setting have largely been overlooked. To address previous gaps, our study will explore the production, interpretation, and processing of subject referring expressions (REs) in L1 Spanish-L2 English instructed and immersed bilinguals, a domain that has been found to be vulnerable in L1 attrition (Chamorro & Sorace, 2019; Gürel, 2019). 81 L1 Spanish-L2 English advanced instructed bilinguals in Spain, 93 L1 Spanish-L2 English immersed bilinguals in the UK, and 33 L1 Spanish functional monolinguals completed two oral video-retelling tasks, a picture selection task (PST), and a self-paced reading task (SPR) which tested the predictions from the Position of Antecedent Strategy (Carminati, 2002), i.e., whether null pronouns bias towards subject antecedents and overt pronouns prefer object antecedents. Additional information was collected from the Bilingual Language Profile (Birdsong et al., 2012), a background questionnaire that provided a continuous language dominance score. Firstly, the production tasks were used to analyse differences in the distribution of 3rd person subject REs in topic continuity (TC) in the three groups using a fine-grained tagset in the UAMCorpusTool (O'Donnell, 2009). The PST (adapted from Tsimpli et al., 2004) contained 20 experimental items where a main clause, with two same-gender antecedents in subject and object position, was followed by a subordinate clause containing either a null or an overt pronoun. Participants selected the picture (subject vs. object bias) that best matched their interpretation. The SPR task presented the same sentences and biased towards a subject or an object interpretation using one of the two pictures presented in the PST task. The data from the PST and SPR were analysed using (generalised) linear mixed-effect models in R (Bates et al., 2015), and the final models of best fit included fixed effects as well as the maximal converging random-effect structure allowed by the design (Barr et al., 2013). The results from the production tasks show that instructed and immersed bilinguals significantly produce more unexpected overt REs in TC than functional monolinguals overall, with additional significant differences between the two bilingual groups. The results from the PST and the SPR tasks indicate that 1) the two bilingual groups differed from functional monolinguals only in the overt pronoun condition, as predicted by the Interface Hypothesis (Sorace, 2011, 2012) and the Activation Threshold Hypothesis (Paradis, 2007) and 2) variability in interpretation preferences of overt pronouns was modulated by language dominance. These results provide further evidence on the nature of L1 morphosyntactic attrition at the early stages as well as new results on the variability found in L1 instructed bilinguals, which has been overlooked in previous research.

Testing the three-stage model of second language skill acquisition

Ryo Maie (University of Tokyo) and Aline Godfroid (Michigan State University).

Skill acquisition theory conceptualizes L2 learning as a type of skill acquisition, comparable to learning how to drive a car or solve arithmetic problems. The dominant view is that L2 skill acquisition is a three-stage process, most notably described as the declarative, transitional, and procedural stage, each of which involves distinct cognitive states for performing the targeted skill. While current L2 research presents indirect evidence for the existence of such distinct stages in language learning, the number and nature of those stages are often assumed a priori and whether or not they actually exist remains untested in SLA. Our primary research aim, therefore, was to test and validate the three-stage model of L2 skill acquisition.

Sixty-five adult learners deliberately learned and practiced how to comprehend and produce a miniature language based on Japanese for 1,056 trials. Comprehension practice took the form of a sentence-picture matching task, and production practice was implemented as a productive maze task. We observed how participants developed their accuracy and speed of performance by observing how accuracy, reaction time (RT), and the coefficient of variability (CV) of RT changed over time with practice. Participants also took a battery of tests on three dimensions of cognitive abilities known to play a primary role in each stage of skill acquisition: declarative memory, procedural memory, and psychomotor ability. The cognitive tests included (i) the Continuous Visual Memory Task and LLAMA-B for declarative memory, (ii) an alternating serial reaction time task and a statistical learning task for procedural memory, and (iii) the alternating serial reaction time task and a two-choice RT task for psychomotor ability.

We analyzed the data from the language practice and the cognitive tests in two steps. First, we fitted a series of hidden Markov models (HMMs) to the RT data with one, two, and three stages, respectively. These mathematical models represented different hypotheses regarding the number of skill acquisition stages. Using RT as input, the HMM estimated the probability of each participant residing in a given stage on a given practice trial. This first step of the analysis revealed that the three-stage HMM was most consistent with the comprehension data, whereas the two-stage HMM best fit the production data.

Taking the learning stages identified by hidden Markov modeling, we then utilized a series of generalized linear mixed models to investigate the nature of these stages. Specifically, we examined whether the individual difference measures of declarative, procedural, and psychomotor ability predicted accuracy, RT, and the CV during practice in each stage. The results showed that (a) declarative memory reliably predicted accuracy, RT, and the CV throughout practice, but its influence diminished as the participants reached more advanced learning stages, and (b) procedural memory predicted RT and the CV in the first and the second stage of practice. Overall, the findings of the study provide the first direct evidence for the three-stage model of L2 skill acquisition in SLA and reveal individual differences a classroom teacher can accommodate to support the development of advanced proficiency in L2 skills.

Enjoyment, anxiety, and language learning motivation in CLIL and non-CLIL: a longitudinal approach

Laurence Mettwie (UNamur), Luk Van Mensel (XTJLU - UNamur) and Benoît Galand (UCLouvain).

The aim of this study is to investigate the impact of a Content and Language Integrated Learning approach (CLIL) on socio-affective outcomes from a longitudinal perspective. 756 French-speaking pupils at primary or secondary school, learning either English or a LOTE (Dutch), in a CLIL track or in traditional foreign language classes, participated in the study. The participants twice completed a questionnaire and several tests over an 18-month interval. The questionnaire included items measuring their emotions in the classroom (anxiety and enjoyment) and motivation for language learning (perceived task-value, expectancy for success, and perceived cost). A range of individual background characteristics, including initial vocabulary knowledge in the target language, were included in the analyses. The findings of the group comparison between CLIL and non-CLIL (between-subject) showed that the CLIL group reported more favorable emotions and motivation for language learning, in line with previous cross-sectional research on socio-affective outcomes. However, the longitudinal results (within-subject) indicated that (a) the effects of CLIL were limited, particularly when initial vocabulary knowledge was factored in, and (b) that the few effects of CLIL vary depending on the tested emotional or motivational variables and on the target language. The findings provide insights for researchers and practitioners on the importance of and interplay between different variables in the CLIL approach as an instructional method.

Vocabulary learning in a novel language after minimal exposure to multimodal input

Imma Miralpeix (Universitat de Barcelona), Ferran Gesa (Universitat de Barcelona) and M^a del Mar Suárez (Universitat de Barcelona).

First Exposure (FE) studies have shown that lexical knowledge can be incidentally acquired at the very first contact with a completely unknown language (Gullberg et al., 2012; Shoemaker & Rast, 2013). However, very little is known about how multimodal input, which has been consistently found to favour language acquisition (Mayer, 2009; Vanderplank 2016), might contribute to intentional vocabulary learning at FE. Similarly, it remains to be explored whether learners' previous exposure to this kind of input in other languages may affect learning in the new one, or which strategies learners may use in this endeavour. To fill the existing gap, a total of 106 Catalan/Spanish adults proficient in English watched a three-minute advert with the audio in English and subtitles in Polish, a language they were not familiar with. Participants were instructed to watch it attentively and try to learn as much as possible of the new language. After a second viewing, they were asked to take a timed meaning recognition test with 15 Polish words and expressions appearing in the advert. They were also administered a background questionnaire enquiring, among other aspects, about their extramural exposure to multimodal input and the strategies they had used to learn from the video. A no-treatment control group of participants with the same characteristics (n=29) was included for comparison purposes: they took the vocabulary test, but they did not watch the advert nor did they answer the questionnaire. Results showed that participants watching the subtitled advert were able to make form-meaning connections, recognizing the meaning of about half of the target Polish vocabulary after only six minutes of video exposure [M=7.33, SD=2.43, 95% CI (6.86, 7.80)]. These scores were also significantly higher than those in the control group (U=234.5, z=-7.015, p<.001). The background questionnaire indicated that frequency of exposure to original version materials was not related to Polish scores, but that previous experience with captioned input was [$\chi^2(10, 62)=21.039$, p=.021 / Phi and Cramer's V: .583]. Furthermore, associations with imagery and with vocabulary in the L1 were among the most recurrent strategies students used to help them learn, along with input factors such as frequency of word occurrence. Findings reveal how multimodal input can help adult vocabulary acquisition at FE, as well as those factors relevant for learning from video viewing at these very first stages. -

References Gullberg, M., Roberts, L., & Dimroth, C. (2012). What word-level knowledge can adult learners acquire after minimal exposure to a new language? *IRAL - International Review of Applied Linguistics in Language Teaching*, 50(4), 239–276. <https://doi.org/10.1515/iral-2012-0010> Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press. Shoemaker, E., & Rast, R. (2013). Extracting words from the speech stream at first exposure. *Second Language Research*, 29(2), 165–183. <https://doi.org/10.1177/0267658313479360> Vanderplank, R. (2016). Captioned media in foreign language learning and teaching. Subtitles for the deaf and hard-of-hearing as tools for language learning. Palgrave Macmillan.

Investigating the cognitive and linguistic underpinnings of L2 Spelling

Heike Mlakar (University of Hildesheim), Joanna Hirst-Plein (University of Hildesheim) and Martin Koch (University of Wuerzburg).

As a central, lower-level component of the writing process, spelling is a complex skill which requires precise recall and production of linguistic and orthographic representation. Continued cognitive overload whilst spelling can have detrimental consequences on the subsequent development of writing skills. While previous studies into young learners' literacy development have focused predominantly on reading skills, empirical findings on the effects of language-related and cognitive predictors of L2 spelling are scarce and findings are mixed. The present study investigates how cognitive and linguistic factors as well as young learners' reading skills affect L2 spelling acquisition. Participants (n=75) were grade 3 and grade 4 learners at two primary schools in Lower Saxony, Germany. The study employed a longitudinal design with six measures of cognitive and linguistic skills (working memory, phonological short-term memory, non-verbal intelligence, phonological awareness, L2 receptive grammar and vocabulary), three measures of reading skills (L2 reading comprehension, L2 oral reading rate and accuracy) and a real word and pseudoword spelling test to explore L2 spelling proficiency. Correlations, multiple regression analyses and a structural equation model were calculated in order to identify major component skills and knowledge bases relevant for young L2 learners' spelling acquisition. Our data revealed that L2 reading comprehension and L2 oral reading accuracy predicted learners' real word spelling performances whereas non-verbal intelligence predicted pseudoword spelling. Learners' cognitive and linguistic abilities did not predict spelling directly, whereas phonological awareness and L2 grammar had an indirect impact on spelling mediated by reading skills. We discuss the results in relation to the componential model of writing (Harrison et al., 2016; Schoonen et al., 2002, 2011) and highlight educational implications.

References

Harrison, G., Goegan, L.D., Jalbert, R., McManus, K.L., Sinclair, K., & Spurling, J. (2016). Predictors of spelling and writing skills in first- and second-language learners. *Reading and Writing, 29*, 69-89. <https://doi.org/10.1007/s11145-015-9580-1>

Schoonen, R., van Gelderen, A., de Glopper, K., Hulstijn, J., Snellings, P., Simis, A., & Stevenson, M. (2002). Linguistic knowledge, metacognitive knowledge and retrieval speed in L1, L2, and EFL writing. In S. Ransdell & M. Barbier (Eds.), *New directions for research in L2 writing* (pp. 101–122). Kluwer Academic Publishers.

Schoonen, R., van Gelderen, A., Stoel, R. D., Hulstijn, J., & de Glopper, K. (2011). Modeling the development of L1 and EFL writing proficiency of secondary school students. *Language Learning, 61*(1), 31–79. <https://doi.org/10.1111/j.1467-9922.2010.00590.x>

Developing productive knowledge of single words and formulaic sequences from explicit instruction

Rebecca Moden (University College London, Institute of Education).

Single words (SW) and formulaic sequences (FS) need to be acquired both receptively and productively for successful communication. However, it is the progression from receptive to productive knowledge which is a particular challenge for second language (L2) learners. Few studies have investigated the impact of explicit instruction on the incremental development of knowledge of SW and FS from a receptive level to different degrees of productive mastery. One approach to developing lexical knowledge is through exposure to vocabulary in rich contexts. Following a first language study (Chilton & Ehri, 2015), the use of cohesive contexts depicting a common theme were expected to strongly activate learners' schemata, resulting in particularly successful learning gains. In addition, while researchers have argued for the added difficulty of learning FS when compared to SW, empirical evidence to support this claim is scarce.

The present study therefore attempted to address these gaps by examining the productive learning of SW and FS acquired in two types of linguistic contexts: cohesive contexts sharing a unified topic and semantic elements, and unconnected sentences describing unrelated events. A comparison of the productive learning gains of SW and FS was also made. The study aimed to examine whether the conditions provided in the pedagogical interventions enhanced knowledge of SW and FS, initially known only receptively, to different degrees of productivity: 1) partial form recall (i.e. the correct form supplied with minor errors), 2) full form recall (i.e. accurate form supplied), and 3) creative use in an original sentence.

Sixty-five advanced postgraduate L2 students learned 34 SW and 30 FS in one of three conditions (cohesive, unconnected, control) in seven weekly one-hour sessions. Both experimental groups encountered the items with their meanings, then the cohesive group (n=20) practised the vocabulary in paragraphs with a common theme, whereas the unconnected group (n=20) practised the lexical items in unrelated sentences. A control group (n=25) only completed the pre-tests and post-tests. The tests measured knowledge of the items at three levels: form recognition (multiple-choice), form recall (gap-filling) and creative use (sentence-writing).

Results of a two-way ANOVA showed that there was a significant main effect of group on the vocabulary items known receptively and learned to a partial and full level of productive knowledge. For both partial and full knowledge, the cohesive group significantly outperformed the other two groups. For full knowledge, the unconnected group also had significantly higher post-test scores than the control group. There was a significant main effect of item type for partial and full knowledge, which showed an overall advantage of FS over SW. Preliminary results also suggest an advantage for the cohesive group on the creative use test. Pedagogical implications will be discussed.

Chilton, W.C., & Ehri, L.C. (2015). Vocabulary learning: Sentence contexts linked by events in scenarios facilitate third graders' memory for verb meanings. *Reading Research Quarterly*, 50(4), 439-458.

Speech rhythm and speaking fluency in pronunciation-unfocused tasks: the role of task complexity

Joan C. Mora (Universitat de Barcelona), Natàlia Fullana (Universitat de Barcelona) and Ingrid Mora-Plaza (University of Barcelona).

Previous research has shown that increasing attention to form through task complexity may enhance L2 learners' lexical and grammatical accuracy (Robinson, 2011), but may negatively affect speaking fluency. The benefits of increased task complexity have also been attested for L2 vowel production accuracy (Solon et al., 2017) and perceived comprehensibility (Gordon, 2021) in pronunciation-focused tasks, but not for speech rhythm and other L2 suprasegmental features in pronunciation-unfocused tasks. Increased cognitive task demands results in lower judgements of L2 pronunciation accuracy and speaking fluency (Crowther et al., 2018), though the relationship between speech rhythm and speaking fluency and how they relate to perceived accentedness and comprehensibility in tasks differing in cognitive complexity is currently under-researched. This study investigated the relationship between speech rhythm and temporal measures of speed, breakdown and repair fluency, and the extent to which rhythm and fluency predicted native listeners' perception of accentedness and comprehensibility, in the speech of eighty-two Catalan/Spanish advanced EFL learners who performed two versions of the Dinner Party task (Ur, 1981) differing along +/- reasoning demands. L2 speech rhythm was gauged through vocalic interval duration metrics previously shown to maximally distinguish Spanish syllable-timed rhythm from English stress-timed rhythm (%V, ΔV , varcoV, nPVI-V; Prieto et al., 2012). Results indicated that learners were more fluent (spoke faster, produced shorter pauses and fewer clause-internal pauses) and exhibited a more target-like (stress-timed) speech rhythm on most metrics (ΔV , varcoV, nPVI-V) in the simple than the complex task, suggesting that increased task complexity was detrimental to both speech rhythm and speaking fluency. Associations between speech rhythm and speaking fluency were weak and measure-dependent, with more target-like speech rhythm occurring in slower speech containing longer pauses. Learners whose speech was more fluent and less syllable-timed (lower %V) were perceived to be less accented and more comprehensible in both simple and complex tasks. Given that associations between rhythm and fluency, accentedness and comprehensibility were found to be highly dependent on the metric used in L2 learners' spontaneous speech, Mahalanobis distances between learners' L2 English speech and native English speech on a two-dimensional space defined by pairs of rhythm metrics (%V and varcoC) are currently being explored to validate the findings concerning the varying relationships between speech rhythm and speaking fluency.

Crowther, D., Trofimovich, P., Saito, K., & Isaacs, T. (2018). Linguistic dimensions of L2 accentedness and comprehensibility vary across speaking tasks. *Studies in Second Language Acquisition*, 40(2), 443-457. Gordon, J. (2021). Pronunciation and task-based instruction: Effects of a classroom intervention. *RELC Journal*, 52(1), 94-109. Prieto, P., del Mar Vanrell, M., Astruc, L., Payne, E., & Post, B. (2012). Phonotactic and phrasal properties of speech rhythm. Evidence from Catalan, English, and Spanish. *Speech Communication*, 54(6), 681-702. Robinson, P. (2011) *Second language task complexity: Researching the Cognition Hypothesis of language learning and performance*. John Benjamins. Solon, M., Long, A. Y., & Gurzynski-Weiss, L. (2017). Task complexity, language-related episodes, and production of L2 Spanish vowels. *Studies in Second Language Acquisition*, 39(2), 347-380. Ur, P. (1981). *Discussions that work: Task-centered fluency practice*. Cambridge University Press.

What do you mean by “Do you like learning English?”

Mitsuhiro Morita (Hiroshima City University), Yoichi Watari (Chukyo University) and Atsushi Mizumoto (Kansai University).

In recent years, there has been a growing body of research on positive psychology in SLA. Among the studies, the enjoyment of learning a second/foreign language has been the focus of investigation, with a particular focus on the effects of the classroom environment (Dewaele & MacIntyre, 2014) and individual learner and teacher variables (Dewaele & Li, 2021; Li, 2022). Studies have found that language learning enjoyment occurs when learners experience success, positive change, novel experiences, as well as teacher recognition and support, a sense of engagement in the classroom, or positive relationships with classmates. Although existing research provides rich insights on enjoyment, it is still not clear how positive emotions change over time, especially during the secondary school years, and the extent to which various language learning factors may influence the experience of enjoyment.

This study focuses on one of the positive emotions, liking, by asking one of the most frequently asked questions of English language learners, “Do you like learning English?” and exploring how responses to this question changed over a two-year period and what factors correlated with the responses. In this study, we collected longitudinal data from two secondary schools. 433 students took a private English language test each year and answered a questionnaire from their 1st to 3rd year at the secondary schools. The questionnaire included “Do you like learning English?”, questions about the English classes in the schools, and other questions related to English learning. The results indicated a positive correlation between the degree of enjoyment in learning English and test scores. Conversely, a decrease in enjoyment was found to correspond to a decrease in test scores. The results also showed that their responses about their feelings about learning English were stable (i.e., most of them liked learning English, disliked learning English, or were in between for three years). More importantly, it was found that the degree of understanding of the English lessons was positively correlated with the emotions toward learning English for two years. However, the relationship between emotions and other factors varied from year to year. Taken together, the first year of English classes may have a significant impact on learners’ feelings about learning English. We will discuss the findings and implications in light of the existing literature.

Dewaele, J.-M., & MacIntyre, P. D. (2014). The two faces of Janus? Anxiety and enjoyment in the foreign language classroom. *Studies in Second Language Learning and Teaching*, 4(2), 237–274.

<https://doi.org/10.14746/ssllt.2014.4.2.5> Dewaele, J.-M., & Li, C. (2021). Teacher enthusiasm and students’ social-behavioral learning engagement: The mediating role of student enjoyment and boredom in Chinese EFL classes. *Language Teaching Research*, 25(6), 922–945. <https://doi.org/10.1177/13621688211014538> Li, C. (2022). Foreign language learning boredom and enjoyment: The effects of learner variables and teacher variables. *Language Teaching Research*, 1–26, <https://doi.org/10.1177/13621688221090324>

Technology-enhanced language learning and learners' language skills: Does the traditional way of teaching and learning truly hinder learners' L2 improvement?

Evita Moulara (University of Birmingham), Petar Milin (University of Birmingham) and Dagmar Divjak (University of Birmingham).

Integrating technology into language classes is now a prevalent tendency in an attempt to facilitate L2 learning. Previously, research on Technology-Enhanced Language Learning (TELL) documented higher achievement scores compared to the traditional way of teaching (Bester & Brand, 2013) by using online learning environments that offer interactional feedback, commercial computer games, mobile applications, and social networking sites to improve L2 learners' knowledge of different linguistic components. However, most studies used university-level L2 learners – i.e., young adults, leaving the effects of TELL at earlier or later ages out of focus and, thus, understudied (e.g., only a fifth of studies has been conducted at secondary educational levels, cf., Chang & Hung, 2019). The present study aimed to fill in that gap and determine whether technology improves the acquisition of a range of lexical or grammatical phenomena for younger teenagers. We consider in particular knowledge of grammar (word order and tenses) and lexicon (polysemous words), alongside listening comprehension and oral production.

This 6-month classroom intervention study was completed at a Lower Secondary School in Greece, where the participants (L1 speakers of Greek) were students aged 12-13, and who were learning English as a Foreign Language. A control group (34 students) was taught in the traditional way using textbooks and paper-based materials while the experimental group (37 students) was exposed to digital resources (e.g., online games and videos, websites, Virtual and Augmented Reality, mobile phones). Both groups had the same teaching goals, lesson structure, and pedagogical approaches, and they were taught by the same teacher. Tests assessing the participants' knowledge of these language phenomena were also the same, and were conducted before and after the intervention to assess any improvement and difference between the two groups. The only (pseudo-) manipulated factor was method of delivery: traditional vs. digital.

The results showed that digital delivery significantly improves vocabulary ($p < 0.0002$) only, whereas other components improve equally for both methods of delivery. In other words, TELL positively contributes to teenagers' lexical knowledge (polysemy), but not knowledge of grammar (word order and tenses), nor their listening comprehension or oral production. This shows that, although recent research focuses on the effectiveness of TELL, and there is pressure to integrate technology in the classroom, an engaging technology-free L2 classroom is still effective and both modes of instruction should be considered for use in today's language classes.

FL learning outcomes after 6 years in primary school: Factors accounting for success

Carmen Muñoz (Universitat de Barcelona).

By the end of primary school, it is not uncommon for children to have already had six years of EFL instruction and to show diverse degrees of success. Such variability may result from the dynamic interaction of affective, cognitive and contextual factors. Whereas affective variables have been a special focus of attention (Mihaljević Djigunović & Nikolov, 2019), very few studies have focused on young learners' language learning aptitude (Roehr-Brackin & Tellier, 2019). Among contextual factors, the family has been identified as an important source of differences. In particular, parents' educational level (Butler & Le, 2018) and experience with English (Lindgren & Muñoz, 2014) have been associated to children's English achievement. Increased amount of contact through extra private lessons and out-of-school use of English have also been reported to influence young learners' performance (Cadierno et al., 2022). In this study we explored the influence of learner (affective and cognitive) and contextual factors on the English language outcomes of 116 learners from five schools in the area of Barcelona. Our research questions are: 1) In what ways are the different learner and contextual variables associated with each other?; 2) What variables contribute the most to young learners' achievement at the end of primary school? Correlational analyses addressing the first RQ found confirmatory evidence of strong associations between different factors, such as language aptitude scores with the scores in the three languages (Catalan, Spanish, and English), and parental education, particularly mother's education, with out-of-school exposure, extracurricular lessons, and children's language aptitude. To answer the second RQ we computed inferential analyses (GLM and LMM) for each outcome variable (General English test, Listening comprehension, Reading comprehension, Written fluency, Written lexical diversity, Written syntactic complexity). Results highlighted the role of language aptitude and parental education. In our presentation we discuss how the learner internal and contextual factors do not operate in isolation from one another shaping the conditions under which L2 learning occurs in these early years.

Butler, Y. G. & Le, V-N. (2018). A longitudinal investigation of parental social-economic status (SES) and young students' learning of English as a foreign language, *System*, 73, 4–15. Lindgren, E. & Muñoz, C. (2013). The influence of exposure, parents, and linguistic distance on young European learners' foreign language comprehension. *International Journal of Multilingualism*, 10, 105–129. Mihaljević Djigunović, J., & Nikolov, M. (2019). Motivation of young learners of foreign languages. In M. Lamb, K. Csizér, A. Henry, & S. Ryan (Eds.), *The Palgrave handbook of motivation for language learning* (pp. 515–533). Palgrave Macmillan. Roehr-Brackin, K. & Tellier, A. (2019). The role of language-analytic ability in children's instructed second language learning. *Studies in Second Language Acquisition*, 41(5), 1111–1131. Cadierno, T., Hansen, M., Lauridsen, J., Eskildsen, S., Fenyvesi, K., & Hannibal Jensen, S. (2022). A multi-factor study of the development of English receptive skills by young Danish children. *Bilingualism: Language and Cognition*, 25(4), 566–578.

Depth of processing, learner aptitude and the acquisition of L2 English grammatical structures

Shweta Nigam (University of Reading).

Depth of processing (DoP), or the cognitive effort employed by the learners in processing grammatical/lexical items in the input (Leow, 2015), has emerged as a key concept in the recent models of instructed second language acquisition. It is argued that high DoP involves greater cognitive effort, which increases the chances of long-term retention of information, potentially facilitating learning. While much research has investigated and provided empirical support for the effectiveness of DoP in relation to the acquisition of lexical items (e.g. Laufer and Hulstijn, 2001), comparatively few studies have explored the significance of DoP in the acquisition of grammatical items (e.g. Adrada-Rafael, 2017; Hsieh et al., 2016). Moreover, little is known about the role of aptitude in mediating learners' DoP of an input. Drawing on Craik and Lockhart's (1972) levels of processing framework, this study aims to fill this gap by investigating the effects of DoP on the acquisition of third conditional and comparative structures in an oral retelling task. The study also examines the mediating role of aptitude on the learners' acquisition of these forms.

The study follows a pretest-posttest-delayed posttest experimental design over four week period involving 120 Grade 8 ESL students in India. They have been randomly assigned to one of the three experimental conditions differing in the level and type of processing: low, high explicit and high implicit. A grammaticality judgement test and an elicited imitation task are used to assess the participants' gains in the knowledge of the target structures. An aptitude test (LLAMA) measures any mediating effects of aptitude on the type of processing. The treatment task requires the participants to listen to an audio-visual story containing the target structures and complete exercises (varying in level of analysis) before listening to the story for a second time and retelling it. The oral performance of the participants is analyzed for fluency, accuracy, and frequency of use of the third conditionals and comparatives.

The findings, which are currently under investigation, will provide insight into the role of DoP in the development of L2 grammatical knowledge and highlight the extent of DoP in explicit and implicit learning conditions. The results will also indicate whether the extent of learning during the treatment is mediated by the learner's aptitude (measured in LLAMA B, D and F).

References Adrada-Rafael, S. (2017). Processing the Spanish imperfect subjunctive: Depth of processing under different instructional conditions. *Applied Psycholinguistics*, 38(2), 477–508.

Craik, F.I.M., & Lockhart, R.S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671-84.

Hsieh, H.-C., Moreno, N. & Leow, R.P. (2016). Awareness, type of medium, and L2 development: Revisiting Hsieh 2008. In R.P. Leow, L. Cerezo, & M. Baralt (Eds.), *A Psycholinguistic Approach to Technology and Language Learning* (pp. 131-150), De Gruyter, Inc.

Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22(1), 1-26.

Leow, R.P. (2015). *Explicit learning in the L2 classroom: A student-centered approach*. New York, NY: Routledge.

Are textbooks a good source for vocabulary learning?

Cadit Nissan Zilbiger (Haifa University) and Batia Laufer (Haifa University).

Three important factors that affect L2 vocabulary learning are the number of encounters, spacing, and quality of engagement with words (Duong et al., 2021; Laufer & Rozovski-Roitblat, 2015; Mohamed, 2018; Nakata, 2015; Uchihara et al., 2019). Since teaching materials, particularly textbooks, are expected to provide the necessary conditions for language learning, we would expect them to fulfill the above three requirements for vocabulary learning. Most textbook evaluations explored the frequency of words, number of occurrences, and aspects of word knowledge in exercises (Brown, 2011; Bergström et al., 2022; Sun & Dang, 2020). Our study assessed a middle school textbook (170 pages) and a workbook (185 pages) as sources for vocabulary exposure and practice. We examined word recycling patterns (number of occurrences and spacing patterns) and different opportunities for engagement with words. 369 words from the national lexical curriculum were annotated for the number of times they appeared in the texts and exercises, and the pages they appeared on, to detect patterns of spacing, and the characteristics of exercises they were practiced in. Each word occurrence was assigned an 'engagement profile' according to a combination of four parameters: Whether the word was necessary for task completion, or not (e.g., word meaning recognition vs. global text understanding), whether it was focused on, or not (e.g., fill-in blanks vs. word occurrence in a text), whether comprehension was required, or not (e.g., necessary vs. unnecessary for text comprehension), whether production was required (e.g., sentence writing vs. meaning recognition). Results showed that the target words appeared between 0 and 97 times in the books, that spacing did not exhibit a definite pattern, and that most tasks practiced receptive knowledge only. We relate the results to opportunities for incidental learning and word-focused instruction.

References Bergström, D., Norberg, C., & Nordlund, M. (2022). Do textbooks support incidental vocabulary learning? – a corpus-based study of Swedish intermediate EFL materials, *Education Inquiry*, DOI: 10.1080/20004508.2022.2163050 Brown, D. (2011). What aspects of vocabulary knowledge do textbooks give attention to? *Language Teaching Research*, 15(1), 83-97. Duong, T., Maribel, Montero-Perez, M., Desmet, P., and Peters, E. (2021) Learning vocabulary in spoken input- and output-based tasks. *TASK Journal on Task-Based Language Teaching and Learning*, 1(1), pp. 100-126. Laufer, B. & Rozovski-Roitblat, B. (2015). Retention of new words: Quantity of encounters, quality of task, and degree of knowledge. *Language Teaching Research*, 19 (6), 687 – 711. Mohamed, A. (2018). Exposure frequency in L2 reading: An eye-movement perspective on incidental vocabulary learning. *Studies in Second Language Acquisition*, 40, 269–293. Nakata, T. (2015) Effects of expanding and equal spacing on second language vocabulary learning: does gradually increasing spacing increase vocabulary learning? *Studies in Second Language Acquisition*, 37, 677-711. Sun, Y., & Dang, T. N. Y. (2020). Vocabulary in high-school EFL textbooks: Texts and learner knowledge. *System*, 93, 102279. Uchihara, T., Webb, S. and Yanagisawa, A. (2019). The Effects of Repetition on Incidental Vocabulary Learning: A Meta-Analysis of Correlational Studies. *Language Learning*. 69(3), pp. 559-599

Researching learner agency from a Complex Dynamic Systems Theory perspective: A multilayered approach

Ryo Nitta (Rikkyo University) and Kyoko Baba (Toyo University).

This study explored how university students develop agency in their second language (L2) learning from a Complex Dynamic Systems Theory (CDST) perspective. Research in second language acquisition (SLA) has advanced by using a computational metaphor to explain certain aspects of L2 learning in an

theoretical sense. However, this approach disregards the fundamental aspect of learning, learner agency. Learner agency is seen as a dynamic and complex construct, described as being "relational, emergent, spatially, and temporally situated" (Larsen-Freeman, 2019, p. 73). To understand how learner agency emerges and evolves over time, we used a multilayered approach that incorporated three psychosocial layers: 1) dispositional traits, 2) characteristic adaptations, and 3) narrative identity (Dörnyei & Ryan, 2015).

We studied two individual cases in an English class of 26 students at a Japanese university over the course of a year. Data was gathered using the Motivational Factors Questionnaire, self-reflection accounts written in their first language, and L2 writing, each focusing on the three layers. Results from the first layer, the questionnaire, showed similarities between the two individuals, with relatively high scores for motivational factors related to pragmatic reasons (e.g., instrumentality) and future self-images in using L2 (e.g., the ideal L2 self). The analysis of the second layer revealed contextualized patterns of engagement in the classroom learning. Specifically, post-writing reflections suggested that both students consistently used self-regulatory processes like goal-setting, self-observation, and self-evaluation to maintain motivation and a sense of agency towards the writing task throughout the year. In the analysis of the students' written narratives in the third layer, common themes emerged (e.g., long-term career goals, immediate learning experiences), with specific episodes highlighting the individuality of their experiences outside the classroom. The synthesis of the three layers for both individual cases provides a comprehensive and multifaceted view of the learners as a whole person.

These findings also provide important insights into the methodology for researching the dynamics and complexity of L2 learners. There have been two opposing research paradigms in learner characteristics: those that identify factors that explain different learning outcomes (e.g., Roshandel et al., 2018) and those that emphasize the uniqueness of individuals in specific contexts (e.g., Ushioda, 2019). Both have developed independently, but the present findings suggest that integration, rather than separation, of these orientations would lead to a deeper understanding of the multidimensional and holistic nature of L2 learners.

References Dörnyei, Z. & Ryan, S. (2015). *The Psychology of the language learner revisited*. London: Routledge. Larsen-Freeman, D. (2019). On learner agency: A Complex Dynamic Systems Theory perspective. *Modern Language Journal*, 103(s1), 61-79. Roshandel, J., Ghonsooly, B., & Ghanizadeh, A. (2018). L2 Motivational Self-System and Self-Efficacy: A Quantitative Survey-Based Study. *International Journal of Instruction*, 11(1), 329-344. Ushioda, E. (2019). Researching L2 motivation: Re-evaluating the role of qualitative inquiry, or the 'wine and conversation' approach. *Contemporary language motivation theory*, 60, 194-211.

Age and L1 effects in the acquisition of verb morphology in English by Chinese and Russian children in an EFL context: evidence from oral narrative production

Athina Ntalli (University of Oxford), Theodora Alexopoulou (University of Cambridge) and Henriette Hendriks (University of Cambridge).

This study investigates the acquisition of tense and agreement morphology in EFL children's oral narratives aiming to address the impact of age of onset and effects due to their L1. We ask two questions: 1) whether L2 children older than 4 will pattern with L2 adults in the domain of inflectional morphology following Meisel (2009) and whether early learners will outperform later learners; 2) whether child learners from a language background that realises tense and agreement features will perform better than learners whose L1 does not. Participants were 147 9- and 12-year-old Chinese and Russian children with 5 years of instruction in English; the 9-year-olds had started learning English at 4 and the 12-year-olds at 7 and were recruited in EF English First private (afternoon) schools in either Shanghai or Moscow. Children were assessed with the Multilingual Assessment Instrument for Narratives (MAIN, Gagarina et al., 2012). We further used the Renfrew Vocabulary task (Renfrew, 1998) as a proxy for proficiency as well as narrative length, lexical diversity, and syntactic complexity in the narratives as additional measures along with the class CEFR level each child attended. Results on copula be, auxiliary be, 3SG-agreement and past tense (regular and irregular) showed low accuracy, high numbers of omissions, asymmetries in the acquisition of morphemes (copula be > auxiliary be > main thematic verbs), and use of the periphrastic structure 'is + verb(x)' (as in 'The boy is take the balloon.') in all groups. Mixed effects logistic regression analyses showed that 12-year-olds consistently outperformed 9-year-olds being more accurate in verb morphology. Chinese were significantly more accurate on copula be than Russian learners, but no significant difference was found in accuracy on auxiliary be and main thematic verbs. Narrative length was consistently a significant predictor of accuracy for copula be, auxiliary be, and main thematic verbs. In light of these results, we first discuss a proficiency discrepancy attested between Chinese and Russian 9-year-olds in our study which might explain Chinese's better performance on copula be. We argue that fieldwork across two countries needs to take a broad view of assessment of proficiency incorporating extralinguistic factors that affect children's rate of learning. Turning to our first question, children seem to pattern with L2 adults supporting Meisel's hypothesis. Older children did better than younger ones due to their more advanced L1 literacy, cognitive and metalinguistic skills as evidenced in their less diverse performance compared to younger ones but also due to more instruction in their day schools. Finally, L1 effects were not pronounced in main thematic verbs but were visible when we considered individual results, the use of bare forms, and patterns within the 'is + verb(x)' structure.

References Gagarina, N., D. Klop, S. Kunnari, K. Tantele, T. Välimaa, I. Balciuniene, U. Bohnacker, & J. Walters (2012). MAIN: Multilingual Assessment Instrument for Narratives. *ZAS Papers in Linguistics*, 56. Meisel, J. M. (2009). Second language acquisition in early childhood. *Zeitschrift für Sprachwissenschaft*, 28(1), 5-34. <doi:10.1515/ZFSW.2009.002> Renfrew, C. (1998). *The Renfrew Language Scales: Word Finding Vocabulary Test*. Milton Keynes: Speechmark.

Tracking the Development of Noun-Adjective Collocations in L2 Learners of Portuguese: A Learner Corpus Analysis

Dogus Oksuz (University of Birmingham), Soroosh Akef (University of Lisbon), Amália Mendes (University of Lisbon) and Patrick Rebuschat (Lancaster University).

Usage-based approaches to language learning view multi-word expressions (MWEs), such as collocations, idioms, binomials and lexical bundles, as integral to the mental lexicon. These MWEs sit alongside single words and are thus understood to be building blocks of language (e.g., Arnon, McCauley & Christiansen, 2017). There is compelling evidence that L1 and L2 speakers differ in their reliance on MWEs. In a computational modelling study based on learner corpus data, McCauley and Christiansen (2017) found that L2 adults used fewer MWEs than L1 speakers. Pullido (2022) identified proficiency and L1 as critical variables affecting L2 speakers' processing of MWEs. Following up on these findings, we examined the extent to which learners' proficiency levels and L1s impact their use of noun-adjective and adjective-noun collocations, (i.e. "association between lexical words", Biber et al., 1999).

Our paper addresses the following research questions. 1). Do learners' proficiency levels affect their use of collocations, as measured by phrasal frequencies and commonly-used association measures MI, Log Dice, and Delta P? 2) Do learners' native languages (18 L1 groups) affect their use of collocations in L2 Portuguese? 3. Do these variables predict learners' errors in the use of collocations?

We analysed data from the COPLE2 learner corpus of European Portuguese to investigate MWE production in L2 learners (n=418). Examining noun-adjective and adjective-noun collocations allowed us to investigate a potential L1 influence in collocation use by those L1 groups which only allow adjectives to precede nouns. We used a Python script to extract noun-adjective collocations from the XML files of the written subset of COPLE2. The data consisted of 1,444 texts and 262,231 tokens, from which 6,575 collocations were extracted. We then gathered frequency counts for these collocations from the written European Portuguese subset of a widely-used Portuguese reference corpus (CRPC), containing 248,974,314 tokens.

The results of our mixed-effects models reveal a complex picture, in which higher proficiency learners do not always produce more frequent and more idiomatic collocations. While learners at B1 level begin to produce frequent collocations, only C1 learners use both frequent and idiomatic collocations consistently. Our finding aligns with McCauley & Christiansen (2017) adult L2 learners experience difficulty in using MWEs. Regarding L1 influence, we found no significant differences in the frequency and idiomaticity of collocations between learner groups whose L1s allow nouns to precede adjectives (e.g. Spanish, Romanian) and adjectives to precede nouns (e.g. Japanese). This shows that differences between word-orders of learners' L1s and target language do not appear to create any additional difficulty for learners' production of collocations. Regarding learners' errors, early analyses show that as collocational frequency increases, learners use them more accurately. Providing evidence from an under-studied language, this study adds to a growing body of evidence showing that L2 speakers rely less on MWEs than L1 speakers.

REFERENCES

- Arnon, I. & Christiansen, M.H. (2017). The role of multiword building blocks in explaining L1-L2 differences Topics in Cognitive Science, 9, 621-636.
- Pulido, Manuel F. (2022) Why are multiword units hard to acquire for late L2 learners? Insights from cognitive science on adult learning, processing, and retrieval. *Linguistics Vanguard*, 8(1), 237-247.

Task complexity effects on L2 pronunciation: exploring trade-offs between lexico-grammar and pronunciation

Mireia Ortega (University of Barcelona), Cristina Aliaga-Garcia (University of Barcelona), Ingrid Mora-Plaza (University of Barcelona) and Joan C. Mora (University of Barcelona).

The Cognition Hypothesis (Robinson, 2011) postulates that a more cognitively complex task generates increased attention to form-function mappings, resulting in greater lexical and syntactic complexity, and more accurate speech, usually at the expense of fluency. Preliminary research suggests that task complexity may be beneficial for L2 phonological development in pronunciation-focused tasks (Gordon, 2021; Solon et al., 2017). However, does increasing task complexity benefit L2 segmental accuracy in pronunciation-unfocused tasks? This study aims to answer this question by examining task complexity effects on learner's perception of task performance, segmental accuracy and comprehensibility and accentedness, and assesses potential trade-offs between lexico-grammatical and pronunciation accuracy. Eighty-two Spanish/Catalan EFL learners performed simple and complex versions of a monologic task (The dinner party; Ur, 1981) differing in \pm reasoning demands in which the contrastive target vowels /æ, ʌ/ and /i:, ɪ/ (n=36) and oral stops /p, t, k/ (n=25) were embedded in key words required to complete the tasks. Learners were given 1.5-minute planning time before the task and rated their task performance (difficulty, mental effort, anxiety) after it on a 9-point scale. Lexico-grammatical accuracy was measured through % error-free AS-units (EFASU). Pronunciation accuracy was measured through Mahalanobis distances between contrastive vowels and between native speakers' and learners' vowel productions in a Bark-normalized vowel space, through VOT duration (ms) for oral stops, and through comprehensibility and accentedness ratings. Results revealed that the complex task was perceived as more difficult, requiring more mental effort, and generating more anxiety. Contrary to the Cognition Hypothesis, complexifying unfocused tasks hindered production accuracy, and led to less distinct target-like vowels, shorter L1-like VOT and overall less comprehensible and more strongly accented speech, in line with Crowther et al. (2016). Interestingly, no trade-offs were observed between lexico-grammar and pronunciation, rather learners with higher EFASU scores also exhibited higher pronunciation accuracy. Overall findings suggest that increasing task demands may be detrimental to pronunciation unless tasks are designed with a specific focus on phonetic form.

Grammatical features in intermediate-level early vs late L2 sentence processing: Age and feature asymmetries

David Öwerdieck (TU Braunschweig) and Holger Hopp (TU Braunschweig).

There continues to be debate about the use of grammatical information in L2 sentence processing (Clahsen & Felser, 2018; Cunnings, 2017; Hopp, 2010). This study compares sentence processing of *wh*-questions in early vs. late L2 learners to test for potential age-induced asymmetries in the use of grammatical features.

Mismatches in grammatical features lead to less cue-based retrieval interference and aid processing of object-first strings among native speakers (Cunnings, 2017). For instance, object questions are harder to process than subject questions in L1 and L2 (Avrutin, 2000; Friedmann et al., 2009; Pontikas et al., 2022); however, a mismatch in number between NP1 and NP2 (1a) makes comprehension of object questions easier in child L1 and early bilingual acquisition (Contemori et al., 2018; Pontikas et al., 2022). Similarly, the presence of a pronoun (2.a) facilitates processing in L1 children compared to cases with two full nouns (2b) (Friedmann et al., 2009).

This study extends the inquiry to sequential L2 learners and covers different types of grammatical feature mismatches (number agreement and NP form). For L1 German-L2 English, adolescents (N = 28, mean age: 13.0 yrs, mean score on Crosslinguistic naming [CLT]: 57%) were compared with adults (N = 32, mean age: 23.9 yrs, mean score on Cambridge placement: 70%) in two picture-selection experiments with auditory stimuli as in (1a-b) and (2a-b). Accuracy and RTs were collected.

The subject-object asymmetry held for both groups (Figure 1), with adults having higher accuracy in both experiments. No significant effect of number mismatch was found, but there was an effect of NP form mismatch, i.e. accuracy was higher on (2a) than (2b) (Figure 2). RTs showed a similar pattern. These findings indicate continuity in sentence processing among early and late L2 learners. The difference between number and NP form mismatches may be due to their different grammatical status (cf. Rizzi, 2018). We will discuss the asymmetries in the use of grammatical features in the context of cue-based retrieval processing models and grammatical models of Relativized Minimality.

(1) a. Which bear are the camels chasing? / Which bears is the camel chasing? b. Which bear is the camel chasing? / Which bears are the camels chasing?

(2) a. Which pilot is she painting? b. Which pilot is the girl painting?

Selected References Avrutin (2000). Comprehension of discourse-linked and non-discourse-linked questions. *Language and brain* (pp. 295-313). • Clahsen & Felser (2018). Some notes on the shallow structure hypothesis. *SSLA*. • Contemori et al. (2018). On-line processing of English *which*-questions by children and adults. *JCL*. • Cunnings (2017). Parsing and working memory in bilingual sentence processing. *BLC*. • Friedmann et al. (2009). Relativized relatives. *Lingua*. • Hopp (2010). Ultimate attainment in L2 inflection. *Lingua*. • Pontikas et al. (2022). Online processing of *which*-questions in bilingual children. *JCL*. • Rizzi (2018). Intervention effects in grammar and language acquisition. *Probus*.

“vallah war geschockt” - Subject-drop in heritage and monolingual speakers of German

Onur Özsoy (Leibniz-Centre General Linguistics), Mariya Burbelko (Leibniz-Centre General Linguistics) and Borbála Sallai (University of Warwick).

German is regarded as a language with optional topic drop in certain registers (Fries, 1988) and under certain grammatical constraints (Trutkowski, 2016). The use of topic drop is often characterized as sub-standard. For example, it is characteristic of children’s speech (Hamann, 2009) and some L2 learners (Clahsen & Hong, 1995). While heritage speakers often acquire German from early age on and can be considered native(-like), they are mainly exposed to their heritage language in the first years. We investigate whether bilingual heritage speakers with a pro-drop heritage language apply topic drop more widely in German, and argue that an expansion in topic drop represents a creative extension of German grammar which points to reorganization at the pragmatics-syntax interface. We focus on the influence that bilingual speakers’ heritage Greek, Russian or Turkish might have on their majority German. Greek is a consistent pro-drop language where null subjects appear regardless of grammatical features. Russian is a partial pro-drop language or a language with abundant subject ellipsis. Turkish is a topic pro-drop language that only requires overt subjects in unspecified contexts. German is a topic-drop language but it is sometimes utilized as a non-pro-drop language (Müller, 2006; Trutkowski, 2016). In this area, little is known about the effect of the heritage language on the majority language. We derive the following research question for a corpus study on RUEG data (Wiese et al., 2021): Does Greek, Russian and Turkish heritage speakers’ expression of (pro)nominal reference in German align with monolingual speakers’ productions? There are four participant groups (Greek-German-bilinguals n=48, Russian-German-bilinguals n=61, Turkish-German-bilinguals n=64, monolinguals n=64). Our data are controlled naturalistic narrations in different settings. We manually annotate the data (455,208 tokens) with respect to subject realization. Currently, we have annotated and analyzed 1,258 subject positions. With the full data, we plan to run Bayesian binomial generalized linear mixed-effects models with realization (overt vs dropped) as a dependent variable and the following independent variables: Group (4 levels), Formality (2 levels), Mode (2 levels), Animacy (2 levels), Age of Onset (continuous) and random effects by Participant. So far, we found that 12% of all subjects are topic-dropped. Currently our sample is too small to make any conclusions on a by-group-level regarding the hypothesis. However, preliminary results (to be taken with caution) point to no considerable differences between monolingual and heritage speakers. We plan to present data from all 237 speakers who produced more than 30,000 subjects. Our findings might reveal how bilingual speakers apply a strategy that is available in the standard grammar and extend it creatively to form a new system of subject drop in German.

References: Clahsen, H., & Hong, U. (1995). Agreement and null subjects in German L2 development: New evidence from reaction-time experiments. *Second Language Research*. • Fries, N. (1988). Über das Null-Topik im Deutschen. *Sprache & Pragmatik*. • Müller, G. (2006). Pro-drop and impoverishment. *Form, Structure, and Grammar*. Berlin: Akademie Verlag. • Trutkowski, E. (2016). Topic drop and null subjects in German. Berlin: De Gruyter. • Wiese, H., Alexiadou, A., Allen, S., et al. (2021). RUEG Corpus.

Multilingual Experience in Language Aptitude and Working Memory

Elifcan Öztekin (Boğaziçi University) and Gülcan Erçetin (Boğaziçi University).

Definitions of language aptitude (LA) and working memory (WM) suggest interactions between them at the construct level (e.g., Wen & Skehan, 2021); while empirical research presents varying results about their interactions. Multilingualism adds an interface to these factors through the extended experience of managing more than two languages (e.g., Cockcroft, 2022; Huang et al, 2021). To provide an overall picture, the present study investigates the relationships between visuospatial and verbal WM and LA in a third language (L3) learning context in Turkey. The participants were 110 English language teaching majors who study an L3 at beginner, intermediate, and upper-intermediate / advanced levels, including a control group with no L3. They took the LLAMA version 3 (Meara & Rogers, 2019), forward digit span (phonological short-term memory), rotation, symmetry (visuospatial WM), and operation (verbal WM) span tasks (Foster et al., 2015). Principal component analysis (PCA) revealed a 3-factor solution, pointing to phonological-verbal (digit span, operation span, and LLAMA D), visuospatial (symmetry and rotation span), and aptitude (LLAMA B, E, F) as separate factors. A multivariate analysis of variance to examine the influence of multilingualism (i.e., the number of semesters the participants took their L3s) on the factors derived in the PCA revealed a statistically significant multilingualism effect on visuospatial WM [$F(3, 74) = 4.180, p < .01, \text{partial } \eta^2 = .145$]. The intermediate L3 students outperformed the upper-intermediate L3 group in the visuospatial factor. All the other pairwise comparisons between the proficiency levels were non-significant. This multilingualism influence on visuospatial WM is discussed in relation to the dynamic positions of the verbal and visual modalities of cognitive factors across the L3 proficiency development.

References Cockcroft, K. (2022): Working memory in multilinguals: factors that assist and impede. *International Journal of Multilingualism*, DOI: 10.1080/14790718.2022.2085710 Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., & Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. *Memory & Cognition*, 43(2), 226-236. Huang, Y., Wang, M., & Rao, H. (2021): Working memory as a partial mediator of the relationship between multilingualism and foreign language aptitude. *International Journal of Multilingualism*. DOI: 10.1080/14790718.2021.1993860 Meara, P. & Rogers, V. E. (2019). *The LLAMA Tests v3*. Cardiff: Lognostics. Wen, Z. & Skehan, P. (2021). Stages of acquisition and the P/E model of working memory: Complementary or contrasting approaches to foreign language aptitude?. *Annual Review of Applied Linguistics*, 41. 6-24.

Development and validation of domain-specific working memory measurements

Junlan Pan (University of York) and Emma Marsden (University of York).

Despite an increasing number of studies into the role of working memory (WM) in explaining variations in L2 learning outcomes, inconsistent findings have emerged that may be due to the different measures used to gauge WM as a multi-faceted construct (Juffs & Harrington, 2011). Systematic reviews of WM tasks used in L2 research highlight the need to develop standardised measurements, so that WM can be measured accurately in ways relevant to L2 learning (Shin & Hu, 2020). This study addressed four main challenges in WM testing research to date: the need for: (1) an underpinning theoretical framework, (2) language-sensitive stimuli to counter potential confounds with L2 knowledge (as most WM tests have been written in English, the L2, itself), (3) examining the potential role of modality effects, and (4) checking reliability and (internal) validity of instrumentation. In the current study, for (1), we based our measurements on the Phonological/Executive Model (Wen, 2016) using two subtests, i.e., Serial Non-word Recall (SNWR) and Complex Span Task (CST) to measure phonological short-term memory and executive control capacity, respectively. For (2), we used participants' L1-Chinese as the encoding language: for SNWR, non-word stimuli conformed to the phonology of Mandarin and avoided real meaning associations (which can be challenging in Mandarin, as individual syllables may correspond to one or more meanings); for CST, we used Chinese stimuli for the sentences, controlled for sentence length and for the location and usage rate of the lexical cues used to elicit comprehension. For (3), we examined the effects of modality by administering the CST in the aural and written modalities using a within-subject design. Data from 165 college-level L1 Chinese - L2 English participants, in two rounds of tests (using two different versions of stimuli, counterbalanced across the aural and written modalities), were analysed using principal component analysis. Although both subtests generally had acceptable internal consistency and unidimensionality, suggesting that they are reliable measures of the WM components, CST in the written modality had lower coefficients. The acceptable quality of the item was evidenced by applying the Generalised Partial Credit Model on Item Response Theory. However, the results of CST in the written modality showed a ceiling effect, indicating that this task lacked challenge for the participants in the current research. Results from linear mixed-effects modelling evidenced that the modality was a significant fixed effect when measuring WM. We aim to showcase a thorough scrutiny of these two psychometric measurements prior to their use in substantive research, a neglected step in promoting methodological rigour in SLA. We provide open materials, data, and analysis code to serve further scrutiny and replication.

References Juffs, A., & Harrington, M. (2011). Aspects of working memory in L2 learning. *Language Teaching*, 44(2), 137–166. Shin & Hu (2022). A methodological synthesis of working memory tasks in L2 research. In J. W. Schwieter, & Z. Wen (Eds.), *The Cambridge Handbook of Working Memory and Language*. Wen, Z. (2016). *Working Memory and Second Language Learning: Towards an Integrated Approach*.

***Predecessors of L2 grit and their complex interactions in online SLA:
Complementary contributions of regression and psychological network analysis***

Michał B. Paradowski (Institute of Applied Linguistics, University of Warsaw) and Magdalena Jelińska (Institute of Applied Linguistics, University of Warsaw).

Learning a foreign language is a long-term process requiring persistence and a willingness to engage in activities that will help develop communicative competence. Viewed from a complex dynamic systems perspective, all these factors form complex, interrelated and interdependent systems that significantly shape the development of second language competence (Larsen-Freeman & Cameron, 2008; Oxford, 2016a, 2016b; Mercer, 2018). An important role on the way to achieving linguistic proficiency is played by L2 grit. However, although a handful of studies provide some insight into the contribution of L2 grit to the development of L2 competence, we still do not know the reasons why learners demonstrate different levels of this trait and subsequently the extent of their L2 achievement (Elahi Shirvan & Alamer, 2022; Teimouri, Sudina & Plonsky, 2021). We identify the most important predictors of L2 grit in the online language learning context, an environment which raises many new challenges for learners. Basing on a literature review, we first identify possible predictors of L2 grit in the generic process of L2 learning, which include language mindsets and two general psychological dispositions: a sense of autonomy and curiosity, and one predictor specifically related to the learning context, i.e., readiness for online learning. A multiple linear regression model built using questionnaire data collected from 615 remote/hybrid language learners from 69 countries (Paradowski & Jelińska, 2023) reveals that L2 grit is most determined by two dimensions of readiness for online learning, and to a lesser extent by learners' autonomy and curiosity (Table 1). The regression model is significant ($F_{6,608} = 97.28, p < .001$) and predicts 48% of variance in the dependent variable, with a very large effect size ($\eta^2 = .49$). To provide a deeper insight into the complex system of interconnected factors, we subsequently apply psychological network analysis. This type of network analysis, capable of illustrating dynamic relationships between individual characteristics and second/foreign language learning, has not yet received much attention from researchers in the SLA field (Freeborn, Andringa, Lunansky & Rispens, 2022). With the use of a regularised partial correlation approach applying a Gaussian graphical model with the least absolute shrinkage and selection operator (LASSO) regularisation technique combined with Extended Bayesian Information Criterion (EBIC) model selection to minimise spurious edges, the analysis provides a clearer and more precise insight into the relationship between L2 grit and the variables examined. The results permitted estimating the relative importance of the factors in the complex network of associations, indicating equally strong direct connections between L2 grit and both dimensions of readiness for online learning, and a much weaker edge linking L2 grit with autonomy (Fig. 1). These patterns of relationships corroborate the need to include a complex systems theoretical perspective into research including individual differences. The findings have practical implications for language teachers in remote and hybrid settings. Paradowski, M.B. & Jelińska, M. (2023). The predictors of L2 grit and their complex interactions in online foreign language learning: Motivation, self-directed learning, autonomy, curiosity, and language mindsets. *Computer Assisted Language Learning*. <https://doi.org/10.1080/09588221.2023.2192762>

Reframing the Role of Gender in the L2 Acquisition of Russian Case

Natalia Parker (University College London) and Clare Wright (University of Leeds).

Though the acquisition of Russian gender and of Russian case have been actively explored in both Second Language Acquisition (SLA) and language pedagogy (e.g., Cherepovskaia et al., 2021; Kempe & Brooks, 2008; Taraban & Kempe, 1999), theoretical accounts remain underspecified how gender may influence acquisition of Russian case morphology and how that affects the learning of case in a beginner-level language classroom.

This paper will present the results of an iterative intervention-based experiment on case acquisition, based on Bock & Levelt's (2002) production model, which argues that morphological noun inflection may depend on multiple steps for function assignment and constituent assembly. Acquisition of two Russian cases (Prepositional and Accusative) was investigated within two gender subsets (M/F). In Russian, feminine gender marking is seen as more complex, since feminine nouns need case inflectional changes from the base form for both case functions, whereas masculine nouns remain uninflected (Accusative) or require simply adding a case marker to the base form (Prepositional). There is a case difference in that Prepositional needs a single case marker form, while Accusative has different case marking for the two genders. These differences in case marking allow for comparison in production accuracy, to test the prediction that more complex feminine marking is harder to produce accurately than masculine forms. Beginner learners, recruited as volunteers in a UK university setting (n= 27), were given an hour's teaching a week for 10 weeks, in a form-meaning integrated approach to introduce the functions and forms of the two cases. Using a pre-posttest design, learners' accuracy in case production was tested using specifically-designed elicitation methodology, to examine potential gender difference patterns in assigning correct form to function in different contexts, such as locationality, temporality, directionality; testing was equally balanced across three different functions per case.

Results in the post-tests (from Mann-Whitney U test) revealed significant impact of gender ($p = 0.049$) for all Accusative-marked nouns, where feminine marking was consistently less accurate. There was also a clear (ns) case function effect for Prepositional marking; both masculine and feminine nouns varied in accuracy, and in different contexts. Within the limitations of this exploratory study, our findings confirm the impact of underlying gender differentiation in Russian case production, but suggest that overtness of morphological marking and context function play a greater role in the processing load than previously understood. We argue that case acquisition research in L2 Russian needs to take the interplay of these factors more into account, both theoretically and in relation to pedagogic practice in the classroom.

Cherepovskaia, N., Slioussar, N. & Denissenko A. (2021). Acquisition of the nominal case system in Russian as a second language. *Second Language Research*. doi:10.1177/0267658320988058

Bock, K., & Levelt, W. (2002). Language production: Grammatical encoding. In (ed.) Altmann, G.T.M., *Psycholinguistics: Critical concepts in psychology*, 5, Routledge, 405-452.

Kempe, V., & Brooks, P. J. (2008). Second language learning of complex inflectional systems. *Language Learning*, 58(4), 703-746.

Taraban, R., & Kempe, V. (1999). Gender processing in native and nonnative Russian speakers. *Applied Psycholinguistics*, 20(1), 119-148.

Maximising the potential of L2 input for early L3 development: The case of plurilingual audiovisual input

Anastasia Pattemore (University of Groningen), Marilyn Lopez (NHL Stenden), Beatriz Cabrera Fernandez (University of Groningen) and Marije Michel (University of Groningen).

In this study we propose the use of plurilingual audiovisual input - combining viewers' L2 English in the soundtrack, and an L3 target language in the subtitles - to promote early acquisition from original version television. Research has consistently shown the positive effects of original version television for second language (L2) development (Montero Perez, 2022). However, mere exposure to audiovisual input is not sufficient for language learning to occur. Processing audiovisual input requires a certain proficiency level - Danan (2004) suggests that an intermediate level is the necessary proficiency threshold - and consequently much of the audiovisual input research to date has been into intermediate and higher proficiency learners. Accordingly, we lack research dedicated to beginner learners, and it remains an empirical question how audiovisual input could be used to support these lower proficiency level learners (pre A1-A2). One solution that audiovisual input researchers have been exploring is viewing with L2 audio but with subtitles in the viewers' L1 (e.g. Gesa & Miralpeix, 2022). Yet, this poses problems for gaining exposure to L2s other than English, as a large proportion of TV shows and films are originally in English, with English audio. Although one could work with English media dubbed into another L2, this may lack ecological validity. In this study we follow the multilingual approach to language learning and teaching where several foreign languages support acquisition of one another (Cenoz & Gorter, 2021).

Forty-six Dutch learners of Spanish with a variety of L1s other than English or Spanish participated in the study, with the majority being L1 Dutch. All were beginner learners of Spanish in their first months of language instruction, but had at least B2 English proficiency. While a comparison group (n=17) only participated in the pre-/post-tests, experimental participants (n=29) watched twelve full-length episodes of a comedy series (264 minutes) in their L2 English with L3 Spanish subtitles over four weeks. The target items were fifty non-cognate (Dutch and English) lexical and multiword units that frequently appeared in the TV series (from five to sixty times), such as *caja fuerte* and *sumidero*. We tested students' learning with a Spanish adaptation of the Vocabulary Knowledge Scale (VKS, Wesche & Paribakht, 1996), also including distractors (n=15). The VKS consists of a self-reporting elicitation scale where learners can identify the level of the target's familiarity and knowledge, from initial recognition to ability to use it in a context. It elicits both self-perceived and demonstrated knowledge. Participants' Spanish (Lextale-Esp, Izura et al., 2014) and English (V_YesNo, Meara & Miralpeix, 2015) vocabulary size was tested as a proxy of proficiency. Initial results indicate L3 vocabulary gains when using the plurilingual subtitled input. Furthermore, students report that they enjoyed this context. In our presentation we will present the complete statistical analyses with participant-level (group, English and Spanish vocabulary sizes) and item-level (pre-/posttest, frequency) variables. In light of the multilingual turn (Cenoz & Gorter, 2021), we will discuss how using plurilingual subtitles could be an excellent method for beginner learners to start exposing themselves to original version television.

The effectiveness of different types of automated feedback in digital language learning games for children

Matthew Pattermore (University of Barcelona) and Roger Gilabert (University of Barcelona).

Language education is increasingly integrating technology into classrooms, and this process has been hastened by the necessities of the recent pandemic. Digital games are a promising avenue for language learning as they can combine intrinsic motivation with significant amounts of meaningful language input and use (Reinhart, 2018). One of the touted benefits of using games for language learning is the increased amount of personalised feedback students can receive (Cornillie et al., 2012). However, automated feedback within digital games is different from classroom-based oral or written feedback. When the software acts as tutor, there is no space for negotiation of meaning, additional clarifications, or flexible output, and additionally the feedback must be explicitly presented in order to align with the narrative nature of the game format. Therefore, many of the traditionally studied types of corrective feedback in SLA may be functionally different or inapplicable in an automated digital context. Further research is required to determine which forms of feedback are most effective in this context.

In this study, we use data from 744 Spanish twelve-year-old EFL students from eight schools around Spain. The learners played 67,623 instances of 15 language-focused minigames, covering 224 different language features applicable to reading skills (e.g. phonology, morphology, syntax), over the period of a year. The minigames were developed as part of the Horizon 2020 iRead project. The system automatically logged each game a student played, as well as any errors they made (n=153,723). The automated feedback given to learners when they made an error was categorised into four types based on the work of Johnson et al. (2017): outcome feedback (e.g. 'try again'), metalinguistic feedback (direct explanation), informational feedback (providing guidance on how to derive the answer), and hints (giving examples). Using an Item Response Theory-based analysis (Debeer et al., 2021), which included controlling for learner ability (a latent trait based on overall success within the game system), minigame and feature difficulty, and improvement over time, we examine the extent to which each type of feedback contributed to successful recovery from errors. Results suggest that non-hint elaborative feedback (i.e. informational and metalinguistic) is most effective in this context.

These results will be discussed in the light of SLA theories regarding feedback, and the implications for language learning game design and pedagogic utilisation.

References:

- Cornillie, F., et al. (2012). ReCALL special issue: Digital games for language learning: challenges and opportunities: Editorial Digital games for language learning: from hype to insight?. *ReCALL*, 24(3), 243-256.
- Debeer, D., et al. (2021). The effect of adaptivity in digital learning technologies. Modelling learning efficiency using data from an educational game. *BJET*.
- Johnson, C. I., et al. (2017). Designing effective feedback messages in serious games and simulations: A research review. *Instructional techniques to facilitate learning and motivation of serious games*, 119-140.
- Reinhardt, J. (2018). *Gameful second and foreign language teaching and learning: Theory, research, and practice*. Springer.

Measuring Aptitude for Explicit and Implicit L2 Learning

Renato Pavlekovic (University of Essex) and Karen Roehr-Brackin (University of Essex).

Language learning aptitude is a consistent predictor of second language (L2) success (Li, 2015, 2016). Research to date suggests that the construct comprises aptitude for explicit and aptitude for implicit learning (Suzuki & DeKeyser, 2017). Moreover, it has been proposed that working memory should be considered a component of aptitude (Wen et al., 2019). In the field of L2 learning, researchers increasingly draw on the LLAMA test battery as a freely available and easy-to-use measure of aptitude. The LLAMA is undergoing continuous improvement, and the most recent version (Meara & Rogers, 2019) has implemented changes to all four subtests. In particular, LLAMA E and F no longer have a two-way multiple-choice answer format, which should improve reliability. Research on implicit aptitude has shown that sound sequence recognition ability (LLAMA D) and implicit sequence learning ability as measured by a serial reaction time task (SRT; Kaufman et al., 2010) load on separate factors, indicating that aptitude for implicit learning may be componential (Granena, 2019, 2020), just like aptitude for explicit learning. In order to further improve the conceptualisation and measurement of aptitude in L2 learning, we investigated whether an adapted, partial-credit scoring system applied to LLAMA E and F would improve consistency of measurement. Furthermore, we examined the relationship between measures of aptitude for explicit and implicit learning and working memory as well as whether the various aptitude subtests and/or working memory would predict L2 English proficiency in 86 L1 Croatian learners aged 15-18. The participants completed the LLAMA battery and a probabilistic SRT task as measures of aptitude for explicit and implicit learning. An automated operation span task (Unsworth et al., 2005) and a forward digit span task (Wechsler, 1997) were used to assess executive working memory and phonological short-term memory, respectively. The Oxford English Language Placement test (Oxford University Press, n.d.) and a gap-fill task served as measures of L2 proficiency. Our findings show substantially improved reliability for both LLAMA E and F on the basis of our partial-credit scoring system. Correlational and factor analyses suggest that aptitude for explicit and implicit learning can be distinguished, as expected. At the same time, we found evidence for a componential structure of implicit aptitude, with sound sequence recognition (LLAMA D) and implicit sequence learning (SRT) loading on separate factors. Intriguingly, factor loadings also indicate that language-analytic ability (LLAMA F) and implicit sequence learning ability (SRT) are related, suggesting that over and above aptitude for explicit and implicit learning, aptitude for item learning and aptitude for pattern recognition and induction may be distinguishable. As executive working memory was associated with the ability to learn novel lexical items (LLAMA B), and as both measures of working memory loaded on the same factor with sound sequence recognition (LLAMA D), we found evidence that working memory can be construed as a component of language learning aptitude. Last but not least, our results demonstrate that both explicit and implicit aptitude can predict learners' L2 proficiency, thus providing evidence for the predictive validity of the LLAMA.

Examining the role of grit in learning L2 grammar: Links to motivation, self-efficacy and achievement

Mirosław Pawlak (Adam Mickiewicz University, Poznań, Poland).

Although grit, understood as perseverance and passion for long-term goals (Duckworth et al., 2007), has only recently attracted the attention of SLA researchers, relevant studies are on the increase. Attempts have been made to relate this construct to other individual difference (ID) factors such as motivation, willingness to communicate or emotions as well as attainment (e.g., Feng & Papi, 2020), to tap into its dynamicity (e.g., Elahi Shirvan et al., 2021), and to explore it among teachers (e.g., Sudina et al., 2021). A major development was the construction of the L2 Grit Scale (Teimouri et al., 2022), gauging perseverance of effort and consistency of interest, which allowed investigating the construct in L2 learning, as is typically the case with other ID factors (Dörnyei & Ryan, 2015). Building on this reasoning, the study examined grit with respect to learning L2 grammar by (1) investigating the link between grit and grammar achievement, and (2) exploring how this relationship is mediated by motivation and self-efficacy. Data were collected from 240 Polish university students by means of the newly developed Grit in Learning L2 Grammar Scale, which is an adaptation of the L2 Grit Scale (Teimouri et al., 2020), the Motivated Behavior Scale and the Self-Efficacy Scale (Piniel & Csizér, 2013). Achievement was operationalized as scores on the grammar-focused component of the mandatory end-of-the-year examination. Confirmatory factor analysis supported a two-factor model of grit for learning L2 grammar. Regression analysis demonstrated that, on the whole, grit was a stronger predictor of grammar attainment than motivated behavior and self-efficacy, with perseverance of effort playing a more important role than consistency of interest. The picture became more nuanced when learners representing three levels of attainment were considered, with the contribution of motivation growing in the case of low-achievers and self-efficacy in the case of high-achievers.

References Dörnyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited*. Routledge.
Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long term goals. *Journal of Personality and Social Psychology*, 92(6), 1087-1101.
Elahi Shirvan, M., Taherian, T., Shahnama, M., & Yazdanmehr, E. (2021). A longitudinal study of foreign language enjoyment and L2 grit: A latent growth curve modeling. *Frontiers in Psychology*, 12.
<https://doi.org/10.3389/fpsyg.2021.720326>
Feng, L., & Papi, M. (2020). Persistence in language learning: The role of grit and future self-guides. *Learning and Individual Differences*, 81.
Mahmoodzadeh, M., & Khajavy, G. H. (2019). Towards conceptualizing language learning curiosity in SLA: An empirical study. *Journal of Psycholinguistic Research*, 48, 333-351.
Piniel, K., & Csizer, K. (2013). L2 motivation, anxiety and self-efficacy: The interrelationship of individual variables in the secondary school context. *Studies in second Language Learning and Teaching*, 3, 523-550.
Sudina, E., Vernon, T., Foster, H., Del Villano, H., Hernandez, S., Beck, D., & Plonsky, L. (2021). Development and initial validation of the L2-teacher grit scale. *TESOL Quarterly*, 55, 156-188.
Teimouri, Y., Plonsky, L., & Tabandeh, F. (2020). L2 grit: Passion and perseverance for second-language learning. *Language Teaching Research*, 26, 893-918

Understanding L2 repair fluency: Perspectives of L1 repair fluency, cognitive fluency, and language anxiety

Pauliina Peltonen (University of Turku), Sanna Olkkonen (University of Turku), Magdalena Szyszka (University of Turku) and Pekka Lintunen (University of Turku).

Repairs (including false starts, repetitions, and different types of self-corrections) have been examined in second language (L2) speech fluency research as an indicator of (dis)fluent speech (utterance fluency; Segalowitz, 2010). However, compared to other dimensions of utterance fluency (speed and breakdown; see Skehan, 2009), repair fluency is not equally well understood. Previous research suggests individual variation in L2 repair use and potential links between first language (L1) and L2 repair fluency, but more research is needed on the roles of cognitive fluency and affective factors in influencing L2 repair fluency patterns. While some previous studies have explored links between cognitive and utterance fluency (e.g., Kahng, 2020) or utterance fluency and affective factors (e.g., Aubrey, 2022), research incorporating all three aspects is rare (but see Zuniga & Simard, 2022).

To fill this gap in research, we combine perspectives of L1 repair fluency, cognitive fluency (attention control), and affective factors (language anxiety, LA) to achieve a comprehensive view of the factors behind L2 repair fluency. Triangulating research methods from SLA and psycholinguistics, we report on a study examining data from L1 Finnish and L2 English monologue speech tasks, a Stroop task in L1 and L2, and surveys for general and task-specific LA from the same participants (advanced users of English; N = 59). We investigated how L1 repair fluency, cognitive fluency, and LA are related to L2 repair fluency with a mixed-methods approach incorporating quantitative (correlations) and qualitative (profiles of frequent repairers) analyses. The study is part of a larger project examining fluency and disfluency in L2 speech.

Our results indicate that the factors behind L2 repair fluency behavior are complex and subject to individual variation. The quantitative correlational analyses indicated that task-specific LA and certain cognitive fluency measures were connected with L2 repair measures, while correlations between L1 and L2 repair fluency measures were weak. The qualitative analyses of repair profiles (n = 7) revealed that, overall, repairs are more common in the L2 than in the L1, but patterns regarding preferences for repair types vary across individuals. The study has methodological implications for psycholinguistic and SLA research into L2 repair fluency and broader implications for L2 teaching and assessment. The methodologically novel approach of the present study, triangulating research methods from SLA and psycholinguistics, brings the field of L2 fluency research forward by illustrating the complex dynamics behind L2 repair behavior.

References

Aubrey, S. (2022). The relationship between anxiety, enjoyment, and breakdown fluency during second language speaking tasks: An idiodynamic investigation. *Frontiers in Psychology*, 13, 968–946. Kahng, J. (2020). Explaining second language utterance fluency: Contribution of cognitive fluency and first language utterance fluency. *Applied Psycholinguistics*, 41, 457–480. Segalowitz, N. (2010). *Cognitive bases of second language fluency*. Routledge. Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30, 510–532. Zuniga, M. & Simard D. (2022). Exploring the intricate relationship between foreign language anxiety, attention and self-repairs during L2 speech production. *System*, 105, 1–10.

CLI on the acquisition of adjective placement and agreement by English and Italian L2 learners of German

Federica Pepe (University of Würzburg).

This paper investigates crosslinguistic influence (CLI) from a morphosyntactic perspective. Although nouns' and adjectives' morphosyntactic features are liable to CLI due to the considerable differences in terms of inflection and word order across languages, little attention has been paid to the acquisition of noun and adjective phrases [2, p. 669] and no consensus on whether the L1 plays a role [2,4,5] or not [1,3] in this area has been reached. Moreover, previous research was mostly conducted with bilinguals [3,5] or untutored students [2] and did not investigate learners' intuitions. To determine the influence of the L1 in terms of adjectival placement and agreement, intermediate learners of German with L1 English (n = 7) and L1 Italian (n = 7) were investigated in a grammaticality judgment task (GJT). The two languages show differences to German (Table 1) which are predicted to transfer to the L2. Table 1

Position	Predicative Position	Agreement Position	Agreement	Italian	+	NA	+	Il tavolo rosso
German	Il tavolo è rosso	+	AN	-	Der rote Tisch	Der Tisch ist		
English	-	AN	-	The red table	The table is red			

The GJT contained 56 experimental items, divided as follows: a)20: target-like b)12: agreement in predicative position (Italian-like), e.g. *Das Wetter ist schöne. (intended: ... schön-ø.) c)12: noun-adjective order (Italian-like), e.g. *Die Mädchen schlaun lesen ein Buch. (intended: Die schlaun Mädchen ...) d)12: missing agreement (English-like), *Die arrogant-ø Sekretärin telefoniert. (intended: Die arrogante Sekretärin ...) Results suggest that, contrary to previous research [4,5] and L1-transfer predictions, Italian learners accepted the target-deviant noun-adjective order extremely infrequently (2.4%) with no significant group difference. Although their L1 does not employ this structure, English participants identified Italian-like items with agreement in predicative position as grammatical more often than Italian speakers (English = 37%, Italian = 19%). Furthermore, the results support the prediction according to which English speakers would be significantly worse in identifying target-deviant items with missing agreement as ungrammatical. In conclusion, the findings suggest that CLI effects are more likely to emerge in judgments concerning the absence of morphosyntactic features: L2 learners correctly reject overtly marked target-deviant features, however L1 influence causes them to incorrectly accept target-deviant features that are not overtly-marked. The results will be discussed in light of the Failed Functional Features Hypothesis and markedness.

- Gass, S., & Ard, J. (1984). The ontology of language universals. In W. Rutherford (Ed.), *Language universals and second language acquisition* (33-68). John Benjamins.
- Parodi, T., Schwartz, D., & Clahsen, H. (2004). On the L2 acquisition of the morphosyntax of German nominals. *Linguistics* 42(3), 669-705.
- Rizzi, S., Gil, A., Repetto, V., Geveler, J., & Müller, N. (2013). Adjective placement in bilingual Romance-German and Romance-Romance children. *Studia linguistica*, 67(1), 123-147.
- Schlyter, S. (2001). Pragmatic rules, C-domain, and language dominance. *Bilingualism: Language and cognition* 4(1), 40-42.
- Volterra, V., & Taeschner, T. (1977). The acquisition and development of language by bilingual children. *Journal of child language* 5(2), 311-326.

The acquisition of Differential Object Marking and split ergativity: a longitudinal study on L1 Dutch learners of Hindi as a foreign language

Aaricia Ponnet (Ghent University) and Ludovic De Cuypere (Ghent University / VUB).

This paper investigates the acquisition of Differential Object Marking (DOM) and split ergativity in Hindi (an under-researched language) as a foreign language based on longitudinal data of five L1 Dutch speakers at a higher education institute in Belgium. A first aim of the study is to verify earlier findings in cross-sectional studies on L2 Hindi and Hindi as a heritage language. These studies have shown that DOM, which is governed by animacy and specificity, and split ergativity, which is governed by transitivity and perfectivity, in Hindi is particularly hard to acquire, yielding high omission rates of the direct object (DO) marker *-ko* and the ergative marker *-ne* for heritage speakers and advanced L2 learners (Baten & Verbeke 2015; Montrul et al. 2012, 2019; Ponnet et al. 2016; Ponnet & Decuypere *fc.*). A second aim is to investigate explanations put forward in earlier studies (Montrul et al. 2012, 2019; Narasimhan 2005, 2020; Baten and Ponnet 2023), which relate to L1 influence, item-based learning or semantic mapping. We elicited semi-spontaneous oral production data using a story retelling task. The learners' data involved four observation points over three semesters. We found high accuracy for split ergativity, and confirmed high omission rates for DOM, with a clear increase of the use of the DO marker towards the last observation point. Using a mixed effects logistic regression analysis, we found a significant association of split ergativity with perfectivity, and a significant increase in DOM marking for human animate, specific DOs, and also for non-human animate and inanimate specific nouns. Our findings suggest that learners acquire split ergativity before DOM, initially expanding the optionality of *-ko* due to the complex syntactic and semantic constraints (which may be enhanced by L1 transfer) before gradually starting to use *-ko* according to its constraints (which seems to emerge via item-based learning).

References: - Baten, K., & Ponnet, A. (2023). Extending PT to split ergative marking and differential object marking. *Processability and Language Acquisition in the Asia-Pacific Region*, 9, 91. - Baten, K., & Verbeke, S. (2015). The acquisition of the ergative case in Hindi as a foreign language. *Theoretical and methodological developments in Processability Theory*, 71-104. - Montrul, S. A., Bhatt, R. M., & Bhatia, A. (2012). Erosion of case and agreement in Hindi heritage speakers. *Linguistic Approaches to Bilingualism*, 2(2), 141-176. - Montrul, S., Bhatia, A., Bhatt, R., & Puri, V. (2019). Case marking in Hindi as the weaker language. *Frontiers in psychology*, 10, 461. - Narasimhan, B. (2005). Splitting the notion of 'agent': case-marking in early child Hindi. *Journal of Child Language*, 32(4), 787-803. - Narasimhan, N. (2020, August 31). The acquisition of differential case-marking: Evidence from Hindi [Conference keynote presentation]. SAFAL 2020, online via Zoom. - Ponnet, A., Baten, K., & Verbeke, S. (2016). The acquisition of differential object marking in Hindi as a foreign language. *Dutch Journal of Applied Linguistics*, 5(2), 101-125. - Ponnet, A., & Decuypere, L. (forthcoming). The acquisition of Hindi split-ergativity and Differential Object Marking by Dutch L1 speakers: systematicity and variation.

'Perfect' bilinguals? A comparison between heritage language speakers and late second language learners

Leonarda Prela (Friedrich Alexander University Erlangen-Nuremberg), Ewa Dabrowska (Friedrich Alexander University Erlangen-Nuremberg) and Miquel Llompart (Universitat Pompeu Fabra).

Several studies suggest that native-like attainment is modulated by biological constraints (Abrahamsson & Hyltenstam, 2008; DeKeyser, 2000; Johnson & Newport, 1989; Lenneberg, 1967). For example, linguistic

outcomes between heritage speakers (HSs) and late bilinguals are expected to differ due to the age of onset of bilingualism, with (early-onset) HSs resembling native monolinguals and late bilinguals rarely achieving nativelike proficiency. However, recent evidence suggests that an early onset is not a guarantee of native-like attainment, as many HSs show signs of incomplete (or divergent) acquisition (Kupisch & Rothman, 2018; Montrul, 2006). Moreover, similarly to late bilinguals, HSs exhibit vast individual differences (Torres et al., 2019), something which has even been attested in monolinguals (Dąbrowska, 2012). These differential outcomes do not only challenge the magnitude of age effects in language acquisition but also question the methodological utility of monolingual native speakers as a benchmark for ultimate attainment. Thus, the current study aims to provide a more complete account of bilinguals' linguistic abilities by comparing all three groups: HSs, late bilinguals, and monolinguals in both of their languages. We recruited 34 HSs of Greek, 35 L1-Greek L2 learners of English, 35 Greek monolinguals and 35 English monolinguals. The groups were matched for age (mean age = 41.6) and education. Participants completed a background questionnaire and an auditory grammaticality judgement task (GJT). Bilinguals completed the GJT in English and Greek, whereas the monolinguals did so only in their native language. Our analyses targeted comparisons between the two bilingual groups in both Greek and English as well as between these groups and the monolingual speakers of each language. We found significant differences between the bilingual groups in each language. In Greek, the late bilinguals and the monolinguals had indistinguishable performances contrary to the HSs' performance, which was much lower. In English, the HSs' performance matched that of the monolinguals' and both groups were better than the late bilinguals. These findings were consistent across both accuracy and reaction time measures. Crucially, differences in performance between HSs and monolinguals/late bilinguals in Greek were considerably larger than differences between late bilinguals and monolinguals/HSs in English (see Figure 1). Additionally, the observed individual differences in HSs were much larger than those of the late bilinguals. The observed deviance from monolingual controls has traditionally led to the characterization of bilinguals' weaker language as incomplete, when in fact the group comparisons themselves have been incomplete. By looking at all groups together, we see that late bilinguals' performance is quite high despite their later exposure to the second language. Furthermore, HSs' nativelike performance cannot always be guaranteed on the basis of early bilingualism onset. Instead, HSs should be treated as a subset of native speakers with divergent acquisition that can be extremely variable and may need to be supported. This is evident from the fact that HSs' performance in Greek ranges from 'perfect' nativelike to very poor performance.

References Abrahamsson, N., & Hyltenstam, K. (2008). THE ROBUSTNESS OF APTITUDE EFFECTS IN NEAR-NATIVE SECOND LANGUAGE ACQUISITION. *Studies in Second Language Acquisition*, 30(4), 481–509. <https://doi.org/10.1017/S027226310808073X>

Dąbrowska, E. (2012). Different speakers, different grammars: Individual differences in native language attainment. *Linguistic Approaches to Bilingualism*, 2(3), 219–253. <https://doi.org/10.1075/lab.2.3.01dab>

DeKeyser, R. M. (2000). THE ROBUSTNESS OF CRITICAL PERIOD EFFECTS IN SECOND LANGUAGE ACQUISITION. *Studies in Second Language Acquisition*, 22(4), 499–533. <https://doi.org/10.1017/S0272263100004022>

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21(1), 60–99. [https://doi.org/10.1016/0010-0285\(89\)90003-0](https://doi.org/10.1016/0010-0285(89)90003-0)

Kupisch, T., & Rothman, J. (2018). Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers. *International Journal of Bilingualism*, 22(5), 564–582. <https://doi.org/10.1177/1367006916654355>

Lenneberg, E. H. (1967). *Biological foundations of language*. Wiley.

Montrul, S. (2006). Incomplete acquisition as a feature of bilingual and L2 grammars. In *Inquiries in Linguistic Development* (pp. 335–359). John Benjamins Publishing Company. <https://doi.org/10.1075/z.133>

Torres, J., Estremera, R., & Mohamed, S. (2019). THE CONTRIBUTION OF PSYCHOSOCIAL AND BIOGRAPHICAL VARIABLES TO HERITAGE LANGUAGE LEARNERS' LINGUISTIC KNOWLEDGE OF SPANISH. *Studies in Second Language Acquisition*, 41(04), 695–719. <https://doi.org/10.1017/S0272263119000184>

Personality as a factor affecting the use of language learning strategies: The case of university students

Jakub Przybył (Adam Mickiewicz University, Poznan) and Mirosław Pawlak (Adam Mickiewicz University, Poznan).

Although personality has been recognised as an important factor influencing the process of second language learning and its outcomes (Dewaele, 2022), empirical evidence in this respect is still limited (Piechurska-Kuciel, 2020). This is also true about the link between the role of personality traits in determining the use of language learning strategies (LLS). While there are studies that have examined the relationship between the two constructs, their results are often inconclusive and most of them have focused on the contribution of extraversion-introversion dimension, neglecting other personality traits (Wakamoto, 2000; Chen & Hung, 2012; Zhou & Intaraprasert, 2015).

Responding to this research gap, we conducted a study in a group of 711 Polish university students, exploring relationships between LLS use, personality traits and attainment in learning English as a foreign language. Data on LLS and personality was collected by means of the modified Strategy Inventory for Language Learning (SILL) (Ardasheva & Tretter, 2013; Oxford, 1990) and the Polish adaptation of the NEO-FFI (Costa & McCrae, 1992), respectively, whereas attainment was established on the basis of university-administered placement tests. The collected data was analysed with the help of structural equation modelling, which is a major improvement on previous research.

We found that high levels of openness to experience were related to more frequent use of all LLS categories, while extraversion and conscientiousness were only positively related to the employment of social and memory/metacognitive LLS respectively. At the same time, significant differences in LLS use were revealed across cohorts of participants characterised by different levels of attainment, with highly open and more proficient students exhibiting greater preference for cognitive and metacognitive LLS than their less open and less proficient counterparts. The study provides empirical support for the claim that personality significantly affects successful L2 learning (cf. Dörnyei & Ryan, 2015).

References Ardasheva, Y., & Tretter, T. R. (2013). Strategy inventory for language learning – ELL student form: Testing for factorial validity. *Modern Language Journal*, 97(2), 474-489. Chen, M.-L., & Hung, L.-M. (2012). Personality type, perceptual style preferences, and strategies for learning English as a foreign language. *Social Behavior and Personality: An International Journal*, 40(9), 1501-1510. Costa, P. T. Jr. & McCrae, R. R. (1992). Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) manual. Psychological Assessment Resources. Dewaele, J.-M. (2022). Personality. In T Gregersen & S. Mercer (Eds.), *The Routledge handbook of the psychology of language learning and teaching* (pp. 112-123). Routledge. Dörnyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited*. Routledge. Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. Heinle & Heinle. Oxford, R. L. (2017). *Teaching and researching language learning strategies. Self-regulation in context*. Routledge. Wakamoto, N. (2000). Language learning strategy and personality variables: Focusing on extroversion and introversion, *IRAL* 38(1), 71-81. Zhou, C., & Intaraprasert, C. (2015). Language learning strategies employed by Chinese English-major pre-service teachers in relation to gender and personality types. *English Language Teaching*, 8(1), 155-169

Optimizing the input for L2 construction acquisition: The role Zipfian input, rules and individual working memory

Manuel Pulido (Penn State).

Usage-based theory has proposed that the acquisition of linguistic constructions is facilitated by Zipfian (or "skewed") input, in which a small set of items are repeated with high frequency (Abbot-Smith & Tomasello, 2006; Bybee & Eddington, 2006). Although early empirical work provided support to this idea (Casenhiser & Goldberg, 2005; Goldberg et al., 2007), subsequent L2 research has provided mixed findings on the role of skewed input (Nakamura, 2012; McDonough & Trofimovich, 2013). However, previous approaches have not explored the impact that cognitive traits (e.g., working memory) have on the effectiveness of input that is skewed (in which a few lexical items have higher frequency) or balanced (in which all lexical items are equally frequent).

The present study tested 82 native English speakers' ability to develop new L2 categories for adjectives, which guide lexical selection in Spanish verbs of "becoming". Participants received training through exposure to the novel verb-adjective combinations. In a 2x2 design, conditions differed in (a) Input Distribution (skewed or balanced) during exposure to adjectives; and in (b) Instruction Type (no rule or explicit rule provided). Learning was measured in an immediate follow up test, and then in a generalization task containing untrained adjectives. Measures of individual working memory span were collected through the Operation-Span task.

The results showed that, when explicit rules were provided, low-working memory learners benefitted from reduced variability in skewed input; in contrast, high-working memory individuals benefitted from balanced input, which better allows for rule-based hypothesis testing. When no rules were provided, learners had lower performance, with no effect of input distribution. The findings help clarify the mixed findings in previous studies, and suggest a way forward for optimizing the L2 input based on individual traits.

Abbot-Smith, K., & Tomasello, M. (2006). Exemplar-learning and schematization in a usage-based account of syntactic acquisition. *Language*, 82, 323-355. Bybee, J., & Eddington, D. (2006). A usage-based approach to Spanish verbs of 'becoming'. *Language*, 82, 323-355. Casenhiser, D. M., & Goldberg, A. E. (2005). Fast mapping between a phrasal form and meaning. *Developmental Science*, 8, 500-508. Goldberg, A. E., Casenhiser, D., & White, T. R. (2007). Constructions as categories of language. *New ideas in psychology*, 25(2), 70-86. Nakamura, D. (2012). Input skewedness, consistency, and order of frequent verbs in frequency-driven second language construction learning: A replication and extension of Casenhiser and Goldberg (2005) to adult second language acquisition. *International Review of Applied Linguistics (IRAL)*, 50, 31-67. McDonough, K., & Trofimovich, P. (2013). Learning a novel pattern through balanced and skewed input. *Bilingualism: Language and Cognition*, 16(3), 654-662.

Cognate vs. Non-cognate Processing in Subtitle Reading: an Eye Tracking Study

Valentina Ragni (University of Warsaw), Breno Silva (University of Warsaw), Agnieszka Szarkowska (University of Warsaw) and Agnieszka Otwinowska-Kasztelanic (University of Warsaw).

The cognate facilitation effect, i.e. faster reaction times to cognates relative to non-cognates, is a well-established phenomenon when bilinguals read isolated L2 words. The effect is smaller for non-identical cognates, and often disappears when reading sentences, especially in high-constraining contexts (for an overview, see Lijewska, 2020). Only few studies have used eye tracking to investigate the processing of cognates either in sentences (e.g., Duyck et al., 2007; Van Assche et al., 2013, Lijewska, 2022) or continuous written input (e.g., Balling, 2013). In audio-visual contexts, only two eye-tracking studies appear to have explored the learning of non-cognates, but not cognates, through subtitle processing (i.e., Montero Perez et al., 2015; Wang and Pellicer-Sánchez, 2022). We have addressed this gap by comparing the processing of cognates and non-cognates when watching subtitled videos in English. We examined whether any differences in subtitle processing occur between cognates and non-cognates using both early eye-tracking measures (e.g., gaze durations, skipping rates) and late eye-tracking measures (e.g., total times, fixation counts). Moreover, we asked whether cognate effects can be found in both monolingual and bilingual viewers.

To this end, we examined data from an experiment where native speakers of English (n=51; female=33; non-binary=4; Mage=21) and Polish advanced learners of English (n=42; Female=34; Mage=23) watched six video excerpts from a documentary with English subtitles at two different speeds: 12 and 20 characters per second. Participants' gaze was monitored with an EyeLink 1000 Plus eye tracker. English-Polish cognates (n=20) and non-cognates (n=20) were selected from the subtitles. All words were controlled (and cognates and non-cognates were matched) for part of speech (noun, adjective), length, frequency of occurrence in SUBTLEX (Brysbaert & New, 2009), number of occurrences (once across the six excerpts), subtitle position (never in first-/last-position in the subtitle, to avoid wrap-up effects), and the presence of a referent in the image. All 40 keywords were assumed known by our bilingual participants. To ensure the keywords were indeed known, they were pre-tested on a group of participants (n=41) with comparable proficiency in English (as measured by the LexTALE; Lemhöfer & Broersma, 2004). Linear mixed-effect models were used to compare cognates and non-cognates in all eye-tracking measures. We assessed whether cognate facilitation occurs in subtitle processing, and whether it is affected by the degree of English-Polish orthographic overlap (Levenshtein distance for cognates), by subtitle speed, and by nativeness (monolinguals vs. English-Polish bilinguals).

In this talk, we will present preliminary results and discuss whether the cognate facilitation effect occurs in subtitle reading. Knowing how cognates are processed in subtitling will help us better understand the specifics of multimodal reading, including whether reading mechanisms differ in native and non-native populations and across different subtitle speeds. Given the ubiquity of subtitling and the prevalence of audio-visual materials in the modern world, understanding these mechanisms in bilinguals seems crucial.

Discrimination of Californian English vowel contrasts by experienced Spanish-Catalan learners

Lucrecia Rallo Fabra (University of the Balearic Islands) and Michael Tyler (Western Sidney University).

L1-Spanish-Catalan speakers often face difficulties when learning the English vowel system (Carlet & Cebrian, 2019; Mora et al., 2022; Rallo Fabra & Romero, 2012). PAM/PAM-L2 (Best, 1994; Best & Tyler, 2007) posits that native language experience shapes perception of L2 sounds, so discrimination of L2 contrasts can be predicted from L2-to-L1 categorization. The current study focused on whether discrimination of the nine Californian English vowel contrasts /i/-ɪ/, /ɛ/-ɪ/, /ɛ/-ə/, /æ/-ʌ/, /æ/-ɑ/, /ɑ/-ʌ/, /ə/-ʌ/ /ə/-ʊ/, and /u/-ʊ/, can be accounted for by the perceptual assimilation of these vowels to L1-Catalan categories. Experienced L1-Spanish-Catalan learners of English performed categorization with goodness rating and AXB discrimination tasks. As shown below, the vowels /i/, /u/, /æ/ and /ɑ/ were categorized, while /ɪ/, /ɛ/, /ʌ/, /ə/, /ɜ/ and /ʊ/, were not consistently mapped onto a single Majorcan Catalan category. Discrimination of the vowel contrasts /i/-ɪ/, /u/-ʊ/, /ɛ/-ə/ and /ʊ/-ə/ was very high (>90%). The contrasts, /ɪ/-ɛ/, /æ/-ɑ/ and /æ/-ʌ/, also were well discriminated (>80%). In contrast, /ʌ/-ɑ/ and /ʌ/-ə/ were discriminated less accurately. The results are discussed in terms of three L2 speech learning models, PAM-L2, SLM-r (Flege, 1995; Flege & Bohn, 2021) and L2LP (van Leussen & Escudero, 2015). In terms of the PAM-L2, the contrast /æ/-ɑ/ was the only instance of a category-goodness (CG) pattern of discrimination. As predicted, discrimination accuracy was moderate-good averaging 0.83. Three contrasts, namely /i/-ɪ/, /æ/-ʌ/ and /u/-ʊ/, were classified as uncategorized-categorized (UC), with different degrees of perceptual overlap, no-overlap for /i/-ɪ/ and partial overlap for /æ/-ʌ/ (72.5%) and /u/-ʊ/ (27.3%). Again, the discrimination scores met PAM-L2's predictions of high accuracy for UC contrasts with no/partial overlap. The low discrimination scores obtained for /ʌ/-ɑ/ and /ʌ/-ə/ are consistent with the complete overlap between the two categories of each contrast, which averaged 82.6% and 53.3, respectively. Finally, discrimination accuracy for the completely overlapping /ɪ/-ɛ/ contrast was higher than would be predicted by PAM-L2. The relationship between perceptual assimilation patterns and discrimination accuracy also meets one of the tenets of the SLM-r and L2LP models, since the two vowels of best-discriminated contrasts /i/-ɪ/, /u/-ʊ/ and /ɛ/-ə/ are quite distant from one another in the acoustic vowel space. In contrast, the two vowels in the poor-discriminated contrasts /ʌ/-ɑ/ and /ʌ/-ə/ overlap one another in the vowel space. Contrary to what has been accepted so far, cross-language similarity can be also assessed acoustically despite the evidence that acoustic measures might not reflect what listeners perceive.

Acquisition of verb morphology in Lithuanian L2: category of person

Jogile Teresa Ramonaite (Institute of the Lithuanian language).

The paper takes the learner variety approach and studies Lithuanian as a second language focusing on the verb morphology. The entrance of a verb into the interlanguage of a learner marks an important step in the acquisition as the utterance can then be anchored around the verb and thus allows a lot more freedom of expression. Even though in the initial stages there is no inflectional morphology, in such morphologically rich language as Lithuanian, there are no bare stems and all forms in the input have some flectional ending. When the learner starts to notice these different endings and to associate them with the different grammatical meaning, the development towards the target language morphology may begin. Lithuanian verbs are inflected for mood, tense, voice, number and person. The focus of this paper is the category of person as acquired in a natural or mostly natural acquisition process. The data analysed was collected using semi-structured (socio)linguistic interviews with 17 speakers of different L1s of various levels of Lithuanian L2 attainment, as well as some specific tasks by the same informants. All of the data (over 27h) was then transcribed and coded for morphology. The form most often used as the base form of the verb by the learners is the 3rd person form ending in -a (also the most frequent one in the target input), however, some other morphological opposition appear rather soon. Drawing on both spontaneous and elicited speech data, the paper presents stages of development of the category of person discussing the relative factors and examining some more problematic instances where some factors (frequency, transparency) play in favour, however, some others (saliency, ambiguity) seem to compete and halt the acquisition of morphology.

References: Andersen, R. W. 1991. Developmental sequences: the emergence of aspect marking in second language acquisition. In T. Huebner & Ch. A. Ferguson (eds.) *Crosscurrents in Second Language Acquisition and Linguistic Theories*, 305-324. Amsterdam/Philadelphia: John Benjamins publishing company. Bernini, G. 1990. Lo sviluppo di paradigmi verbali nelle varietà elementari di apprendimento dell'italiano lingua seconda. In G. Bernini & A. Giacalone Ramat (eds.) *La temporalità nell'acquisizione di lingue seconde*, 81-101. Milano: Angeli. Dressler, W. U., W. Mayerthaler, O. Panagl & W. U. Wurzel. 1987. *Leitmotifs in Natural Morphology*. Amsterdam/Philadelphia: John Benjamins publishing company. Ellis, N.C. 2022. Second language learning of morphology. *Journal of the European Second Language Association*, 6(1), 34-59. Klein, W. & C. Perdue. 1997. *The Basic Variety (or: Couldn't Natural Languages be Much Simpler?)*. *Second Language Research* 13. 301-347. MacWhinney, B. 2000. *The CHILDES Project. Tools for Analyzing Talk*. Third Edition. Mahwah, NJ: Lawrence Erlbaum Ass. Perdue, C. & W. Klein. 1992. Why does the production of some learners not grammaticalize? *Studies in Second Language Acquisition* 14. 259-272.

'Believing and Achieving' in learning Arabic: the relationship between 'motivational selves' and proficiency in a multiglossic context

Anna-Maria Ramezanzadeh (University of Oxford), Robert Woore (University of Oxford) and Lars-Erik Malmberg (University of Oxford).

The learning of Arabic as an L2 has seen strong global growth in recent years (Lo, 2019). Yet, it is seen as challenging to learn, described by US Foreign Service Institute as 'exceptionally difficult for native English speakers' (<https://exchanges.state.gov/cls>). Contributing to this difficulty is the language's multiglossic nature, consisting of standard, classical, and colloquial varieties, each with their own realms of application.

Despite the central importance of motivation for language learning, scant research has investigated motivation for learning Arabic. This is theoretically and pedagogically interesting, since the language's multiglossic nature may affect learners' motivation: the variety being taught may not always align with their aspirations for language use. Addressing this gap in the literature, our study examines the motivation of secondary school students in England studying for national examinations at age 16 (the General Certificate of Secondary Education or GCSE) in Arabic, taking Dörnyei's (2009) Second Language Motivational Self System (L2MSS) as its theoretical lens.

Research conducted within this framework has mainly focused on the linear relationship between imagined future selves and learners' outcomes. However, recent studies into self-efficacy suggest that the impact of learners' self-beliefs may not be linear but reciprocal (Talsma et al., 2018). Rather than a 'believe-achieve' relationship, where self concepts lead to subsequent achievement, these two constructs were found to influence each other over time. Thus, our study used Structural Equation Modelling to investigate the strength and direction of influence between (a) motivational self concepts within the L2MSS and (b) language proficiency. It also examined whether reciprocal, rather than linear relationships are present amongst these variables.

Data was collected from 265 GCSE Arabic students at two time-points. Motivational self concepts were measured in relation to the three Arabic varieties: Modern Standard; Classical; and Dialect. A new instrument measuring learners' Current, Ideal, and Ought-to selves was designed, along with language proficiency tests. Semi-structured interviews were conducted with 95 participants on their motivational self concepts and language proficiency.

Findings revealed significant regression paths between performance and self-concept, though not reciprocal relationships. This corroborates the findings of Talsma and colleagues (2018), who reported similar results amongst younger learners, and challenges existing L2MSS research, which emphasizes the role of the ideal self in promoting learning outcomes. Interview data – analysed via a sequential coding approach – shed light on the mechanisms underlying these findings. We discuss the implications of our findings for policy, pedagogy, research and theory in relation to both Arabic learning and teaching and the study of motivation in SLA.

Dörnyei, Z. (2009). The L2 Motivational Self System. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self*. *Multilingual Matters*. 9-42.

Lo, M. (2019). *The Arabic Classroom: Context, Text and Learners*. Routledge

Talsma, K., Schüz, B., Schwarzer, R., & Norris, K. (2018). I believe, therefore I achieve (and vice versa): a meta-analytic cross-lagged panel analysis of self-efficacy and academic performance. *Learning and Individual Differences*, 61(January), 136–150.

Phonaesthetics and L2: Does enjoying the sound of foreign languages support L2 learning?

Susanne Reiterer (University of Vienna), Vita Kogan (University College London) and Anna Winkler (University of Vienna).

With the spread of easy-to-access language-learning apps and online platforms, recreational L2 learning – learning a foreign language/L2 as a hobby – is getting momentum. Often time the choice of L2 is based on the phonesthetic properties or how pleasant a particular language sounds to a prospective learner. Phonesthetics as a source of learner motivation has been never explored before, even though it might have an emotive and pedagogical potential to support language learning even in the context where L2 choice is assigned by the institution or circumstances (e.g., the use of synesthetic activities that emphasize specific L2 phonetic features in Wrembel (2010)).

In a series of studies, we asked participants (n=145) to rate a number of European languages in terms of eroticism, beauty, well-structuredness, and status/prestige and also asked about their familiarity with the languages, L1 and L2 backgrounds, demographics, mobility, musical/singing experience, personality traits, and how much they liked the speakers' voices (Kogan & Reiterer, 2021; Reiterer et al., 2020; Winkler et al., in press). Overall, polyglots rated the languages of the experiment significantly higher perceiving them as more erotic and beautiful in comparison to their monolingual peers. This finding suggests that polyglots might experience the so-called “phonetic chill” – the aesthetic pleasure one derives from listening to the sounds of a foreign language, that might inspire and stimulate their learning. The qualitative data – participants' free-style comments – support this assumption. That being said, the picture is more complex with both the inherent phonological features of the languages (e.g., sonority, syllabic structure, intonation, and rhythm) and the sociocultural stereotypes about the languages (e.g., French sounds beautiful because of the associations of romance and hedonism) driving the phonaesthetic judgments. In Study 3, we also observed the significant effect of personality traits on phonaesthetic preferences with introverts rating languages higher for eroticism and beauty and neurotics lower across the board. It is to be determined if the phonetic chill is a prerequisite or a result of extensive language learning and/or exposure to foreign languages (mobility). That being said, using phonaesthetics as a pedagogical device to complement the L2 classroom might be an efficient way to increase learners' motivation and support the acquisition of L2 phonetics and phonology.

References

Kogan, V. V., & Reiterer, S. M. (2021). Eros, beauty, and phonaesthetic judgements of language sound. We like it flat and fast, but not melodious. Comparing phonetic and acoustic features of 16 European Languages. *Frontiers in human neuroscience*, 15, 578594.

Reiterer, S. M., Kogan, V., Seither-Preisler, A., & Pesek, G. (2020). Foreign language learning motivation: Phonetic chill or Latin Lover effect? Does sound structure or social stereotyping drive FLL?, *Psychology of learning and motivation*, 72, 165-205.

Winkler, A., Kogan, V., & Reiterer, S.M. (in press). Phonaesthetics and Personality – Why we do (not only) especially prefer Romance languages. *Frontiers in Language Sciences*.

Wrembel, M. (2010). Sound symbolism in foreign language phonological acquisition. *Research in Language*, 8, 1–14.

Teacher Judgements as a Predictor of Students' Vocabulary Knowledge

Pablo Robles-García (University of Toronto, Toronto, Canada), Christopher Nicklin (Rikkyo University, Tokyo, Japan), Jeffrey Stewart (Tokyo University of Science, Tokyo, Japan), Joseph P. Vitta (Kyushu University, Fukuoka, Japan) and Stuart McLean (Momoyama Gakuin University, Osaka, Japan).

When developing word lists for pedagogical purposes, word-frequency is routinely implemented to identify which words L2 learners are more or less likely to know (Hashimoto, 2021). However, some researchers suggest that teachers often prefer to rely on their own intuition rather than using corpus-based vocabulary lists for making decisions about the words they want to teach in the classroom (Dang & Webb, 2020; Sánchez-Gutiérrez et al., 2022). Although teacher judgments are a commonly used strategy for vocabulary selection in the L2 classroom, little is known about the accuracy of such judgments when predicting L2 learners' vocabulary knowledge. This study investigated the effectiveness of word-frequency and teacher judgments in determining students' vocabulary knowledge and compared the predictive powers of both approaches when estimating word difficulty. Twenty-nine L2 Spanish teachers were asked to predict how likely their students would know words from 3K-LEx (Robles-García, 2020), a 216-word Yes/No test that measures knowledge of the most frequent 3,000 words in Spanish. The accuracy of their responses was compared with the 3K-LEx results of 1,075 L2 Spanish learners in North America. To examine if the results were generalizable to other L2 learning contexts, 394 L2 English students in Japan completed a 70-word Yes/No test measuring knowledge of the first 14,000 words in English, and 15 Japan-based English language instructors predicted the likelihood of learners knowing the words from that test. Results revealed that for both language contexts the median teacher rater could assess difficulty with an accuracy roughly comparable to frequency. Additionally, bootstrapping was employed to demonstrate that the aggregation of as few as three sets of teacher judgements was sufficient to improve upon frequency as a predictor of vocabulary knowledge, with no improvement displayed by the aggregated judgements of more than approximately 10 teachers. Finally, a multiple regression model combining frequency and teacher judgements did not substantially improve the prediction of word difficulty compared to models with teacher judgments as the lone predictor. These findings suggest that aggregating teacher judgements could provide a simple way for language program managers and teachers to develop word lists that are more suitable for their learners than typical, frequency-based lists.

References

Dang, T. N. Y., & Webb, S. (2020). Vocabulary and good language teachers. In C. Griffiths & Z. Tajeddin (Eds.), *Lessons from good language teachers* (pp. 203–218). Cambridge University Press.

Hashimoto, B. (2021). Is frequency enough? The frequency model in vocabulary size testing. *Language Assessment Quarterly*, 18(2), 171–187. <https://doi.org/10.1080/15434303.2020.1860058>

Robles-García, P. (2020). 3K-LEx. Desarrollo y validación de una prueba de amplitud léxica en español [Development and validation of a lexical breadth test in Spanish]. *Journal of Spanish Language Teaching*, 7(1), 64–76. <https://doi.org/10.1080/23247797.2020.1770465>

Sánchez-Gutiérrez, C., Robles-García, P., & Pérez Serrano, M. (2022). L2 Spanish vocabulary in US universities: Instructors' beliefs and reported practices. *Language Teaching Research*, 1–23. <https://doi.org/10.1177/13621688221074443>

Are phraseological units processed holistically? An eye-tracking study on L1 and L2 speakers of Italian.

Maria Roccaforte (University "La Sapienza" of Rome), Veronica D'Alesio (University "La Sapienza" of Rome), Irene Fioravanti (University for Foreigners of Perugia), Luciana Forti (University for Foreigners of Perugia) and Stefania Spina (University for Foreigners of Perugia).

In the last decades, there has been a growing interest in the processing L1 and L2 phraseology. It has been argued that phraseological units are holistically stored and retrieved directly by L1 speakers, while learners tend to analyse them in a more compositional way (Wray, 2002; Van Lancker Sidtis, 2012). However, faster processing of phraseological units nowadays is not believed to be completely attributed to the holistic storage of phraseological sequences (Siyanova-Chanturia, 2015). Indeed, hybrid models of comprehension of phraseological units (Titone & Libben, 2014) claim that both direct retrieval of phraseological sequences and compositional processing of word constituents take place during on-line processing with phraseological units' processing advantage depending on different properties with an access to their elements (Carrol & Conklin, 2020). To investigate the issue of holistic or hybrid processing in L1 and L2 phraseology, an eye-tracking study was carried out with L1 speakers and learners (intermediate and advanced) of Italian. We focused on a specific type of Verb+Noun phraseological sequence, collocations. We sought to answer two research questions: 1) How do L1 and L2 speakers of Italian process typical and atypical Verb+Noun collocations? 2) Does proficiency influence the processing of typical and atypical collocations in L2 speakers of Italian? Collocations were derived from a reference corpus of Italian and manipulated lexically by substituting the verb of the collocation with a synonym (e.g., *passare + esame*, 'to pass + exam' vs. **attraversare + esame*, 'to pass through + exam'), and grammatically by inserting an agreement error in the article (e.g., *aprire gli occhi*, 'to open the eyes' vs. *aprire *lo occhi*, 'to open the[singular] eyes[plural]) and in the noun (e.g., *vivere una esperienza*, 'to live an experience' vs. *vivere una *esperienze*, 'to live an experiences'). Typical and atypical collocations were embedded in context sentences and participants read the sentences for comprehension while their eye-movements were recorded. Mixed-effect modeling showed that L1 processed typical collocation faster than L2 speakers. No difference was found between intermediate and advanced learners. Further, L1 speakers and advanced learners processed atypical collocations differently, fixating lexically manipulated collocations longer than grammatically manipulated collocations. On the contrary, no clear distinction between lexical and grammatical manipulation processing was found in intermediate learners. Further, learners' proficiency resulted to not modulate any effect on the processing of typical and atypical collocations. However, as the proficiency increased advanced learners tended to process typical and atypical collocations in a similar way to L1 speakers. Our results show that there is an access by both L1 and L2 speakers to the internal structure of collocations, suggesting that the direct retrieval takes place in processing typical collocations, while the processing cost induced by atypical phrases is resolved by a word-by-word analysis.

References Carrol, G., & Conklin, K. (2020). Is all formulaic language created equal? Unpacking the processing advantage for different types of formulaic sequences. *Language and Speech*, 63(1), 95-122.

Siyanova-Chanturia, A. (2015). On the 'holistic' nature of formulaic language. *Corpus Linguistics and Linguistic Theory*, 11(2), 285-301.

Titone, D., & Libben, M. (2014). Time-dependent effects of decomposability, familiarity and literal plausibility on idiom priming: A cross-modal priming investigation. *The Mental Lexicon*, 9(3), 473-496.

Van Lancker Sidtis, D. (2012). Two-Track Mind: Formulaic and Novel Language Support a Dual-Process Model. In M. Faust (Ed.), *The Handbook of the Neuropsychology of Language* (pp. 342-367). Wiley-Blackwell.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press.

Order of exposure matters: a case study of the article system in English

Laurence Romain (University of Birmingham), Petar Milin (University of Birmingham) and Dagmar Divjak (University of Birmingham).

Since research within learning theory has found that order of exposure affects learning (e.g., Kamin 1969), we consequently argue that the order in which relevant information is learned/taught ought to be crucial for SLA (cf. Ellis 2006). Using the English article system as our case study, we ran a pre-registered, online intervention study testing the effect of order of exposure/teaching with B2/C1 level learners of an article-less L1 (Polish).

Using corpus data, we first identified which elements in the context are the most informative for the choice of article. Then, we ran computational simulations of learning using an error-driven associative algorithm (Widrow and Hoff 1960) on a dataset of 2000 discourse chunks extracted from the BNC (1000 spoken, 1000 written) that we manually annotated with the five most commonly used variables, i.e., Hearer Knowledge (HK), Specificity of the referent (SR), countability, number and elaboration (Huebner 1985). In the first simulation, we fed the algorithm our data in random order of exposure. We found that different articles are learned from different cues but that HK leads in the acquisition of *a* and *the*. We then tested order of exposure by designing two ordered datasets in which the cues for the variable HK or the cues for SR are introduced only halfway through training. The results from the simulations suggest that HK is crucial in determining article choice and when introduced after SR it does not get associated sufficiently and, hence, utilised fully: the acquisition of the article system is hindered by less potent cues.

Accordingly, we divided our 72 learners into two training groups: one group received the HK-first treatment, i.e., they were trained using the HK variable before we introduced the SR variable, while the other group received SR-first treatment. We predicted that the HK-first group would improve their performance significantly more than the SR-first group and the knowledge the HK-first group built up would generalise better, enabling them to tackle more challenging contexts. Preliminary analyses of the pre-, post- and delayed post-tests results show that both groups performed extremely well at pre-test already (scoring an average of 46/50), leaving little scope for improvement post-test. However, in addition to the items included on the pre-test, the post-test included “hard” items, on which the HK-first learners performed significantly better ($p = 0.0083$) than their SR-first counterparts.

Our findings clearly demonstrate that relevant information in the right order is crucial in the context of SLA too. Even more striking is that this principle robustly shows its effect in advanced learners whose performance is at ceiling.

Selected references

- Ellis, Nick C. (2006). Language acquisition as rational contingency learning. *Applied Linguistics*, 27(1), 1-24.
- Huebner, Thom. (1985). System and variability in interlanguage syntax. *Language Learning*, 35(2), 141-163.
- Kamin, Leon J. . (1969). Predictability, surprise, attention and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and Aversive Behaviour*. New York: Appleton-Century-Crofts.
- Widrow, B., & Hoff, M. E. . (1960). Adaptive switching circuits. Paper presented at the WESCON Convention Record Part IV.

Task and Time Effects on Revision Behaviours in Computer-Mediated Synchronous Collaborative Writing

Xin Rong (University College London).

Collaborative writing, defined as an activity that requires writers to co-construct a text, has gained increasing attention over the past decades as a tool to promote writing development. To date, however, researchers have exclusively studied collaborative writing processes taking a sociocultural perspective. Little is known about the writing behaviours and underlying cognitive activities in which learners engage when they perform collaborative writing tasks. This study, inspired by Kellogg's (1996) and Rijlaarsdam and Van den Bergh's (1996) work, aimed to mitigate this gap. Specifically, it intended to examine (a) the extent to which task complexity affects learners' revision behaviours and associated cognitive activities when they write collaboratively, (b) the temporal distribution of revision behaviours and related cognitive processes during collaborative writing, and (c) the interaction between task complexity and time distribution of revision behaviours.

The participants were 56 advanced Chinese learners of English, who were randomly assigned to 28 pairs. Each pair completed two collaborative writing tasks in a counterbalanced order. The two tasks took the form of reading-to-write tasks and differed in cognitive task complexity. The less complex version required participants to summarise one text. The more complex version, on the other hand, asked them to synthesise three texts, which was assumed to pose greater cognitive demands as it required participants to integrate information and build new organisational patterns. Text length and linguistic complexity was controlled across the task versions. When performing the tasks, participants used separate laptops sitting face-to-face to each other. Their keystrokes were logged to capture revision behaviours. Immediately after the second task performance, 16 participants from eight pairs engaged in one-to-one stimulated recall interviews. Revision behaviours were coded in terms of location (e.g., below word) and the related stimulated recall comments were analysed for reasons for revisions (e.g., content planning). The resulting frequency indices were obtained for the whole writing process and five equally timed intervals to assess the temporal distribution of revision behaviours.

Preliminary results, based on a sample of 12 participants and six recall sessions, found the following patterns. Participants more frequently revised smaller than larger textual units, and focused more on content and lexis than organisation or syntax. Increased task complexity elicited more frequent below-word and below-clause revisions, while full words were more often revised in the less complex task. In terms of the temporal changes, frequency of revision showed a rising trend overall. The specific patterns, however, revealed that, as time passed, participants made more frequent revisions to larger textual units but revised smaller units to a lesser degree. In addition, they paid increasingly more attention to language use during the writing process. An interaction effect also emerged: participants revised more frequently in the first three intervals than the final two intervals during the more complex task, as compared to the less complex task.

These results will be discussed in relation to trends observed for individual writing in previous research. The implications of this study will also be discussed for theoretical models of writing and pedagogical practice.

Auditory Processing as Perceptual, Cognitive, and Motoric Abilities Underlying Successful Second Language Acquisition

Kazuya Saito (University College London), Magdalena Kachlicka (Birkbeck, University of London), Yui Suzukida (Juntendo University), Ingrid Mora-Plaza (University of Barcelona), Yaoyao Ruan (Oxford University) and Adam Tierney (Birkbeck, University of London).

A growing amount of attention has been given to examining the domain-general auditory processing of individual acoustic dimensions (e.g., frequency, duration) as a key driving force for adult L2 acquisition (Kachlicka et al., 2019). To date, auditory processing has traditionally been assessed via a psychoacoustic discrimination task in which researchers examine how small a difference participants can perceive in the spectral and temporal content of sounds. Comparatively, the interaction view (Kraus & Banai, 2007) has posited that auditory processing can be conceptualized as a multifaceted phenomenon wherein perceptual, cognitive, and motoric abilities are reciprocally interwoven at multiple levels. Specifically, auditory processing covers not only the perception of acoustic details (perceptual acuity) but also the direction of attention to particular acoustic dimensions (attentional control) and the conversion of audio information into motor action (audio-motor integration).

To test this hypothesis, a total of 102 Chinese learners of English with a wide range of L2 experience and proficiency levels were recruited in the UK. First, they took a battery of auditory processing tests which were assumed to tap into acuity (discrimination of pitch and duration differences; Kachlicka et al., 2019), attention (detection of pitch and duration repetitions; Symons et al., 2021), and integration (reproduction of melody and rhythm; Saito et al., 2021). Then, they completed both vowel and prosody identification and grammaticality judgement tasks to index their L2 phonological and morphosyntactic proficiency. Finally, we surveyed their biographical backgrounds (e.g., age, length of residence, daily L1 and L2 use) and working memory (via forward and backward digit span).

The results of factor analyses showed that the tests scores tapped into essentially distinct components of auditory processing (acuity, attention, and integration). Further, the results of logit mixed-effects analyses demonstrated that the perceptual, cognitive, and motoric model of auditory processing demonstrated medium-to-large predictive power for L2 phonology proficiency ($R^2 = .447$) and L2 morphosyntax proficiency ($R^2 = .381$) and remained significant even after all the relevant biographical background and working memory variables were controlled for. The patterns reported here not only concur with the extant literature which has noted a small-to-medium link between acuity and L2 phonology (e.g., $R^2 = .15$ to $.25$ in Kachlicka et al., 2019) and L2 morphosyntax ($R^2 = .05$ to $.15$ in Saito et al., et al., 2021) but also lend empirical support to our hypothesis that the inclusion of neighboring abilities (attention and integration) can explain additional variance in language learning outcomes. Our argument here concurs with ongoing claims that auditory processing, comprised of perceptual, cognitive, and motoric components, serves as a bottleneck for language learning (Kraus & Banai, 2007).

Kachlicka et al. (2019). Successful L2 learning is tied to robust domain-general auditory processing. *Brain and Language*, 192, 15-24. Kraus, N., & Banai, K. (2007). Auditory-processing malleability. *Current Directions in Psychological Science*, 16(2), 105-110. Saito et al. (2021). Domain-general auditory processing partially explains L2 learning in classroom settings: *Language Learning*, 71, 669-715. Symons et al. (2021). Dimension-selective attention and dimensional salience modulate cortical tracking of acoustic dimensions. *NeuroImage*, 244, 118544.

Different weightings of the agent-first strategy and a semantic cue in L1 and L2 sentence processing

Sarah Schimke (LMU Munich), Gregory Poarch (University of Groningen), Freya Gastmann (TU Dortmund University), David Öwerdieck (TU Braunschweig), Damian Stier (TU Dortmund University) and Holger Hopp (TU Braunschweig).

Across languages and types of language users, subject-initial sentences have been shown to be easier to comprehend than object-initial sentences, suggesting that language users rely on a general agent-first processing heuristic, according to which the first-mentioned NP is assigned the role of subject and agent by default (Slobin & Bever, 1982). L2 learners have been found to privilege processing heuristics over deeper morphosyntactic processing to a larger degree than native speakers (Clahsen & Felser, 2006). It is less clear, however, how semantic cues are weighted in L1 and L2 processing. On the one hand, some studies have found that learners (over-)rely on such cues, e.g., animacy (Jackson & Roberts, 2010). On the other hand, the rapid integration of cues from different linguistic levels has sometimes been found to be more difficult in L2 than in L1 processing (e.g., Hopp, 2018). The current study investigates whether adolescent language users ($n = 114$, age range: 12-14 years) make use of animacy to restrain or even override the agent-first strategy in their L1, German, and their L2, English. In a visual-world eye-tracking task, subject-initial (1,3) and object-initial (2,4) WH-questions were auditorily presented. The first NP was either animate (1,2), supporting the agent-first strategy, or inanimate (3,4), facilitating an object-initial interpretation. We assessed fixations to target pictures (Figure 1), decision latencies, and accuracies of sentence-final interpretations.

1. Welches Tier umarmt denACC Pinguin? (subject-first; animate) Which animal hugs the penguin? 2. Welches Tier umarmt derNOM Pinguin? (object-first; animate) Which animal does the penguin hug? 3. Was umarmt denACC Pinguin? (subject-first; inanimate) What hugs the penguin? 4. Was umarmt derNOM Pinguin? (object-first; inanimate) What does the penguin hug?

Mixed-effects modelling returned interactions of Language, Word Order, and Animacy in all measures (Figure 2). In L1 German, accuracy was at ceiling in all conditions. Processing measures revealed interactions between animacy and word order: for animate first NPs, target fixations were earlier and decisions were faster for subject- than for object-initial sentences; for inanimate NPs, the reverse held. In L2 English, eye movements and decision latencies revealed a strong processing advantage of subject-initial over object-initial sentences, irrespective of animacy. Only in sentence-final interpretation did animacy interact with word order. The timing differences in the use of semantic information point to a different weight of the agent-first strategy and of semantic cues in L1 and L2 processing. Participants were able to incrementally integrate semantic cues in their L1, but not in their L2, where they only used them in sentence-final interpretation. We will discuss theoretical implications of these findings.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, 27(1), 3–42.

Hopp, H. (2018). The bilingual mental lexicon in L2 sentence processing. *Second Language*, 17, 5–27.

Jackson, C.N., & Roberts, L. (2010). Animacy affects the processing of subject-object ambiguities in the L2. *Applied Psycholinguistics*, 31, 671–691.

Slobin, D.I., & Bever, T.G. (1982). Children use canonical sentence schemas: A crosslinguistic study of word order and inflections. *Cognition*, 12(3), 229–265.

The foreign language effect and vaccine hesitancy: a COVID-19 study

Monika Schmid (University of York) and Karen Roehr-Brackin (University of Essex).

There is a substantial body of research demonstrating that the recruitment of rational processes is increased in hypothetical decision-making scenarios when the underpinning information is presented in a foreign language (e.g. Costa et al., 2014). Few studies have attempted to transpose this effect to real-life situations, but it has been found to apply to some extent in medical scenarios (e.g. Geipel, Grant & Keysar, 2022; Hayakawa, Pan & Marian, 2022). This study aims to investigate whether the Foreign Language Effect could be beneficial in the context of vaccine hesitancy – a phenomenon that was counted among the 10 greatest challenges to worldwide health in 2019 by the World Health Organisation (<https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019>, accessed 20/01/2022) and which was subsequently thrown into sharp relief by the Covid-19 pandemic. This means that ways have to be found to convince vaccine-hesitant individuals to trust scientific evidence relating to the risks and benefits of vaccines in the face of irrational and counterfactual beliefs amplified by disinformation campaigns. The present study thus investigates whether vaccine campaigns could benefit from the Foreign Language Effect (FLE) in order to overcome vaccine hesitancy. We conceptually replicated a recent study on Covid-19 vaccine campaigns (Freeman et al., 2021) in order to assess whether information can more successfully overcome vaccine hesitancy when presented in a foreign language as well as how other factors such as language proficiency impact on the FLE. 436 participants who had thus far declined a Covid-19 vaccine, who were native speakers of either English, French or German and second language speakers of another of these three languages were randomly presented with information in either their L1 or their L2, as well as with a short proficiency assessment. Their attitude towards the Covid-19 vaccine was assessed before and after the intervention, following the methodology by Freeman et al. We conclude that there may be the potential in particular for moderately vaccine-hesitant individuals with English as a foreign language to respond more positively to information presented in this language than in their native language. Conversely, participants who are native speakers of English and negatively predisposed against the Covid-19 vaccine react more negatively to information they read in a foreign language compared to their native language.

References Costa, A., Foucart, A., Arnon, I., Aparici, M., & Apesteguia, J. (2014). "Piensa" twice: On the foreign language effect in decision making. *Cognition*, 130(2), 236-254. Freeman, D., Loe, B. S., Yu, L. M., Freeman, J., Chadwick, A., Vaccari, C., ... & Lambe, S. (2021). Effects of different types of written vaccination information on COVID-19 vaccine hesitancy in the UK (OCEANS-III): a single-blind, parallel-group, randomised controlled trial. *The Lancet Public Health*, 6(6), e416-e427. Geipel, J., Grant, L. H., & Keysar, B. (2022). Use of a language intervention to reduce vaccine hesitancy. *Scientific reports*, 12(1), 1-6. Hayakawa, S., Pan, Y., & Marian, V. (2022). Language changes medical judgments and beliefs. *International Journal of Bilingualism*, 26(1), 104-121.

Encoding and decoding adjectival agreement in instructed Italian L3 learning

Katrin Schmiderer (Universität Innsbruck) and Barbara Hinger (University of Graz).

In recent years, the majority of psycholinguistic and neurolinguistic research assumes that language comprehension and production rely on the same processing mechanisms (see Kempen et al., 2012). Lenzing (2021) proposed an “Integrated Encoding-Decoding Model” with a single syntactic processor that is accessed by both L2 grammatical encoding and decoding and develops in concordance with the predictions of acquisitional stages postulated in PT (Pienemann 1998). However, empirical studies investigating language comprehension and production processes from a PT perspective (see Buyl & Housen 2013; Buyl 2019) have yielded divergent results.

The present paper seeks to investigate whether the productive and receptive acquisition of attributive and predicative adjective agreement in Italian as an L3 moves along similar developmental trajectories and is governed by the same processing procedures. The participants of the study (n=29) were in their second year of instruction in an upper secondary school in Austria; the control group consisted of 18 Italian L1 speakers. In order to elicit the mentioned structures, four communicative tasks providing different contexts for spontaneous oral speech production were used; comprehension data was collected from an oral sentence-matching task (Verhagen 2011), in which the participants’ reaction times and error rates when matching grammatical and ungrammatical sentences were recorded. Based on the learners’ developmental sequences identified in the production data learners were grouped for the statistical analysis of receptive data.

In the paper we discuss the research findings which seem to confirm that learners can produce attributive adjective agreement before predicative agreement as predicted in the PT-hierarchy (Di Biase/Bettoni 2015). Furthermore, we discuss the presence of grammaticality effects in reaction times in learners who have acquired attributive and predicative adjective agreement in production, and its absence for learners who have not or only partially acquired adjective agreement. Finally, the results are compared to L1 speakers.

Buyl, A. (2019): Is morpho-syntactic decoding governed by Processability Theory? In: Lenzing, A./Howard, N./Roos, J. (eds.): *Widening Contexts for Processability Theory. Theories and issues*. Amsterdam: John Benjamins, 73-104.

Buyl, A./Housen, A. (2013): Testing the applicability of PT to receptive grammar knowledge in early immersion education. Theoretical considerations, methodological challenges and some empirical results. In: Flyman Mattsson, A. /Norrby, C. (eds.): *Language acquisition and use in multilingual contexts. Theory and Practice*. Lund: Travaux de l’Institut de linguistique de Lund, 13-27.

Di Biase, B./Bettoni, C. (2015): The development of Italian as a second language. In: Bettoni, C./Di Biase, B. (eds.) *Grammatical Development in Second Languages. Exploring the boundaries of Processability Theory*, EUROSLA Monograph Series (3), 117-147.

Kempen, G./Olsthoorn, N./Sprenger, S. (2012): Grammatical workspace sharing during language production and language comprehension: Evidence from grammatical multitasking. *Language and Cognitive Processes* 27 (3), 345-380.

Lenzing, A. (2021): *The Production-Comprehension Interface in Second Language Acquisition. An Integrated Encoding-Decoding Model*. London: Bloomsbury.

Pienemann, M. (1998): *Language Processing and Second Language Development: Processability Theory*. Amsterdam: John Benjamins.

Verhagen, J. (2011). Verb placement in second language acquisition: Experimental evidence for the different behavior of auxiliary and lexical verbs. *Applied Psycholinguistics*, 32(4), 821-858.

Exploring different implementations of repeated reading for incidental vocabulary learning

Raquel Serrano (Universitat de Barcelona) and Ana Pellicer-Sánchez (Institute of Education, University College London).

There is evidence that in order to promote incidental vocabulary learning from reading, multiple exposures to the same target words are needed (Webb, 2007). There is not much evidence, however, as to whether novel words are learned more successfully if they appear in the same text read multiple times or in different texts. Although repetitions in the same context might lead to faster recognition of novel vocabulary, reading the same text multiple times could also cause boredom and lack of attention to new words (Liu & Todd, 2016). Additionally, encountering words in varied contexts has been claimed to support the establishment of richer semantic associations (e.g., Ferreira & Ellis, 2016). Besides the role of context, the spacing between repetitions of novel words has also been shown to affect their learning and retention when learned in isolation, in sentences, or when they appear in longer texts (Edmonds et al., 2021), although it is still unclear what the optimal spacing is to promote incidental vocabulary learning from reading. The aim of the present study was to investigate the effect of context (same vs. different), as well as spacing (massed vs. spaced) on the incidental learning of vocabulary through repeated reading. One hundred EFL college students from Spain were exposed to two sets of 10 novel words. One set appeared in the same text read three times and another one in three different texts in a counterbalanced fashion. The four experimental texts were around 950 words long each, and each included two instances of the target words. Approximately half of the participants read the texts in a massed format (repeated readings in the same session), while the other half did so under a spaced schedule (one reading per week). Participants' knowledge of the target vocabulary was assessed immediately after the last reading and one week later through a form-recognition, a meaning recall, and a meaning recognition test. Additionally, measures of general vocabulary and proficiency were also obtained. The results of the GLMMs show differences between conditions, with massed and same creating less difficulty for immediate learning than spaced and different, but easier conditions did not always lead to better long-term retention. The results will be interpreted following the Desirable Difficulties Framework (Suzuki et al., 2019).

References

Edmonds, A., Gerbier, E., Palasis, K., & Whyte, S. (2021). Understanding the distributed practice effect and its relevance for the teaching and learning of L2 vocabulary. *Lexis*, 18.

Ferreira, R. A., & Ellis, A. W. (2016). Effects of contextual diversity on semantic decision and reading aloud: Evidence from a word learning study in English as a second language. *Studies in Psychology*, 37, 162–182

Liu, Y.-T., & Todd, A. G. (2016). Implementation of assisted repeated reading techniques for the incidental acquisition of novel foreign vocabulary. *Language Teaching Research*, 20(1), 53–74.

Suzuki, Y., Nakata, T., & DeKeyser, R. (2019). The desirable difficulty framework as a theoretical foundation for optimizing and researching second language practice. *Modern Language Journal*, 103(3), 713–720.

Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, 28(1), 46-65.

Learner perceptions of difficult L2 phonological contrasts and orthographic influence across writing systems

Louise Shepperd, Sam Hellmuth and Leah Roberts

Research investigating the relationship between orthography and second language (L2) phonology has grown considerably over the past 20 years in both size and complexity (Hayes-Harb & Barrios, 2021). This is unsurprising given the ubiquity of written language as a foundation for adult language acquisition. However, only a few studies have explored the influence of L2 written input across writing systems (Han & Kim, 2017; Showalter & Hayes-Harb, 2013), and participants are rarely invited to reflect on how orthography influences their phonological acquisition (Bassetti et al., 2020). The present study investigated the influence of written input on adult Arabic-speakers' accuracy in lexical-encoding of a difficult L2 phonological contrast, as well as learners' perceptions of the influence and importance of written input. / / Adult Arabic-speaking learners of English (n=114), with varying English proficiency levels, completed an online word learning study. This consisted of learning 12 English pseudowords presented auditorily, accompanied by an image and written input. Target words were minimal pairs beginning with /f-v/ or /m-n/, where /f-v/ is a difficult L2 contrast for L1 Arabic-speakers and /m-n/ is a well-established contrast in both Arabic and English. Target words were presented with either Arabic spelling, English spelling or no written input. The visual distinction in English spelling was predicted to promote lexical-encoding of the contrast, whereas the shared <f> grapheme for both /f/ and /v/ in Arabic was expected to exacerbate perceptual difficulty. This was tested in a YES-NO audio-picture matching task, where half the trials matched and the other half mismatched target words with the images of their minimal pair item. In a post-test questionnaire, participants reflected on the influence of written input in the study and broader beliefs about the importance of written input when learning new words. / / Analysis of L1 Arabic accuracy data with mixed-effects models revealed a significant interaction between phonological contrast and written input: any written input, English or Arabic, had an inhibitory effect on encoding the distinction between /f-v/ words. This was at odds with participants' overwhelmingly positive perceptions about the influence of written input during the study, particularly English spelling. Qualitative Content Analysis of questionnaire open responses revealed diverse perceptions of the support and/or hindrances associated with exposure to written input in different scripts, as well as the relevance of previous (language) learning experiences. Overall, these findings highlight the need to consider learner beliefs about literacy and language learning when investigating phonological acquisition with multimodal input.

Bassetti, B., Mairano, P., Masterson, J., & Cerni, T. (2020). Effects of Orthographic Forms on Second Language Speech Production and Phonological Awareness, With Consideration of Speaker Level Predictors. *Language Learning*, 43, 155.

Han, J.-I., & Kim, J.-Y. (2017). The Influence of Orthography on the Production of Alphabetic, Second-Language Allophones by Speakers of a Non-alphabetic Language. *Journal of Psycholinguistic Research*, 46, 963-982.

Hayes-Harb, R., & Barrios, S. (2021). The influence of orthography in second language phonological acquisition. *Language Teaching*, 1-30.

Showalter, C. E., & Hayes-Harb, R. (2013). Unfamiliar orthographic information and second language word learning: A novel lexicon study. *Second Language Research*, 29(2), 185-200.

The role of L1 typology and L2 proficiency in predicting null-subject transfer

Jye Smallwood (University of Cambridge).

This paper is based on my distinction-awarded dissertation as part of my MSc at the University of Oxford. The project is a large-scale investigation on the omission of subject pronouns by learners of null-subject languages in an effort to explore transfer effects by language typology. A null-subject language is one where pronouns are either optional or unnecessary to make reference to a primary subject. In these languages, (which includes, Spanish, Italian, Arabic, Mandarin Chinese, Russian, Swahili, and many others) dropping a pronoun does not result in ambiguity. However, English is a non-null-subject language and a subject pronoun is always required in referring expressions. The aim of the present study is to determine if this difference in subject parameters means that learners are more prone to omit pronouns in second language (L2) English production. The first research question of the present study asks, 'do null subjects in the L1 transfer to L2 English?' This question becomes more complex when we consider that there are more than one type of null-subject language. How null subjects are licensed is in 'canonical' null-subject languages like Spanish is quite different from 'radical' null subject languages like Japanese. Typological classification, then, becomes an additional consideration and the basis of research question two. The literature has suggested, however, that regardless of L1, the effects of language transfer are still persistent at advanced levels of L2 proficiency. This means, that learners are still omitting pronouns in L2 English—beyond the B2 level—just as they would in their null-subject first language (L1). Research question three queries whether proficiency, rather than L1, is the key factor in accurate L2 grammatical production.

The majority of the research on these questions has focused on smaller-scale studies of L1 Spanish or Italian learners of English. To address this, I use data from the EFCAMDAT learner corpus to sample a large and typologically diverse set of learners. This corpus contains over a million writing samples from 174,000 learners representing 198 nationalities. For this study, writing samples from learners with L1 Arabic, L1 Chinese (Mandarin), and L1 Russian were selected and analysed for their accuracy in pronoun production and omission. The results of chi-square and logistic regression analysis indicate that null-subject transfer is limited. There was a clear effect of L1, but not typology. L1 Arabic and L1 Chinese learners had nearly identical rates of accuracy ($p < .001$) but L1 Russians performed far worse in comparison. A quadratic (U-shaped) function in performance by proficiency was observed for all L1s, however. Learners did not omit pronouns at the elementary (A1) level. By A2, there was a downward trend, indicating an increase in pronoun omission, which continued to the lower-intermediate (B1) level. By the upper-intermediate (B2) level, accuracy in pronoun production matched that at the initial A1 level, before fully resolving at the advanced (C1/C2) level. These results are contrary to those in previous studies on null-subject transfer which found a more linear pattern in accuracy. The empirical data provides evidence that L1 influences the acquisition of certain features in L2. This has implications on the learnability of certain features in SLA. The data also suggests that language development is non-linear, and learners may struggle with transfer at the A2/B1 threshold. However, in the absence of experimental data to support these findings, the implications are tentative pending further investigation.

Is L2 parsing qualitatively different from native parsing? The role of context and individual differences in Working Memory Capacity and proficiency

Ehsan Solaimani (University of Essex), Florence Myles (University of Essex) and Laurel Lawyer (University Of Essex).

There is increased interest in the way native (L1) and second language (L2) readers resolve structural ambiguities in real-time, and it is generally assumed that investigating L2 disambiguation can provide insight into the processes underlying L2 parsing. Specifically, sentences containing relative clause (RC) ambiguities such as (1) have featured prominently in the native and L2 processing literature (e.g., Pan et al., 2015). (1) Someone shot the servant DP1 of the actress DP2 [RC who was on the balcony] In (1), the RC who was on the balcony has two potential host sites for its attachment. Preference for attaching the RC to the first DP, the servant, is referred to as DP1 attachment and preference for attaching it to the second DP, the actress, is referred to as DP2 attachment. Despite the wealth of research, however, there is no consensus on how readers process these structures, especially as far as L2 processing is concerned. For instance, little is known about the role of previous context on RC attachment preferences, or about the way in which L2 parsing preferences are impacted by individual differences in proficiency and working memory capacity (WMC). To contribute to a better understanding of RC disambiguation in English, two self-paced reading (SPR) tasks examined attachment preferences in isolated sentences (SPR 1) and in context (SPR 2), taking into account the impact of L1, proficiency, and WMC. Context was operationalised in terms of preceding sentences that biased attachment to either DP1 (DP1-supporting context) or DP2 (DP2-supporting context). The participants included 44 L1-French and 48 L1-Persian readers, in addition to 40 native English readers as the control group. The results of SPR 1 showed that attachment preferences were a function of WMC in all groups, with higher WMC associated with DP2 attachment. This is consistent with the Chunking Hypothesis (Swets et al., 2007), according to which, unlike low WM readers, high WM individuals chunk the complex DP (DP1-of-DP2) and the RC as one prosodic phrase and therefore prefer the linearly closer antecedent DP2 to ease processing burden. Similarly, the results of SPR 2 indicated that while there was no Reaction Time difference in a DP2-supporting context between DP1- and DP2-attached conditions, DP1- was favoured to DP2-attachment in a DP1-supporting context. Furthermore, context effects interacted significantly with proficiency for the L1-Persian readers, with high proficiency associated with stronger context effects. This suggests that that while the less proficient L1-Persian readers might differ from the other two groups in attachment preferences, they converge on a native-like interpretation at higher proficiency. Overall, the results suggest that L1 and L2 attachment preferences are a function of, among others, previous context and individual differences in WMC and proficiency.

Does more time lead to better L2 writing performance?

Myeongeun Son (Gangneung-Wonju National University) and Jongbong Lee (Cyber Hankuk University of Foreign Studies).

The amount of time assigned to completing a task has been a crucial topic in assessing second language (L2) writing assessments (e.g., Weigle, 2002). The controversy centers on whether more assigned time leads to better L2 writing performance. Previous studies on L2 writing assessment have yielded mixed results regarding the effect of different time allotments (e.g., Knoch & Elder, 2010; Khuder & Harwood, 2015; Lee et al., 2021; Wu & Erlam, 2016). Most of these studies manipulated only one writing task to investigate the effects of time allotment. However, considering task difficulty and the cognitive processes required to complete a task, the effects of time allotment may differ depending on the task type. Specifically, for a complex task, L2 learners may benefit from having more time, which they can use to monitor and edit their writing carefully. In contrast, for simple tasks, they may not require this additional time, leading to negligible effects when time constraints differ. To address this research question, this study examines the extent to which different time allotments affect L2 learners' writing performance in diverse tasks.

In this study, 128 Japanese intermediate-level EFL learners completed five picture description tasks, one narrative, and two argumentative writing tasks for a combined 30 minutes. All participants had learned English in instructional settings and were familiar with the three task types through previous English standardized test experiences. The learners were assigned to five different time allotment groups, which were given between two and four minutes (2, 2.5, 3, 3.5, and 4 minutes) for each of the picture description tasks and between three and five minutes (3, 3.5, 4, 4.5, and 5 minutes) for each of the narrative and argumentative tasks in counterbalanced order. The order of the tasks was also counterbalanced. The students used Microsoft Word (with the spelling and grammar check functions disabled) to complete the tasks. Writing performance was analyzed in terms of syntactic and lexical complexity using a syntactic complexity analyzer (Lu, 2010) and Coh-Metrix (McNamara et al., 2014), respectively. Accuracy was analyzed in terms of error-free T-units/total T-units by two coders, and fluency was analyzed by number of words.

Linear mixed models revealed that the learners' writing performance was affected by the allotted times, regardless of task type, and interaction was observed between time and task. In other words, the effects of time allotment differed across task types. Specifically, the L2 learners' linguistic outcomes were significantly more complex and fluent when the learners completed an argumentative task with a longer time constraint. However, the effects of different time allotments were negligible in narrative tasks. Regarding accuracy, no difference emerged between task types and allotted times. This result corroborates some, but not all, existing research. The findings are discussed in terms of their theoretical and pedagogical implications.

L2 Speaking Anxiety: A Subjective and Objective Analysis

Gisela Sosa-López (University of Barcelona).

Psycho-social factors such as Foreign Language Anxiety (FLA) are an important source of individual differences in the acquisition of foreign languages [1]. FLA is known to have a negative effect on language achievement, since it hinders learners' ability to perform in a FL class successfully [2]. Also, FLA has always been studied as a static construct; however, it varies in intensity and duration, and fluctuates over time [3]. Despite extensive previous research on FLA, L2 speaking anxiety (the feelings of nervousness when speaking the target FL) and the reason of its fluctuations has been under-researched, even in instructed FL classrooms where learners report speaking-oriented activities, especially oral presentations, to be highly anxiety-evoking [4, 5]. FLA has been assessed subjectively through questionnaires; however, more recent research [3] includes physiological measures of emotional arousal. Nonetheless, the relationship between subjective and physiological measures of speaking anxiety are still not well understood. The present study investigates the distress levels generated by two oral presentation contexts, their cause as well as the alignment between post-task subjective perceptions of speaking anxiety (by speakers and their audience) and physiological measures of emotional arousal (heart rate and electrodermal response) during task performance. Sixty-seven EFL learners performed 4 speaking tasks in counterbalanced order in the L2 comprising two presentation modes (individual vs. audience) and two task types (descriptive vs. argumentative), and comparable baseline speaking tasks in the L1. After task performance, participants self-assessed their distress with questionnaires and the idiodynamic approach [5], in which learners watched their videos and self-assessed their levels of distress that accompanied their presentation. Additionally, stimulated recalls were carried out for the L2 oral presentations. Finally, proficiency, vocabulary size and working memory were controlled variables. It is hypothesized that these data, currently under analysis, will yield main effect of speaking task and context such that argumentative speaking tasks performed in front of an audience will generate the highest levels of anxiety. The stimulated recalls will provide informative insights into why learners experience L2 speaking anxiety. Finally, it is hoped that the current study will uncover converging explanatory associations among the various quantitative and qualitative data obtained.

References

- [1] Teimouri, Y., Goetze, J., & Plonsky, L. (2019). Second language anxiety and achievement: a meta-analysis. *Studies in Second Language Acquisition*, 41(2), 363-387. [2] Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, 70(2), 125-132. [3] Gregersen, T., MacIntyre, P. D., & Meza, M. D. (2014). The motion of emotion: Idiodynamic case studies of learners' foreign language anxiety. *The Modern Language Journal*, 98(2), 574 - 588. [4] Price, L. M. (1991). The subjective experience of foreign language anxiety: Interviews with highly anxious students. In E. K. Horwitz and D. J. Young (Ed.) *Language anxiety: From theory and research to classroom implications*, pp.101–108. Prentice Hall. [5] Young, D. J. (1990). An investigation of students' perspectives on anxiety and speaking. *Foreign Language Annals*, 23, 539–553.

Adjective intensification in young multilingual learners of L2 Italian and German from South Tyrol

Stefania Spina (Università per Stranieri di Perugia / Eurac Research, Bolzano/Bozen), Aivars Glaznieks (Eurac Research, Bolzano/Bozen) and Andrea Abel (Eurac Research, Bolzano/Bozen).

Intensification is of considerable interest in the field of SLA. The possibility of grading a given quality, adding nuances of intensity to linguistic productions, can be expressed by lexical (extremely cold), morphological (ice-cold) and syntactic (as cold as possible) means. Previous studies have shown that intensifiers are also pragmatic resources (Hinkel 2003): the expression of the intensity of a quality conveys at the same time the subjective attitude of the speaker towards the propositional content. Moreover, the relationship between the intensifying and the intensified forms may result in highly conventional word combinations, which pertain to the domain of phraseology (Hendrikx et al. 2019). Being at the interface of lexicon, grammar, pragmatics and phraseology, intensification represents a challenge for L2 learners, who are required to acquire the linguistic and pragmatic knowledge associated with intensifiers, as well as the collocational patterns established between intensifying and intensified forms. Previous studies have investigated adjective intensification in advanced learners of English (Lorenz, 1998; Hinkel, 2003), while few have focused on other languages (Hendrikx et al. 2019), and on younger learners (Hasselgård 2022; Pérez-Paredes & Díez-Bedmar 2012). Some of their most relevant results highlight learners' overuse of all-purpose, delexicalized adverbial intensifiers compared to native speakers (Lorenz 1998), as well as a longitudinal tendency towards target-like uses of intensification, increased by target language exposure (Hendrikx et al. 2019). This study aims to provide new insights into L2 acquisition and use of intensification, by answering the following research question: Are there differences in the way L2 Italian and L2 German young learners from the multilingual area of South Tyrol acquire over time and use adjective intensification in written texts? On the basis that "language contact situations are a privileged vantage point for the observer to tease out all the manifestations of intensification in everyday language use" (Fiorentini & Sansò 2017: 176), we have investigated large multilingual corpora of two under-researched languages, focusing on how adjective intensification evolves over time within a multilingual repertoire of young learners, and thereby contrasting the two learner groups.

References Fiorentini, I., & Sansò, A. (2017). Intensifiers between grammar and pragmatics: A lesson from a language contact situation. In M. Napoli & M. Ravetto, Eds., *Exploring Intensification. Synchronic, diachronic and cross-linguistic perspectives*, 173–192. Benjamins. Hasselgård, H. (2022). Adverb-adjective combinations in young writers' English (EL1 and EL2). *Nordic Journal of Language Teaching and Learning*, 10(2). Hendrikx, I., Van Goethem, K. & Wulff, S. (2019). Intensifying constructions in French-speaking L2 learners of English and Dutch: cross-linguistic influence and exposure effects. *International Journal of Learner Corpus Research*, Vol.5:1, 63–103. Hinkel, E. (2003). Adverbial markers and tone in L1 and L2 students' writing. *Journal of Pragmatics*, 35(7), 1049–1068. Lorenz, G. 1999. Adjective intensification - learners versus native speakers: A corpus study of argumentative writing. *Rodopi*. Pérez-Paredes, P. & Díez-Bedmar M.B. (2012). The Use of Intensifying Adverbs in Learner Writing. In Y. Tono, Y. Kawaguchi, M. Minegishi (eds), *Developmental and Crosslinguistic Perspectives in Learner Corpus Research*, Benjamins, 105–124.

Intentional and incidental vocabulary learning: The role of historical linguistics in the second language classroom

James Stratton (University of British Columbia).

Building on the implicit and explicit learning paradigm (e.g., Norris & Ortega, 2000), the present study examines whether knowledge of language history can be beneficial when learning historically related languages. As Germanic languages, English and German share many cognates – words that trace back to the same ancestral form. However, due to various historical changes, many German cognates are no longer easily recognizable to most naïve English-speaking second language (L2) learners of German.

While many scholars have called for the use of historical instruction in the L2 classroom (e.g., Wolff, 1993), to date no studies have tested the effects of historical instruction on German vocabulary learning under ecologically valid experimental conditions. Knowledge of language history, such as the Second Germanic Sound Shift can explain several differences between English-German cognates. For instance, the interdental fricative (Modern English “th”) shifted to the stop [d] in German, explaining differences such as “Ding”–“thing”, “Dorn”–“thorn”, and “Durst”–“thirst”. Declarative knowledge of relevant historical semantic changes may also be beneficial for learners. For instance, making explicit connections between cognates such as German *sterben* ‘to die’ and its English counterpart ‘starve’ may aid memory retention, as learners can draw upon a representation (i.e., ‘starve’) that already exists, a strategy that has been shown to be effective in vocabulary learning (Hall, 2002).

Thirty-five English-speaking L2 learners of third-semester German at an American university were assigned to either an intentional or incidental learning condition. The intentional group received explicit instruction on two historical sound shifts (Second Germanic Sound Shift, Ingvaenic Palatalization) and relevant historical semantic changes (semantic broadening, narrowing, pejoration, amelioration). In contrast, the incidental group carried out various activities that exposed learners to cognates through reading and oral communication tasks. Following a pre/post/delayed-post-test design, vocabulary knowledge was assessed on an isolated translation task containing 126 words (63 cognates, 63 distractors). Forty-two cognates were words that both learning conditions encountered during their six 20-minute training interventions, with 21 words included that they would not encounter outside of the assessments.

Results indicate that the intentional group significantly outperformed the incidental group, suggesting that knowledge of language history can be beneficial when learning historically related languages. Declarative knowledge of the historical changes led to significantly greater vocabulary gains and it also provided learners with a tool kit to correctly predict the meaning of several cognates they had not previously learned about. The results are discussed in the context of implicit and explicit learning, intentional versus incidental vocabulary acquisition, with broad implications for vocabulary learning and language teaching.

Selected References Hall, C. J. (2002). The automatic cognate form assumption: Evidence for the parasitic model of vocabulary development. *IRAL*, 40, 69–87. Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50, 417–528. Wolff, R. A. (1993). The history of the language as an instructional aid. *Die Unterrichtspraxis*, 26, 27–35.

Comparison of morphological analyzers for L2-Korean written corpora

Hakyung Sung (University of Oregon) and Gyu-Ho Shin (University of Illinois Chicago).

With the recent advancement of computational approaches to L2 research, automatic processing of learner corpora has been gaining momentum for a better understanding of L2 development. Research on L2 Korean in this context is extremely thin, mostly concerning error analysis (Kim et al., 2016; Lee et al., 2012), lexico-grammatical measurement (Lim et al., 2022; Nam & Hong, 2014), and item extraction such as functional morphemes (Shin & Jung, 2022; Jung, 2022) or clausal constructions (Shin & Jung, 2021). While these studies often proceed with automatic tokenization/tagging of the L2 data, less is known about whether and how reliably the currently available morphological analyzers work for L2-Korean data. Indeed, there are language-specific challenges to the automatic processing of L2-Korean texts. For example, it is unclear how the properties of Korean (e.g., agglutination, overt case-marking via dedicated markers, suffixation, scrambling/omission of sentential components) are faithfully reflected in these analyzers (Shin & Jung, 2021). In addition, it seems that L1-based language models/pipelines are applied directly to L2 data without serious resort to learner language characteristics (e.g., spacing/spelling errors, novel combinations of words/chunks; Kyle, 2021; Meurers & Dickenson, 2017). To overcome these caveats, the present study compares various Korean morphological analyzers with respect to L1/L2-Korean written corpora and see how they handle learner language.

Methods. We first developed a gold-standard test set for L2-Korean texts, by using portions of the National Institute of Korean Language learner corpus (27,299 written and 2,541 spoken text files collected from 2015 to 2020). We randomly sampled 100 sentences from argumentative essays in the dataset. Each sentence was independently annotated by two native speakers of Korean using the Sejong tag set, and the initial annotation was cross-validated. For a reference L1-Korean corpus, we used a test set (n=989) of Google Korean Universal Dependency Treebank (UD Korean GSD) comprising 6K sentences from online blogs and news (Chun et al., 2018; McDonald et al., 2013). Finally, we selected nine Korean open-access morphological analyzers (based on various computational algorithms [ranging from statistical models to deep-learning models] and training data), applied them in parallel both to the L2-Korean dataset (as a target) and to the L1-Korean dataset (as a reference), and compared their performance on tokenization and POS-tagging.

Results. The deep-learning-based analyzers achieved the highest accuracy of both tokenization and tagging while the tagging accuracy of the L2 dataset was comparatively lower than that of the L1 dataset (Table 1). When focusing on the three analyzers with the highest accuracy, they demonstrated by-tag performance asymmetries in the L2 dataset (Table 2), especially low performance on verb-related tags (VV: verb/main; VX: verb/auxiliary). These findings call for the need to empirically investigate, and thoroughly refine, the currently available (L1-based) pipelines to cope with resources on understudied languages/registers.

References Chun et al. (2018). Proceedings of LREC; Jung (2022). Korean Linguistics 18(2); Kim et al (2016). Teaching Korean as a Foreign Language 44; Kyle (2021). International Journal of Learner Corpus Research 7(1); Lee et al (2016). Language Facts and Perspectives 38; Lim et al (2022). Natural Language Engineering; McDonald et al (2013). Proceedings of ACL; Meurers & Dickinson (2017). Language Learning, 67(S1); Nam & Hong (2014). The Journal of the Humanities for Unification, 57; Shin & Jung (2021). International Journal of Learner Corpus Research 7(1); Shin & Jung (2022). Australian Review of Applied Linguistics

Gaming as an extramural activity for L2 incidental vocabulary acquisition

Iwarin Suprapas (The University of Sheffield) and Beatriz González-Fernández (The University of Sheffield).

Second Language Acquisition (SLA) research has highlighted the importance of out-of-classroom (a.k.a. extramural) input for successful second/foreign language (L2) development (Richards, 2015). This is further accentuated in contexts where L2 input and classroom time is limited (Peters, 2018). Certain extramural activities (e.g., reading, TV viewing) have been found to promote incidental lexical acquisition (Webb, 2020). However, no research to date has thoroughly explored the effectiveness of gaming for vocabulary acquisition, despite it being one of the most popular leisure activities among L2 learners (Rodgers & Heidt, 2021). This exploratory study aims to investigate whether vocabulary can be learnt from playing a computer game and the influence of learner-specific factors (i.e., L2-English vocabulary size and gaming habits) on learning gains.

The study adopts a pretest-posttest design. Twenty-five L2 gamers from different L1s at university level completed the Vocabulary Size Test (Nation & Beglar, 2007) and a background questionnaire on their gaming habits. One week later, participants completed an introductory quest in a computer game (Skyrim) which had been modified to include 20 target non-words (retrieved from ARC nonword Database) that replaced object names in the target quest. Participants were also instructed to collect and report the number of coins as an index of their engagement with the game. Two tests were administered to assess receptive and productive knowledge of the target non-words' form-meaning mapping immediately after completing the quest and again one week later. Repeated measures ANOVAs showed significant gains in receptive and productive mastery of form-meaning mapping immediately after playing the game, with most of this knowledge being retained one week later. Additionally, significant positive correlations were found between vocabulary gains (target-nonwords learnt) and learner-specific factors (participants' prior vocabulary size and gaming habits), with the level of engagement with the game showing a stronger correlation. Overall, the findings indicate that gaming is an effective activity for vocabulary development, suggesting that games can serve as valuable extramural resources to promote L2 acquisition. The implications for L2 teaching and research will be discussed.

Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9–13. Peters, E. (2018). The effect of out-of-class exposure to English language media on learners' vocabulary knowledge. *ITL - International Journal of Applied Linguistics*, 169(1). Richards, J. C. (2015). The changing face of language learning: Learning beyond the classroom. *RELC Journal*, 46(1), 5–22. Rodgers, M. P. H., & Heidt, J. (2021). Levelling up comprehensible input and vocabulary learning. In *Pop Culture in Language Education: Theory, Research, Practice*. Routledge. Webb, S. (2020). Incidental Vocabulary Learning. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 225–239). Taylor & Francis Group.

Enhanced new vocabulary learning in adolescent L2 French learners through the use of pre- and post-viewing activities

Lea Suter (University of Fribourg), Raphael Berthele (University of Fribourg) and Anita Thomas (University of Fribourg).

Audio-visual inputs such as films and series are increasingly used in foreign language classes as they provide authentic exposure to the language. Several studies suggest they support the learning of new vocabulary (e.g. Muñoz et al., 2021; Rodgers & Webb, 2020), and this learning may be enhanced through the use of pre/post-viewing activities (Schmitt, 2008). However, these theories have not yet been studied in the context of adolescent L2 French learners. It is therefore necessary to study the effect of pre/post-viewing activities, and investigate the importance of activity sequence (i.e. pre vs post) to effectively enhance the learning experience when utilising audio-visual inputs in foreign language teaching. This study investigates the effect of presence and sequence of viewing activities on the learning of target words. Seventeen target words of differing word class, length, frequency and occurrence were selected from three 13-15 minute excerpts from the French TV series *Plan Cœur* (total 51 target words; Netflix, 2018). A within-subjects design study was then undertaken in compulsory secondary school in Switzerland, where 97 beginner level L2 French students, aged 13-14, watched the three excerpts with French captions. The participants watched each excerpt in a different condition (episode only, episode and pre-viewing activity, episode and post-viewing activity). In the activities, the participants had to recognise the correct meaning (translation) of target words. Three immediate and one delayed post-test of the same type (meaning recognition) were administered to four upper-level classes (HI) and two lower-level classes (LO). We fit a mixed effects logistic regression model to our data which includes random intercepts for each student, each item and each class. The immediate post-test results show that learners in the HI track have a higher (+20%) probability to recall the meaning of the target words than learners in the LO track. With an activity (regardless of pre or post), the probability of a correct recall significantly increases by 15% (HI) and 18% (LO), and there is no significant interaction with track. Similar patterns were observed with the delayed post-test results; however these were not significant. These results highlight the importance of pre/post-viewing activities when learning new vocabulary through audio-visual input in foreign language teaching.

Muñoz, C., Pujadas, G., & Pattemore A. (2021). Audio-visual input for learning L2 vocabulary and grammatical constructions. *Second Language Research*. May 2021, 1–25. DOI: 10.1177/02676583211015797

Plan Cœur (2018). French TV series by Noémi Saglio, Julien Teisseire and Chris Lang. Netflix.

Rodgers, M.P.H. & Webb, S. (2020). Incidental vocabulary learning through viewing television. *ITL - International Journal of Applied Linguistics*, 171 (2), 191–220. <https://doi.org/10.1075/itl.18034.rod>

Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research*, 12 (3), 329–363. DOI: 10.1177/1362168808089921

The moderating role of L2 automaticity in the predictive power of L1 fluency for L2 utterance fluency

Shungo Suzuki (Waseda University) and Judit Kormos (Lancaster University).

Oral fluency has been found to be a robust indicator of second language (L2) proficiency (Baker-Smemoe et al., 2014; Tavakoli et al., 2020) and thus is commonly regarded as an essential construct in various assessment contexts of L2 speaking skills. In accordance with the conceptualization of L2 fluency as a proxy for L2-specific oral proficiency, scholars have continuously examined the validity of utterance fluency (UF) measures in terms of the independence from language-general processing skills. Such language-general aspects of fluency have been operationalized as the covariance between L1 and L2 UF measures (Bradlow et al., 2017; Duran-Karaoz & Tavakoli, 2020). Although prior research has shown that the L1-L2 UF link is cross-linguistically robust (Bradlow et al., 2017), previous studies were largely limited to the combination of stress-/syllable-timed languages. In addition, previous studies suggested that the strength of the association between L1 and L2 fluency can vary as a function of L2 oral proficiency (Derwing et al., 2009; Huensch & Tracy-Ventura, 2017). To further explore the generalizability of the L1-L2 UF link, the current study examines the moderator effects of L2 oral proficiency on L1-L2 UF link in the context of argumentative speech produced by L1 Japanese (mora-timed language) learners of English (stress-timed language).

A total of 104 Japanese learners of English completed two argumentative speaking tasks in L2 English as well as another argumentative task in L1 Japanese. Their self-reported university placement test scores suggested that most of them were placed on the B1–B2 levels of the CEFR scale. Their speech samples were transcribed and annotated for calculating a set of UF measures, capturing three major components of fluency—speed, breakdown and repair fluency (Tavakoli & Skehan, 2005). To operationalize L2 oral proficiency with high compatibility with UF performance, we adopted the measurements of cognitive fluency (De Jong et al., 2013; Segalowitz, 2010), using a set of linguistic knowledge test scores, covering vocabulary, grammar and pronunciation. Two separate factor scores were computed by means of confirmatory factor analysis (CFA): linguistic resources (breadth of linguistic knowledge) and processing speed (speed of linguistic processing).

A series of generalized mixed-effects modelling showed that all of the L2 UF measures were significantly predicted from the corresponding L1 UF measures the small-to-medium effect size of the overall predictive power of L1 fluency (Marginal $R^2 = 5.2$ – 24.7%). The predictive power of L1 UF was ignorable in articulation rate, self-correction ratio, and false start ratio (Marginal $R^2 = 5.2$ – 7.1%), while the relatively strong predictive power was indicated in mean length of run, filled pause ratio, and self-repetition ratio (Marginal $R^2 = 17.1$ – 24.7%). The results also revealed that the significant moderator effects of L2 processing speed on the L1-L2 UF link were found only in speed fluency measures—articulation rate, speech rate and mean length of run. These findings suggest that the L1-L2 UF link is subject to the dimensionality of fluency measurements and the degree of L2 automaticity.

Assessing the role of accent exposure in reducing foreign language pronunciation learning anxiety

Yui Suzukida (Juntendo University, University College London).

Anxiety has been considered one of the key psychological factors that affect various dimensions of L2 performance and learning. This is because anxiety influences behavioural, psychological, and cognitive aspects of language use and learning. For instance, anxious L2 learners have been found to have a lower level of motivation towards learning a L2, less willingness to communicate in their L2, and as a result are assumed to experience impaired attention impairment, taking longer to store information in memory, and impaired memory retrieval, hampering smooth L2 processing and production. In the context of L2 pronunciation learning (the focus of the current study), researchers consider pronunciation to be particularly anxiety-evoking due to learners' negative attitudes towards non-native accents and embarrassment relating to their own L1-influenced accents. While the introduction of diverse forms of the target language through classroom instruction is often suggested (Matsuda, 2003), not many attempts have been made to examine the effectiveness of such instruction for anxiety reduction (Bashori et al., 2020).

To respond to the current knowledge gap, the present study was set to evaluate the effectiveness of exposure to multiple varieties of English in reducing L2 pronunciation anxiety. A total of 104 Japanese learners of English in Japan were recruited and allocated into three groups: a treatment group ($n = 44$) and two control groups ($n = 30$ per group). The treatment group listened to different English varieties via the IDEA database (Meier & Muller, 2023), discussed the comprehensibility of the speech, and identified noticeable features of the speech. The first control group received pronunciation training focusing on segmental and suprasegmental features related to comprehensibility, and the second control group did not receive any instruction related to L2 speaking and listening but engaged in grammar and vocabulary exercises. The participants' degree of pronunciation anxiety (Baran-Łucarz, 2016) was measured via a questionnaire before and after the 10 weekly interventions.

The results of repeated measures of ANOVA showed that there was a group difference between the degree of anxiety at pre- and post-test, and the post-hoc analyses revealed that the degree of pronunciation anxiety of the treatment group significantly decreased whereas that of the two control groups did not change or increased. The result suggested that (a) exposure to different varieties of English contributed to the reduction of L2 learners' anxiety related to L2 pronunciation learning and use, and (b) learners' anxiety levels may be resistant to change even after they have the opportunity to learn and practice pronunciation.

References

Baran-Łucarz, M. (2016). Conceptualizing and Measuring the Construct of Pronunciation Anxiety: Results of a Pilot Study. In M. Pawlak (Ed.) *Classroom-Oriented Research* (pp. 39–56). Springer, Cham. Bashori, M., van Hout, R., Strik, H., & Cucchiari, C. (2020). Web-based language learning and speaking anxiety. *Computer Assisted Language Learning*, 35, 1058–1089. Meier, P. & Muller, S. (2023). IDEA: International Dialects of English Archive. Retrieved June 3, 2020, from <https://www.dialectsarchive.com/> Matsuda, A. (2003). Incorporating world Englishes in teaching English as an international language. *Tesol Quarterly*, 37, 719-729.

Vocabulary Knowledge and Language Proficiency as Predictors of Academic Achievement in the Vietnamese context

Csaba Zoltan Szabo (University of Nottingham Malaysia), Barry Lee Reynolds (University of Macau) and Phat Cao (University of Nottingham Malaysia).

For more than a decade, the Pearson Test of Academic English (PTEA) scores have been extensively used as a linguistic parameter for English-medium degree programmes. However, compared to IELTS and TOEFL, research on the predictive validity of the PTEA is much scarcer (i.e., Durrant, et al., 2015; Riazi, 2013; Wilson, 2016). The extant research does show that standardised tests of English proficiency, such as the IELTS are relatively restricted in predicting academic success while linguistic aspects, like vocabulary size or knowledge have been found to contribute significantly to academic performance (e.g., Daller & Xue, 2009; Masrai & Milton, 2017). This study therefore examined the predictive validity of the PTEA by investigating the relations between language proficiency as measured by the PTEA scores, vocabulary size by the vocabulary size test (VST) (Nation and Beglar, 2007) and academic performance by grade point average (GPA). Moreover, the association between academic performance and the three main individual factors (i.e., autonomous learning activities, language learning attitudes and course perception and evaluation) were also investigated via a language background questionnaire adapted from Durrant, et. al.'s (2015) study. The participants of this study include pre-service English as a Foreign Language (EFL) student-teachers (N = 120) at one teacher-training university in Vietnam. The findings show that employing both the PTEA and VST scores can explain 25% the variance of participants' academic grades, which indicates a superior predictive validity compared to previous studies on IELTS and the TOEFL. The partial least squares structural equation modelling (PLS-SEM) results also highlight the predictive relevance of the three previously-mentioned individual factors for academic success. Implications are also discussed as they can contribute to the decisions on admissions and language support courses based on language proficiency tests, lexical measures as well as individual factors.

Repetition and incidental learning of multiword units: a conceptual replication study of Webb et al. (2013)

Paweł Szudarski (University of Nottingham), Elke Peters (KU Leuven) and Eva Puimège (KU Leuven).

There is increasing evidence that activities like L2 reading (Pellicer-Sánchez, 2017) and reading-while-listening (Webb & Chang, 2020) can promote incidental learning of multiword units. Webb, Newton and Chang (2013) examined the effect of reading-while-listening on learners' knowledge of multiword units, as well as the effect of number of exposures, in a between-participants design with four experimental conditions (one, five, ten, or fifteen exposures). Their findings revealed large effects of repetition on various aspects of knowledge of multiword units, with considerably larger gains in the 15 exposures condition compared to the other conditions. This finding has not been replicated in other studies on incidental learning of multiword units, with most research finding much smaller repetition effects (e.g., Pellicer-Sánchez, 2017). In a conceptual replication of Webb, Newton, and Chang (2013), the present study investigates the effect of repetition on incidental L2 learning of multiword units. In order to assess the reproducibility of findings across different learning contexts, we are collecting data from high school and university students of L2 English in Belgium and Poland. The present study also addresses some limitations of this initial study (e.g., task duration effects resulting from a between-participant design) by adopting a within-participants design. In an online experiment, participants read and listen to a modified graded reader in which target multiword units (e.g., buy time) occur 1-15 times. Participants are randomly assigned to four versions of the text, in which counterbalanced item sets are repeated a different number of times. To measure learning of the multiword units, we use pre- and posttests of form recall (L1 to L2 translation) and form recognition (multiple choice). At present, 119 participants have already completed the experiment, with the aim to collect data from 280 students in total. Posttest scores will be analyzed in mixed effects logistic regression models, with item-level gain score as outcome variable. In addition to the experimental conditions analyzed by Webb et al., we will also factor in learner (L1, education level, prior vocabulary knowledge) and item variables (corpus frequency, congruency). The results from preliminary analyses show that significant gains were made in both test formats, and that repetition positively affected learning, in line with the original study. Final analyses on the full data set will reveal to what extent learning is moderated by variables such as L1 background and education level. The results will have important implications for L2 theory and pedagogy, by expanding our knowledge of the role of exposure frequency in incidental learning of multiword units at different stages of learning and for learners from across different L1 backgrounds. The study also contributes to the improvement of methodological rigor by testing the feasibility of conducting L2 learning experiments online. Pellicer-Sánchez, A. (2017). Learning L2 collocations incidentally from reading. *Language Teaching Research*, 21, 381–402. Webb, S., & Chang, A. (2022). How does mode of input affect the incidental learning of collocations? *Studies in Second Language Acquisition*, 44, 35-56. Webb, S., Newton, J., & Chang, A. (2013). Incidental learning of collocation. *Language Learning*, 63, 91–120.

Using multiword sequences in a dialogic speaking task across levels of proficiency

Parvaneh Tavakoli (University of Reading), Takumi Uchihara (Tohoku University) and Svetlana Mazhurnaya (University of Reading).

Examining multiword sequences (MWSs) has been a burgeoning area of research with the existing psycholinguistic evidence suggesting that the fast and accurate processing of MWSs lead to an increased speed in speech processing and performance (Siyanova-Chanturia & Pellicer-Sánchez, 2019). Previous research (e.g., Eguchi & Kyle, 2020, Kyle et al., 2018; Tavakoli & Uchihara, 2020) has indicated that with the development of proficiency, L2 users tend to use a greater number of high-frequency and more sophisticated MWSs. The existing body of research on the use of MWSs across proficiency levels, however, is limited in that it is mainly based on either written data or spoken data elicited in monologic tasks (e.g., Tavakoli & Uchihara, 2020), with no evidence of whether use of MWSs in a dialogue increases with proficiency development. The current study is aimed at investigating this gap by examining the use of MWSs in a dialogic task across assessed levels of proficiency in Test of English for Educational Purposes (TEEP). Providing evidence about the relationship between MWSs and proficiency is central to a) understanding the speech production process in a dialogic mode, and b) facilitating the automated assessment of proficiency.

The data comprises 129 speech samples from speakers at six proficiency levels of the Test of English for Educational Purposes (TEEP) from 5.0 to 7.5 levels (equivalent to A2-C1 CEFR levels) performing a dialogic discussion task. The data were transcribed and subjected to MWSs analysis in TAALES (Kyle et al., 2018) for bigram and trigram measures of frequency and sophistication (i.e., MI and Delta P). The analysis of mixed-effects linear regression was used to investigate the relationship. The results suggest that learners of the lowest proficiency level (5.0 group) produced fewer sophisticated two-word sequences than the other five groups, although no differences were observed across the remaining five groups. Furthermore, two groups of higher proficiency levels (6.0 and 7.5) produced more sophisticated three-word sequences than did the lowest proficiency group (5.0). No significant differences were found for frequency measures across groups. These findings indicate that the ability to use MWSs may not be as sensitive to general oral proficiency required in dialogic tasks as in monologic tasks. It can be argued that the interactive and dynamic nature of a dialogue and its characterizing features of short turns and interruptions may affect the way MWSs are used in a dialogue. Implications will be discussed in relation to the development of an L2 speech production model and language assessment.

References

- Eguchi, M., & Kyle, K. (2020). Continuing to explore the multidimensional nature of lexical sophistication: The case of oral proficiency interviews. *The Modern Language Journal*, 104(2), 381-400.
- Kyle, K., Crossley, S., & Berger, C. (2018). The tool for the automatic analysis of lexical sophistication (TAALES): version 2.0. *Behavior research methods*, 50(3), 1030-1046.
- Siyanova-Chanturia, A., & Pellicer-Sánchez, A. (2019). *Understanding formulaic language: A second language acquisition perspective*. Routledge.
- Tavakoli, P., & Uchihara, T. (2020). To what extent are multiword sequences associated with oral fluency? *Language Learning*, 70(2), 506-547.

Clitic placement in adult L2 European Portuguese: when L1 and L2 acquisition meet

Joana Teixeira (Universidade NOVA de Lisboa), Alexandra Fiéis (Universidade NOVA de Lisboa) and Ana Madeira (Universidade NOVA de Lisboa).

Recent research has suggested that phenomena acquired late in L1 may be challenging in bilingual development, mainly due to input factors (Tsimpli, 2014; Sorace, 2014). Clitic placement in European Portuguese (EP), which depends on various factors (syntactic, lexical, semantic...), constitutes an appropriate testing ground for this hypothesis, since research on L1 acquisition (L1A) found that EP-speaking children begin by generalising enclisis to proclisis contexts and acquire knowledge of some of these contexts very late (Costa, Fiéis & Lobo, 2015). According to Costa et al. (2015), the acquisition of proclisis contexts follows a sequence (negation>negative subjects/subjunctive complement clauses>adverb já 'already'> adverbial clauses>quantified subjects). The contexts acquired later are subject to variation in adult grammars. Preliminary studies suggest that the acquisition route of proclisis may be similar in L1 and L2 EP (Gu, 2019, 2022). However, the following questions remain open: Are the proclisis contexts that are acquired late in L1 EP challenging in L2A? Is the developmental route of proclisis contexts in L2 EP similar to the one observed in L1A? Is convergence with the L2 possible wrt clitic placement?

To address these questions, this study investigates clitic placement in L2 EP. Participants were 25 L1 EP speakers and 30 L1 Spanish-L2 EP adult learners (10 intermediates, 10 advanced, 10 near-natives). As clitic placement is taught explicitly, we used two tasks that tap primarily into implicit knowledge (Ellis, 2005): an elicited oral production task (EOPT) and a timed acceptability judgement task (TAJT), which tested enclisis and proclisis in sentences without a proclisis trigger, negative sentences, indicative and subjunctive complement clauses, adverbial clauses, and quantified subjects.

Results in the EOPT show that, like the natives, learners prefer enclisis in the absence of proclisis triggers. In negation contexts and in subjunctive complement clauses, all groups display a preference for proclisis. In the other contexts, the native and near-native groups clearly prefer proclisis. Advanced learners prefer proclisis in indicative complement clauses and produce similar rates of enclisis and proclisis with adverbial clauses and quantified subjects. Intermediate learners exhibit optionality in indicative complement clauses and prefer enclisis in the remaining contexts. In the TAJT, all groups exhibit the same preferences regarding proclisis as in the EOPT, except the intermediate learners, who display optionality with adverbial clauses and quantified subjects.

These results show that enclisis stabilizes earlier than proclisis, which develops sequentially, following a route similar to that observed in L1A: the first contexts to be acquired are negative and subjunctive clauses. The contexts where the overgeneralization of enclisis is more persistent are sentences with adverbial clauses and quantified subjects. These proclisis contexts are described as less categorical in native grammars, which is confirmed by the control group's results. This may give rise to greater input variability, adding complexity to the acquisition task. As a result, prolonged input exposure is required to discover the patterns of clitic placement in both L1 and L2 EP. This study thus confirms the hypothesis that input alone can account for the late acquisition of some grammatical properties in L1 and L2A.

L2 use or happy music reduce bilinguals' morals. Why the moral foreign-language effect may have to do little with emotions

Dieter Thoma (University of Mannheim).

In moral decision-making, people normally have a deontological bias, i.e., they prefer options that comply with moral norms regardless of consequences (e.g., do not harm or kill humans) to utilitarian ones that maximize societal welfare (e.g., harm one for the benefit of many). Many studies found this deontological bias to be reduced if unbalanced bilinguals make moral decisions in their weaker second language (L2), compared to their dominant first language (L1) (Hayakawa et al., 2017). However, the reliability of the so-called moral Foreign Language Effect (FLE) and its causes are debated. Most language researchers claim that understanding a dilemma task in L2 triggers weaker emotional reactions, which in turn, would reduce the deontological bias. This implies that emotionality increases the bias. However, research from monolingual settings suggests that, if anything, emotion increases the deontological bias, e.g., in that happier decision-makers behave more utilitarian (Gawronski et al., 2018).

To resolve this apparent paradox, we investigated in how far the moral FLE is moderated by incidental emotion induction in reading and listening. We designed a 2 (language: L1 German | L2 English) x 2 (emotion induction: neutral | happy) x 2 (mode: reading | listening) experiment with language and mode as between-subject and emotion induction as within-subject factor. German-English bilinguals (N = 186) participated online and self-assessed their proficiency in both languages. In the main experiment, they read six moral dilemmas including personal force and four filler decision tasks, while listening to neutral or happy ambient music. Each dilemma required a binary choice between a deontological and utilitarian option, and participants rated their feelings after each decision as manipulation check for emotion induction. The data were analyzed with logistic mixed-effects regression models. The main results replicated a moral FLE with a higher proportion of utilitarian choices in L2. Happy music reduced the deontological bias in L1 more than in L2. Thereby, the FLE was significantly smaller in happy, relative to neutral trials. However, while this pattern was clear in the reading task, emotion induction made no difference when listening to the dilemmas in L1 or L2. Further, the FLE reduced consistently with increasing L2 proficiency. These findings indicate that the FLE has little to do with language-dependent bilingual emotions, yet it seems to vary with language processing fluency in any language.

References Gawronski, B., Conway, P., Armstrong, J., Friesdorf, R., & Hütter, M. (2018). Effects of incidental emotions on moral dilemma judgments: An analysis using the CNI model. *Emotion, 18*(7), 989–1008. <https://doi.org/10.1037/emo0000399> Hayakawa, S., Tannenbaum, D., Costa, A., Corey, J. D., & Keysar, B. (2017). Thinking More or Feeling Less? Explaining the Foreign-Language Effect on Moral Judgment. *Psychological Science, 28*(10), 1387–1397. <https://doi.org/10.1177/0956797617720944>

Evaluating explicit inductive and deductive corrective feedback on EFL 3rd sg -s acquisition: Introducing analogy-based corrective feedback

Kavita E. Thomas (University of Gävle).

While corrective feedback (CF) research largely compares the relative merits of implicit and explicit CF, few evaluate CF where explicitness remains constant (Li, 2010). Often explicit metalinguistic CF is compared to exemplar-based CF like recasts where evidence of errors is more implicit. Yet although recasts may go unnoticed, Li's meta-analysis found no significant difference between recasts and metalinguistic CF on immediate and short-delay posttests with greater retention for recasts on long-delay posttests. However, explicitness and mode (rule-based vs. exemplar-based) both differ between recasts and metalinguistic CF. Considering mode, rule-based (metalinguistic) CF involves deductively determining whether given rules apply while exemplar-based CF (e.g., direct corrections, analogy-based CF, recasts) enables inductively discovering underlying rules from examples. However, while direct corrections and recasts reproduce learners' utterances, correcting errors, analogy-based CF provides corrected syntactically identical utterances, lexically distinct, but close in meaning to the learners' erroneous utterances.

Analogy-based CF is motivated by mutual analogical alignment research (Gentner et al, 2003) advocating bidirectional mapping involving comparing and aligning relational structures. In L2 learning, analogy-based CF following syntactic errors requires learners to recognise analogous syntactic structures and align them in order to discover errors. According to Gentner et al, analogical encoding particularly benefits novices, who have shallower representations focused on surface-specific features, pushing discovery of deeper relational structure (akin to usage-based schema formation) and promoting transfer to new targets. Further motivation comes from syntactic alignment in dialogue which describes emergence of exemplar-driven syntactic convergence phenomena in interaction; in an L2 context, McDonough and Mackey (2006) linked learning to syntactically aligned but lexically distinct learner productions. Furthermore, unlike direct corrections, analogy-based CF requires deeper processing (Craik & Lockhart, 1972) to detect errors and revise.

A quasi-experimental classroom-based study on oral CF effectiveness on English 3rd sg -s with upper secondary Swedish learners (n = 49) is presented varying mode while keeping explicitness and revision prompting constant. Metalinguistic, direct correction and analogy-based CF groups and a control group are compared in a pretest, posttest, delayed posttest design with receptive grammaticality judgment tests and productive sentence completion tests. Results indicate significant delayed explicit knowledge gains for all corrective feedback on some receptive measures. There were however different trends over testing times for the different types of feedback, where analogy-based feedback resulted in worst performance immediately after treatment, rising for the delayed posttest unlike the other two feedback types which tended to have the highest means immediately after treatment, dropping with time.

References Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11. Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology*, 95. Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, 60. McDonough, K. & Mackey, A. (2006). Responses to recasts: Repetitions, primed production and linguistic development. *Language Learning*, 56(4).

Crowdsourced comparative judgement for L2 writing assessment: is high reliability still possible when texts are homogeneous in proficiency and diverse in topic?

Peter Thwaites (UCLouvain), Charalambos Kollias (National Foundation for Educational Research (NFER)) and Magali Paquot (UCLouvain).

Comparative Judgement (CJ) is a method of assessment in which judges (who may be experts, peers or even novices) decide which of two pieces of student work is “better”. These comparisons can be used to automatically rank each piece of work, and from these rankings a grade can be derived. Research from a range of fields has demonstrated that CJ generates reliable, valid, and efficient evaluations of learner outputs (Jones et al., 2019; Wheadon et al., 2020).

This presentation reports on findings from a project which explores the potential for CJ (in tandem with crowdsourcing) to contribute to the assessment of L2 texts included in widely used learner corpora that unfortunately often do not come with reliable proficiency measurement. A recently-published preliminary study (Author, 2022) suggested CJ’s suitability for this purpose. However, in focusing on relatively short writing samples (median = 272 words), using a single essay prompt and covering the full spectrum of L2 proficiency, the study provided optimal conditions for such results.

A next step is to find out whether CJ continues to be reliable when learner texts are longer, more homogeneous, and/or include responses to a range of prompts. Since all of these changes are likely to complicate the task of choosing the stronger of each pair of essays (van Daal et al., 2019), it is important to test their effects on the decisions made by judges. The two studies reported here therefore involve the comparative judgement of two sets of texts; firstly 48 learner essays taken from the International Corpus of Learner English, constituting responses to a single prompt and of a relatively narrow proficiency range (roughly B1-C1), and of greater length than the texts in the earlier study (median = 548 words); and secondly another set of texts, similar except that they comprised responses to five different topics. In addition to being evaluated using CJ, all texts were triple-graded using rubric-based methods. The presentation will compare the reliability and efficiency of these two assessment approaches.

The results contribute both to learner corpus research, where CJ can potentially enrich learner corpora with accurate measures of text proficiency; and to the more general goal of exploring reliable, valid, and efficient alternatives to rubric-based L2 writing assessment.

References Jones, I., Bisson, M., Gilmore, C., & Inglis, M. (2019). Measuring conceptual understanding in randomised controlled trials: Can comparative judgement help? *British Educational Research Journal*, 45(3), 662–680. <https://doi.org/10.1002/berj.3519>

van Daal, T., Lesterhuis, M., Coertjens, L., Donche, V., & De Maeyer, S. (2019). Validity of comparative judgement to assess academic writing: Examining implications of its holistic character and building on a shared consensus. *Assessment in Education: Principles, Policy & Practice*, 26(1), 59–74. <https://doi.org/10.1080/0969594X.2016.1253542>

Wheadon, C., Barmby, P., Christodoulou, D., & Henderson, B. (2020). A comparative judgement approach to the large-scale assessment of primary writing in England. *Assessment in Education: Principles, Policy & Practice*, 27(1), 46–64. <https://doi.org/10.1080/0969594X.2019.1700212>

The impact of input on the Harmonious Bilingual Development in French-Russian children

Ekaterina Tiulkova (University of Toulouse 2), Barbara Köpke (University of Toulouse 2) and Vanda Marijanović (University of Toulouse 2).

While input has been shown to play a major role in bilingual development (De Houwer, 2018; Unsworth, 2016), we do not yet have precise knowledge of the quantity and quality of input necessary to foster harmonious bilingualism in children growing up in a multilingual environment. When “subjective well-being is not negatively affected by factors relating to a bilingual setting”, the child experiences Harmonious Bilingual Development (De Houwer, 2020: 63). Several studies examined the relationship between children’s bilingual proficiency and their subjective well-being as well as the role of parental and school effort of promotion of dual language acquisition in social-emotional and behavioral skills (Han, 2010; Sun, 2022). However, the concept of Harmonious Bilingual Development still needs to be further explored through the collection of various personal testimonies. Therefore, we investigated the relationship between input and social-emotional and behavioral skills among 37 five-year-olds from French-Russian families residing in France. For this purpose, parent questionnaire and activity journal were used to assess children’s current language exposure as well as diversity of input (speakers, activities, media, reading). Children’s social-emotional and behavioral skills were assessed by parents with the Strengths and Difficulties Questionnaire. Language proficiency was determined through oral fluency (speech and articulation rate, mean syllable duration) in a picture-story retelling task in Russian and French. The participants were exposed to both languages from birth and are all considered as proficient in French. Five children, though, expressed difficulties with the narration task in Russian. A series of variables were taken into account for the data analyses, e.g., SES, birth order, family size. We will briefly present the research procedure including the data collection conditions (COVID period) and discuss the data with respect to our major interrogation about the input’s impact on harmonious bilingual development. Preliminary results show the direct link between more balanced input in both languages and better social-emotional and behavioral skills in bilingual children. We will discuss how our findings might support more positive attitudes towards early bilingualism among the parents and concerned educators.

De Houwer, A. (2018). The role of language input environments for language outcomes and language acquisition in young bilingual children. In D. Miller, F. Bayram, J. Rothman & L. Serratrice (Eds.), *Bilingual cognition and language: the state of the science across its subfields* (pp. 127-153). John Benjamins. De Houwer, A. (2020). Harmonious bilingualism: Well-being for families in bilingual settings. In S. Eisenclas & A. Schalley (Eds.), *Handbook of home language maintenance and development* (pp. 63–83). De Gruyter Mouton. Han, W. J. (2010). Bilingualism and socioemotional well-being. *Children and Youth Services Review*, 32, 720–731. Sun, H. (2022). Harmonious Bilingual Development: The Concept, the Significance, and the Implications. In O. S. Tan, K. K. Poon, B. A. O’Brien, & A. Rifkin-Graboi (Eds.), *Early Childhood Development and Education in Singapore* (vol. 2, pp. 261-280). Springer Singapore. Unsworth, S. (2016). Quantity and quality of language input in bilingual language development. In E. Nicoladis & S. Montanari (Eds.), *Bilingualism across the lifespan: Factors moderating language proficiency* (pp. 103-121). De Gruyter Mouton.

Bilingual children's Theory of Mind abilities: The benefits of biliteracy exposure

Jacopo Torregrossa (Goethe University Frankfurt), Christiane Bongartz (University of Cologne), Andrea Listanti (University of Cologne) and Sonja Eisenbeiß (University of Cologne).

Studies on factors that affect bilingual children's Theory-of-Mind abilities (ToM) focus on executive functions [1], metalinguistic awareness [2] and sociolinguistic variables [3]. Most of these studies rely on comparisons between monolinguals and bilinguals, which may overshadow the internal variability among bilinguals. The few studies considering only bilinguals show that balanced bilinguals in terms of language exposure have better ToM [4], which supports a sociolinguistic account of variation in bilinguals' ToM. However, these studies do not consider whether ToM benefits from certain aspects of language exposure more than others. The present study investigates the impact of different language-exposure variables on children's ToM, distinguishing between language and literacy exposure in and outside school. Furthermore, it explores which components of ToM are most affected by these language-exposure variables.

We tested 116 bilingual children (M: 9;4; 7;5-12) with 3 language-combinations, Greek-Italian in Greece (N:40), German-Italian in Germany (N:28) and Italian-English in Italy (N:48). They were recruited from 3 bilingual schools differing in the type of biliteracy implemented: in two schools, one language was the main medium of instruction (unbalanced biliteracy), corresponding to the societal minority language in one case (unbalanced_minority) and the majority one in the other (unbalanced_majority). In the third school, instruction was balanced across the two languages, considering both amount of instruction and language integration across content activities. The children were administered the "advanced" silent-films ToM-task in [5] (meant for children between 9 and 13) – featuring questions on first- and second-order false belief (1FB and 2FB) – in their dominant language, as assessed using a vocabulary-task [6]. The answers of the ToM-task were coded for explicitness, i.e., whether relevant mental states were explicitly mentioned (1) or not (0), following [5]. The parents were administered a questionnaire targeting children's language (e.g., family, outdoor activities) and literacy exposure (book reading, writing activities) outside school over time [7]. We extracted two indexes related to the minority societal language, a language- and a literacy-exposure one.

A generalized-linear-mixed-effects analysis revealed that the children were less accurate in 2FB than 1FB. The effect of false-belief type (1FB vs. 2FB) is less visible among the children exposed to balanced biliteracy, who also performed better in 2FB than their peers exposed to unbalanced biliteracy (Figure1a). We found a main effect of literacy-exposure (Figure1b) and gender (with females performing better), but no effect of vocabulary, age, language-exposure (minority language) and SES.

The results show the relevance of balanced biliteracy for the development of bilingual children's ToM in the age-range considered. Literacy exposure to the societal minority language boosts ToM across-the-board. The use of an integrated biliteracy approach in school leads to an advantage in 2FB. Balanced language use does not seem to be associated with any advantage. Given the relevance of literacy-exposure compared to language-exposure variables for ToM-development, the results are more consistent with a metalinguistic rather than sociolinguistic account of variation in bilinguals' ToM, whereby biliteracy has a general positive impact on children's metarepresentational abilities (e.g., [8]).

Motivation after ten years: learner profiles with a time dimension

Elsa Tragant (Universitat de Barcelona).

While children's first experiences with English tend to be very positive (i.e., Butler, 2017), less is known about how their motivations change over time. In this study our major focus is students' L2 motivation in their final year of secondary education. We adopt a learner-centered approach by identifying broad learner profiles among a small group of successful and less successful high school learners of English in Catalonia.

Research questions: What motivational profiles can be identified among the less and the more successful students? How do students' motivational trajectories in primary school align with their motivation in grade 10?

Learner profiling assumes heterogeneity among students and groups them according to common traits. The first study on motivational profiles was carried out by Csizer and Dörnyei (2005) who identified four learner profiles with the least and the most motivated learners showing the lowest and highest levels of integrativeness. Later studies have also found integrativeness or closely related constructs to be relevant grouping factors. L2 self-concept comes out as another important trait across several studies.

Interesting as motivational profiles are, knowledge of how these profiles have been reached over time can provide valuable insights. Nevertheless, only Zheng et al. (2020) along with the present study adopt a temporal dimension. The study expands on work by Muñoz (2017) and the ELLiE project (Enever, 2011) and focuses on 14 successful and 6 less successful learners. The primary source of data are yearly semi-structured interviews complemented with teachers' interviews, parents' questionnaires and language aptitude scores and a process of analysis similar to the retrodictive qualitative approach (Dörnyei, 2014) is followed. Our qualitative analyses yielded 2 profiles each among the least and most successful students. In defining these profiles, in addition to self-determination theory, one key element was how students experienced the distal goal of developing a good command of English. Students' language trajectories also show that the high school period is a time of important development with five and three students undergoing decremental and incremental motivational trajectories respectively. Overall, there is evidence of a general lack of stability among students' future-oriented selves and a strong interconnection between motivation, proficiency and contextual factors. Butler Y. (2017). The dynamics of motivation development among young learners of English in China. In J. Enever and E. Lindgren (Eds.), *Early language learning. Complexity and mixed methods* (pp. 167–185). Bristol: Multilingual Matters.

Csizer, K. and Dörnyei, Z. (2005). The internal structure of language learning motivation and its relationship with language choice and learning effort. *The Modern Language Journal*, 89, 19–36.

Enever, J. (Ed.) (2011). *ELLiE. Early Language Learning in Europe*. British Council.

Dörnyei, Z. (2014). Researching complex dynamic systems: 'retrodictive qualitative modelling' in the language classroom. *Language Teaching*, 47(1), 80-91.

Muñoz C. (2017). Tracing trajectories of Young learners: Ten years of school English learning. *Annual Review of Applied Linguistics*, 37, 168-184.

Zheng Y., Lu X., Ren, W. (2020). Tracking the evolution of Chinese learners' multilingual motivation through a longitudinal Q methodology. *The Modern Language Journal*, 104 (4), 781-803

English language proficiency, academic English skills and academic outcomes in higher education: Differences between students with English as first language, additional language and foreign language

Danijela Trenkic (University of York) and Selma Babayigit (UWE Bristol).

This study examined the difference in language and literacy skills between native and non-native English speakers in a university setting and their impact on academic grades. Three groups of university students in the UK were compared: international students who speak English as a foreign language (EFL, n=60), home students who speak English as an additional language (EAL, n=59), and home students who speak English as their first language (ENS, n=63). The study used psychometric tests to assess the students' English language proficiency (ELP), including vocabulary and grammar, as well as their reading and writing skills in English. The students' average grades were obtained from university records. The analysis controlled for differences in working memory and fluid intelligence, as well as socio-economic status (SES).

ELP measures. The ENS students demonstrated significantly stronger ELP skills compared to the EAL students, who, in turn, had stronger ELP skills than the EFL students. The size of the difference was substantial in vocabulary, as measured by the Peabody Picture Vocabulary Test (Dunn & Dunn, 2007) and moderate in grammar (Hartshorne et al, 2018).

Reading measures. The EFL and EAL students read at a slower pace compared to the ENS students resulting in a lower understanding of the text. This was particularly evident when the reading was timed (the Nelson-Denny Reading Test; Brown et al, 1993), where the two non-native English speaking groups (NENS) ran out of time to answer all the questions. However, when allowed to read at own pace, there were no group differences in reading comprehension. After controlling for reading speed, covariate analyses revealed that vocabulary and working memory were key factors in reading comprehension, implying that allowing extra time in assessments could enable NENS students to capitalise on their cognitive strengths.

Writing measures. The ENS students outperformed both the EAL and EFL students on writing speed (measured as a number of words in a two-minute written summarization task), text length (produced in a ten-minute essay-writing task), and essay quality. The analysis of covariates suggests that vocabulary, grammar, and processing speed in English were the major contributors to essay quality.

Academic grades. The ENS students had higher average grades than the two NENS groups, and they were more likely to be in the top 25% of academic scores, whereas the NENS students were more likely to be in the bottom 25% of academic scores. This indicates that language proficiency and academic English skills are likely to be contributing factors to the difference in academic success. In a linear regression model that combined linguistic and non-linguistic factors, the strongest and only unique predictor of average grades was essay quality.

The findings of this study have significant policy implications and suggest that addressing the English vocabulary bias and providing additional time in assessments could help to create a fairer environment for students from non-native English speaking backgrounds. Furthermore, the results indicate that a brief ten-minute essay-writing task can be used to identify students who would benefit the most from additional language support during their studies.

Language mediation and vocabulary enhancement through active subtitling - Insights into a L2-Dutch language learning project at German schools

Lukas Urbanek (Institute of Dutch Philology, University of Münster).

The effectiveness of subtitles accompanying audiovisual input for foreign language learning has been widely researched in different contexts (e.g., Montero-Perez 2022). In comparison, active subtitling by language learners to promote linguistic competences - higher education L2-English contexts excluded - is still underresearched (Lertola 2019). Even though the 'technology turn' has facilitated access to user-friendly subtitling tools for those outside the professional translation sector (Díaz-Cintas 2013), and subtitling as a learning activity is increasingly discussed as profitable for language mediation and Task-Based Language Teaching (TBLT), acceptance towards non-professional audiovisual translation in secondary school classrooms remains limited (Bobadilla-Pérez & Carballo de Santiago 2022). For this reason, taking into account existing proposals for embedding subtitling tasks into language classrooms (e.g. Talaván & Lertola 2022), an interlingual subtitling activity based on TBLT was conducted for L2-Dutch classes in German secondary schools and was carried out by teachers with more than 100 learners (without subtitling experience). Besides subtitling into German as a while-task, there were also pre- (audiovisual introduction to subtitling) and post-tasks (feedback/discussion). In addition to final questionnaires on teacher/learner perceptions of the activity, learners were presented a randomized vocabulary test with 34 target items before and after the intervention (similar to Peters & Webb 2018) and subtitling strategies used by learners were elicited following Gottlieb's (1992) classification. The results show significant vocabulary development in all participating classes and an overall positive feedback on the task from both teachers and learners. The learners' perceived need for content reduction as a characteristic of a good language mediation task is also evident in the subtitle files studied, as paraphrase and condensation were frequently used in addition to word-for-word transfer as a strategy. The suitability of subtitling for foreign language learning at schools in Germany, as assumed by Blell (2016), is thus confirmed empirically.

Montero Perez, M. (2022). Second or foreign language learning through watching audio-visual input and the role of on-screen text. *Language Teaching*, 55(2), 163–192. <https://doi.org/10.1017/S0261444821000501>. Lertola, J. (2019). Audiovisual translation in the foreign language classroom. *Research-publishing.net*. <https://doi.org/10.14705/rpnet.2019.27.9782490057252>. Bobadilla-Pérez, M., & Carballo de Santiago, R. J. (2022). Exploring Audiovisual translation as a didactic tool in the Secondary school foreign language classroom. *Porta Linguarum Revista Interuniversitaria De Didáctica De Las Lenguas Extranjeras*, 81–96. <https://doi.org/10.30827/portalin.vi.22307>. Talaván, N., & Lertola, J. (2022). Audiovisual translation as a didactic resource in foreign language education. *Encuentro*, 30, 23–39. <http://dx.doi.org/10.37536/ej.2022.30.1931>. Peters, E., & Webb, S. (2018). Incidental Vocabulary Acquisition through Viewing L2 Television and Factors that Affect Learning. *Studies in Second Language Acquisition*, 40(3), 551–577. <https://doi.org/10.1017/S0272263117000407>. Blell, G. (2016). Mehrsprachige Filme. Fächer- und sprachenübergreifende Filmarbeit. In G. Blell et al. (Eds.), *Film in den Fächern der sprachlichen Bildung*. (pp. 307–331). Schneider Verlag Hohengehren. Gottlieb, H. (1992). Subtitling - a new university discipline. In C. Dollerup & A. Loddegaard (Eds.), *Teaching Translation and Interpreting*. (pp. 161–172). Benjamins. <https://doi.org/10.1075/z.56.26got>. Díaz-Cintas, J. (2013). The technology turn in subtitling. In M. Thelen & B. Lewandowska-Tomaszczyk (Eds.), *Translation and Meaning. Part 9*. (pp. 119–132). UPM.

Can human capital and linguistic distance measures explain country TOEFL scores over the period 2005–2020?

Roeland Van Hout (Center for Language Studies, Radboud University Nijmegen) and Frans Van der Slik (Centre for Language Studies, Radboud University, Nijmegen).

International students are required to proof their mastery of English when entering higher education. In addition, learning English in a non-English spoken environment enhances career and income opportunities. Consequently, the English language testing industry has become a billion dollar enterprise (Azoth Analytics, 2019). Interestingly, the Educational Testing Service in Princeton, US, has reported systematic differences over the time span 2005–2020 on their Test of English as a Foreign Language, henceforth TOEFL. These differences pertain to countries/regions and native languages. In our study, we address these differences in relation to two important groups of indices, one on human capital, at the country level, and one on linguistic distance, based on the L1s. Van Hout & Van der Slik (submitted) have found that both the Human Capital Index (HCI; see Kraay, 2019) and linguistic distance measures explain a substantial part of the variation in another L2 English proficiency test, the Education First's SET scores (EF, 2021). The HCI comprises various components: survival, health, and the number of years of school a child can attain given prevailing enrolment rates. "(...) the HCI measures the expected future human capital of a child born today, given current education and health outcomes for the young." (Kraay, 2019:4). In previous studies, we successfully applied linguistic distance measures to predict language learners Dutch L2 scores (e.g., Van der Slik, 2010, Schepens et al., 2022), yet we applied these measures successfully to English L2 scores, too (Van Hout & Van der Slik, submitted). In this presentation, we focus on answering the question to what extent (1) human capital and other country-related distinctions, and (2) linguistic distance relate to TOEFL scores and how they interact over time. We used mixed regression analyses (LMER) and generalized additive mixed models (GAMM) to reveal the patterns of relationships between the predictors, including the effect of introducing a time lag between a changing human capital index and its future effect on English L2 proficiency scores. Our main conclusions are that both human capital, as measured by the gross enrolment ratio of secondary schooling, and linguistic distance, as measured by lexical, morphological, and phonological distance, substantially explain the variation in TOEFL scores over a period of 16 years, including a clear overall time lag of seven years between the gross enrollment ratio and the language proficiency scores.

Azoth Analytics (2019). Global English proficiency test market: Insights, trends and forecast 2019-2024. Education First. (2021). EF EPI EF English Proficiency Index. A ranking of 112 countries and regions by English skills. <https://www.ef.nl/epi/>. Kraay, A. (2019). The World Bank human capital index: A Guide. The World Bank Research Observer, 134, 1-33. Schepens, J., Van der Slik, F., & Van Hout, R. (2022). Linguistic dissimilarity increases age-related decline in adult language learning. *Studies in Second Language Acquisition*, <https://doi.org/10.1017/S0272263122000067> Van der Slik, F. (2010). Acquisition of Dutch as a second language. *Studies in Second Language Acquisition*, 32(3), 401–432. <https://doi.org/10.1017/S0272263110000021>

The vocabulary demands of English and French L2 textbooks: A cross-lingual corpus study

Amaury Van Parys, Vanessa De Wilde, Lieve Macken and Maribel Montero Perez (Ghent University)

Studies have determined the vocabulary demands of a variety of input types (e.g., novels, audio-visual media) by calculating lexical profiles (e.g., Nation, 2006), i.e., categorisations of the vocabulary into word frequency levels. This allows researchers to determine the required vocabulary knowledge to reach 95% and 98% understanding (or coverage) of the words in the input, which are believed to be necessary to reach minimal and optimal comprehension of a text's contents, respectively (Laufer & Ravenhorst-Kalovski, 2010). Though textbooks are a vital source of input in the L2 classroom, lexical profiling research into L2 textbooks is limited. Moreover, existing studies tend to focus exclusively on English. Therefore, a corpus consisting of English and French L2 textbook reading materials was compiled (ca. 300,000 words per L2) in order to investigate (RQ1) what the vocabulary demands are, (RQ2) how these demands evolve across all six years of secondary education, and (RQ3) how target language influences the (evolution of) demands.

Typically, lexical profiling research relies on word families as lexical unit, but recent research has shown that these may overestimate the vocabulary knowledge of learners who struggle with morphology (e.g., Brown et al., 2022). In light of this criticism and bearing in mind the higher morphological demands of French, we opted for flemmas and individual word types instead and created a custom Python script to ensure cross-language comparability. The lexical profiles were supplemented with indices of lexical diversity and density.

Results showed that knowledge of the 15,000 most frequent flemmas was required to reach 98% coverage of the first grade English reading materials, as opposed to 11,000 for French, despite French instruction already starting in primary school and English instruction not until secondary school. From there, a gradual and moderately systematic increase in demands across grade levels was observed in both the English and French segments of the corpus. However, some grade levels did not differ in terms of lexical profiles. In addition, English consistently required more vocabulary knowledge than French (ca. 5,000 flemmas more in each grade level). Overall, these findings suggest that publishers take into account the considerably higher English vocabulary knowledge that has been observed in adolescents as a consequence of out-of-school exposure (Peters et al., 2019).

References:

- Brown, D., Stoeckel, T., Mclean, S., & Stewart, J. (2022). The Most Appropriate Lexical Unit for L2 Vocabulary Research and Pedagogy: A Brief Review of the Evidence. *Applied Linguistics*, 43(3), 596-602. <https://doi.org/10.1093/applin/amaa061>
- Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size. *Reading in a Foreign Language*, 22(1), 15-30.
- Nation, I. S. P. (2006). How Large a Vocabulary is Needed For Reading and Listening? *The Canadian Modern Language Review*, 63(1), 59-82. <https://doi.org/10.3138/cmlr.63.1.59>
- Peters, E., Noreillie, A., Heylen, K., Bult, B., & Desmet, P. (2019). The Impact of Instruction and Out-of-School Exposure to Foreign Language Input on Learners' Vocabulary Knowledge in Two Languages. *Language Learning*, 69(3), 747-782. <https://doi.org/10.1111/lang.12351>

Does exposure to multiple languages influence children’s pragmatic abilities? A meta-analysis.

Elise van Wonderen (Amsterdam Center for Language and Communication), Kimberley Mulder (Amsterdam Center for Language and Communication), Judith Rispen (Amsterdam Center for Language and Communication) and Josje Verhagen (Amsterdam Center for Language and Communication).

Despite an often reported delay in the vocabulary and grammar development of multilingual children compared to monolingual peers, recent studies have found that multilingual children might outperform monolinguals in their pragmatic abilities (e.g., Fan et al., 2015; Siegal et al., 2009; Yow & Markman, 2011). Specifically, findings indicate that multilinguals may be more attentive to the social and physical context of conversations, either as a compensatory strategy to overcome any gaps in vocabulary or grammar knowledge, or as a result of constantly monitoring their interlocutors’ language abilities. Current evidence is mixed, however, and it remains unclear whether there are indeed differences in pragmatic skills between monolingual and multilingual children. We therefore conducted a meta-analysis to establish whether, overall, multilingual children outperform monolingual children in pragmatic abilities, and if so, whether this varies across pragmatic abilities.

Based on a systematic database search, we included 109 effect sizes from 29 papers, representing data from 1,115 unique multilingual and 931 unique monolingual children between the ages of one and twelve years. The pragmatic abilities investigated in these studies fell into four broad categories: sensitivity to non-verbal information (14 studies), understanding non-literal language (5 studies), informativeness in referencing (6 studies), and other (5 studies). We used a correlated and hierarchical effects model and robust variance estimation to account for dependent effect sizes (Pustejovsky & Tipton, 2022). Results showed that multilingual children did not outperform monolingual children across pragmatic abilities, Hedges’ $g = 0.09$, 95% CI [-0.16, 0.33], nor in any of the separate pragmatic categories (sensitivity to non-verbal information: $g = 0.16$ [-0.15, 0.48]; understanding of non-literal language: $g = 0.32$ [-0.69, 1.34]; level of informativeness in referencing: $g = -0.02$ [-0.75, 0.72]). However, significant residual heterogeneity in effect sizes indicated that there was systematic variance in effect sizes within each pragmatic category. Future research could explore further moderators of pragmatic differences between multilingual and monolingual children (e.g., language proficiency), as well as employ research designs that can distinguish between hypotheses on why such differences may arise.

References: Fan, S. P., Liberman, Z., Keysar, B., & Kinzler, K. D. (2015). The exposure advantage: Early exposure to a multilingual environment promotes effective communication. *Psychological Science*, 26(7), 1090–1097. Pustejovsky, J. E., & Tipton, E. (2022). Meta-analysis with Robust Variance Estimation: Expanding the Range of Working Models. *Prevention Science*, 23(3), 425–438. Siegal, M., Iozzi, L., & Surian, L. (2009). Bilingualism and conversational understanding in young children. *Cognition*, 110(1), 115–122. Yow, W. Q., & Markman, E. M. (2011). Young bilingual children’s heightened sensitivity to referential cues. *Journal of Cognition and Development*, 12(1), 12–31.

Exploring Inter-Individual Differences in L2 Receptive Development among Adults with Diverse Educational Backgrounds

Marieke Vanbuel (UGent) and Bart Deygers (Ghent University).

This study investigates variations in the receptive language development of adult second language (L2) learners with diverse educational backgrounds. L2 listening and reading are considered distinct but highly interrelated skills (Jeon & Yamashita, 2014). However, most studies on receptive gains are based on samples of adult learners with a tertiary education background, potentially limiting their generalizability (Andringa & Godfroid, 2020). Since first language (L1) reading skills influence L2 reading skills to some extent (Yamashita & Shiotsu, 2017), findings based on highly educated respondents may not translate to a more educationally diverse sample. What research is available on low-educated adult L2 learners suggests that educational background is an important predictor of L2 proficiency (Kurvers, 2015) and, to a lesser extent, language gains (Deygers & Vanbuel, 2022). Often, however, large differences in gains between students with similar educational backgrounds remain unexplained, even when other background variables are taken into account (Author, 2022). These observations lead to an underexplored research question that is vital to the generalizability of receptive skills research: Can we identify patterns of L2 receptive development across L2 adult learners with diverse educational backgrounds? This study employs multilevel regression, hierarchical cluster analysis and K-means clustering (Crowther et al., 2021) to identify groups of adult L2 learners based on their L2 listening and reading gains (i.e., random intercept and slope) on a standardized test. While attending L2 classes, an educationally diverse sample of 551 adult learners of Dutch as a second language took the same computer-adaptive listening and reading tests twice over a four-month period. A background questionnaire was administered to collect demographic data. Preliminary analysis reveals two clusters, indicating that L2 listening and reading skills do not necessarily develop similarly across the whole group. Cross-cluster variation is primarily explained by track type, educational background and knowledge of English. The results confirm the role of educational background and linguistic distance in L2 receptive skills development, underline the importance of sampling beyond higher educated L2 learners, and fuel hypotheses regarding why educational background may impact not only receptive test scores, but also receptive score gains.

Andringa, S., & Godfroid, A. (2020). Sampling Bias and the Problem of Generalizability in Applied Linguistics. *Annual Review of Applied Linguistics*, 40, 134–142. Crowther, D., Kim, S., Lee, J., Lim, J., & Loewen, S. (2021). Methodological Synthesis of Cluster Analysis in Second Language Research. *Language Learning*, 71(1), 99–130. Deygers, B., & Vanbuel, M. (2022). Gauging the impact of literacy and educational background on receptive vocabulary test scores. *Language Testing*, 39(2), 191–211. Huettig, F., & Mishra, R. K. (2014). How Literacy Acquisition Affects the Illiterate Mind—A Critical Examination of Theories and Evidence. *Language and Linguistics Compass*, 8(10), 401–427. Jeon, E. H., & Yamashita, J. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language Learning*, 64(1), 160–212. Kurvers, J. (2015). Emerging literacy in adult second-language learners: A synthesis of research findings in the Netherlands. *Writing Systems Research*, 7(1), 58–78. Yamashita, J., & Shiotsu, T. (2017). Comprehension and Knowledge Components That Predict L2 Reading: A Latent-Trait Approach. *Applied Linguistics*, 38(1), 43–67.

Fluency in task-based peer interaction: Exploring the effects of interactional instruction on primary school EFL learners

Alexandra Vraciu (Universitat de Lleida) and Elisabet Pladevall Ballester (Universitat Autònoma de Barcelona).

Young Learners' (YLS) oral fluency has been generally analysed through monologic tasks (Enever, 2011). Nonetheless, L2 performance in interaction also needs to be approached from a broader perspective, taking into account the fact that the boundaries of individual performance are blurred in dialogue as learners also have collaborative resources to overcome breakdowns and maintain fluent speech (Sato, 2014). The dialogic condition of the tasks significantly affects learners' individual oral fluency and, at the same time, individual fluency resources play an important role in maintaining interactional fluency (Michel, Kuiken and Vedder, 2007; Peltonen, 2017). Fluency in interaction is to date unexplored with YLS, even more so as a result of a pedagogical intervention fostering task-based peer interaction in a primary education context.

The present study analyses the impact of task-based peer interaction and interactional with/without grammar instruction on the individual and interactional fluency of primary school learners of EFL. Two groups of 4th grade Catalan/Spanish bilingual children aged 9-10 (n=25 per group) participated in an 8-week pedagogical intervention (50 minutes per week). A third group (n= 25) acted as control and did not receive any treatment. The two intervention groups did 30 minutes of focused task-based peer interaction (both collaborative and information-gap tasks) every week. Before taking part in the tasks and for 20 minutes per week, one of the groups was given explicit instruction on interactional strategies (i.e conversational adjustments and corrective feedback strategies) and the second group received explicit grammatical (present continuous) and interactional instruction. Oral production data was collected before and after the intervention by means of a focused collaborative picture narrative task. Individual fluency was measured through number of turns, number of L2 words, number of L2 words per turn and the use of self-repetitions both in the L1 and the L2. Following Peltonen (2017), interactional fluency was measured through the use of other repetitions and completions both in the L2 and the L1.

Results suggest that the pedagogical intervention had a moderate impact on the two experimental groups, while the control group experienced no changes. More specifically, interactional fluency development was modest across both groups whereas significant gains were observed in individual fluency, particularly in the interaction group, with a decrease in the number of turns and self-repetitions and an increase in L2 words per turn. The group who received interactional and explicit grammar instruction appears to be more dysfluent than the other experimental group at the end of the intervention in terms of individual measures, with an increase in the number of turns and almost no variation in terms of self-repetitions and L2 words per turn, yet outperforms it with regard to dialogue maintenance strategies, in particular other repetitions. Combined allocation of communicative effort to focus on form and interaction may have affected individual and interactional dimensions of fluency in the case of this group. We discuss our findings with regard to the models L2 oral production in SLA and YL pedagogies.

Foreign language aptitude, learning environments, and motivational self-determination in Austrian primary school learners. Reflections on a pilot study

Thomas Wagner (University College of Education Upper Austria) and Claudia Resch (University College of Education Salzburg - Österreichisches Zentrum für Begabtenförderung und Begabungsforschung (ÖZBF)).

Gifted learners often require special materials and learning environments in order to develop and maintain motivation and perform according to their potentials. This also applies to EFL learning in primary school. This paper reports an Austrian pilot study on the role of foreign language aptitude and learning environments in co-determining self-regulatory, material-related motivation (Deci & Ryan, 2002). A cross-sectional within-subjects aptitude-treatment interaction study (Vatz et al., 2013) examined, in how far gifted education materials and learning architectures are motivational for young EFL learners with high foreign language aptitude. Experimental treatment consisted of an open and self-regulatory gifted education intervention. For this treatment, four primary EFL experts researched and rated 24 EFL materials based on five criteria from EFL and aptitude research as well as gifted education. Based on inter-rater reliability, 12 materials were selected and edited for experimental treatment. The control consisted of a standardised teacher-centred EFL lesson as traditionally delivered in Austrian primary classrooms. For the empirical pilot study, foreign language aptitude was assessed in 57 pupils from three primary schools, using slightly adapted LLAMA-F and LLAMA-B tests (Meara & Rogers, 2019). All learners were exposed to the regular control treatment, while 15 pseudo-randomly selected high and low aptitude learners also engaged with the self-regulatory gifted education intervention. The learners' motivation was assessed using the SRQ-A scale (Müller et al., 2007) after each intervention, as well as guided interviews after experimental treatment. In addition, the self-regulatory learning intervention was videotaped and coded for pupils' engagement with the materials. We hypothesised a disordinal relationship between aptitude and teaching intervention with regard to motivation; learners with high aptitude scores were expected to be motivated by the open learning intervention and rather demotivated by the traditional one, while learners with low aptitude would show an inverse pattern. While preliminary quantitative results show an approximation towards such an aptitude by treatment interaction, data-triangulation revealed inconsistencies. We discuss these results and outline potential implications for EFL primary teaching and gifted education. Moreover, we reflect on this pilot study in terms of research method, intervention design, and motivational measures, outlining avenues for further research within and beyond traditional ATI approaches.

Deci, E. L., & Ryan, R. M. (2002). Overview of self-determination theory: An organismic dialectical perspective. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 3–33). Rochester: Rochester UP. Meara, P. M., & Rogers, V. E. (2019). *The LLAMA tests v3*. Cardiff: Lognostics. Müller, F. H., Hanfstingl, B., & Andreitz, I. (2007). *Skalen zur motivationalen Regulation beim Lernen von Schülerinnen und Schülern: Adaptierte und ergänzte Version des Academic Self-Regulation Questionnaire (SRQ-A) nach Ryan & Connell*. Wissenschaftliche Beiträge aus dem Institut für Unterrichts- und Schulentwicklung. Klagenfurt: Alpen-Adria-University. Vatz, K. M., Tare, M., Jackson, S. R., & Doughty, C. J. (2013). Aptitude-treatment interaction studies in second language acquisition. In G. Granena & M. H. Long (Eds.), *Sensitive periods, language aptitude, and ultimate L2 attainment* (pp. 272–292). Amsterdam: John Benjamins.

The Optimal Lag for Intentional and Incidental Language Learning

Neil Walker (University of Central Lancashire), Padraic Monaghan (University of Lancaster), Sarah Chadwick (University of Lancaster) and Patrick Rebuschat (University of Lancaster).

Several recent studies have investigated the optimal lag between study sessions for second language learning (e.g. Kasprowicz et al., 2019; Serfaty & Serrano, 2022). However, there is a lack of consensus as to whether shorter or longer lags are preferable. One theory of the underlying mechanism of the lag effect, the Reminding Account (Benjamin & Tullis, 2010), predicts that the optimal lag will be one that provides the learner with a desirable level of difficulty.

In this study (pre-registered on OSF), we tested whether the optimal lag for learning form-meaning connections (four artificial pre-nominal determiners, *gi*, *ro*, *ul* and *ne*, expressing both animacy and distance; see Williams, 2005 for a fuller description of the target mappings) was affected by several factors that might influence the desirable difficulty of learning: the type of learning (intentional or incidental), the newness of the tested item (presented previously during training or new items that required a generalisation of the rules), and an individual memory difference (declarative memory).

Two hundred participants were randomly assigned to one of five groups which differed only in the lag between the three exposure blocks of this study (0-day, 1-day, 2-day, 4-day and 7-day lag). In each exposure block, participants were intentionally taught one aspect of the determiner use (either animacy or distance, counterbalanced across participants), while the other aspect remained hidden and had to be acquired incidentally. Next, participants were presented with 48 sentences, each containing one of the artificial determiners followed by a noun and were asked to make a meaning judgment about the intentional condition. After the three exposure blocks presented at different lags, participants were finally tested on their knowledge of both the animacy and distance aspects of the determiners in a 96-item two-way-forced-choice test administered 35 days after the last exposure block. Declarative memory was tested via the MLAT-V.

Results revealed a significant learning advantage for three of the lag groups (1-day, 4-day and 7-day lag groups but not the 2-day lag group) over the massed (0-day lag) group. However, there was no significant difference in learning between the spaced groups (1-day, 2-day, 4-day and 7-day lag groups), suggesting that there was no one optimal lag. Importantly, lag was shown to have differing effects on intentional and incidental learning, with only 1-day and 2-day lag groups showing evidence of learning for the incidental aspect. In exploratory analysis, we found no difference in test scores between previously trained items and new items that required a generalisation of rules. Declarative memory, contrary to our predictions, did not interact with lag, suggesting that those with stronger declarative memory do not require a wider optimal lag. The results indicate that more than one mechanism may be at work in the lag effect.

Benjamin, A. S., & Tullis, J. (2010). *Cognitive Psychology*, 61(3), 228-247. Kasprowicz, R. E., Marsden, E., & Sephton, N. (2019). *The Modern Language Journal*, 103(3), 580-606. Serfaty, J., & Serrano, R. (2022). *Applied Psycholinguistics*, 1-38. Williams, J. N. (2005). *Studies in Second Language Acquisition*, 27(2), 269-304.

What kind of skill-integrated language activities are effective in improving English proficiency?

Yoichi Watari (Chukyo University), Mitsuhiro Morita (Hiroshima City University) and Atsushi Mizumoto (Kansai University).

This study aims to clarify which aspects of English language proficiency can be improved by implementing skills-integrated language activities (SILAs) in secondary school classrooms. At the secondary school level in Japan, the governmental educational policy mandates the enhancement of SILAs that effectively leverage all four skills. In the classroom environment, it is more ecologically valid to examine the processes of second language acquisition on the premise of SILAs, as opposed to those that center on each individual skill. However, studies examining the effects of SILAs as independent variables on language proficiency are scarce, and the effects of long-term implementation in the classroom setting are not always clear. Shintani, Li, and Ellis (2013) conducted a meta-analysis comparing the effects of comprehension-based instruction (CBI) and production-based instruction (PBI) on grammatical knowledge, and found that (1) for receptive knowledge, CBI had a greater effect than PBI, but the difference diminished in the delayed tests, and (2) for productive knowledge, both had similar effects in the short term, but PBI was more effective in the delayed tests. Subsequently, while many studies have been conducted to examine the development of specific skills under various conditions and designs or to identify individual differences that influence the results (Pawlak, 2017), it remains unresolved which specific skills can be dependably enhanced through SILAs, where both comprehension and production are cyclically required. In this study, we collected longitudinal data from two secondary schools (N = 1,016) over a span of two years, 2018 and 2019. The students annually took a private-sector English language test as their proficiency measure and answered a questionnaire about the classroom activities they had experienced in lessons as well as their learning habits and English-related experiences outside the classroom. Overall, the learners demonstrated $d = 0.77$ growth in listening, $d = 1.04$ in speaking, $d = 0.59$ in reading, and $d = 0.73$ in writing over the course of one year. A generalized linear mixed model analysis reveals that writing activities based on reading and/or listening will be effective for improving reading, writing, and speaking skills. Additionally, the result shows that prepared speech activities would lead to improvement in writing skills. On the other hand, debate and discussion activities were not shown to be effective in improving any of their skills. We discuss our findings in relation to the theoretical perspectives from SILAs. In addition, we will expound upon the methodological considerations inherent in collecting longitudinal data within the classroom setting.

Pawlak, M. (2017). Individual difference variables as mediating influences on success or failure in form-focused instruction. In E. Piechurska-Kuciel, E. Szymańska-Czaplak, & M. Szyszka (Eds), *At the crossroads: Challenges of foreign language learning* (pp. 75–92). Springer. https://doi.org/10.1007/978-3-319-55155-5_5

Shintani, N., Li, S. & Ellis, R. (2013). Comprehension-based versus production-based grammar instruction: A meta-analysis of comparative studies. *Language Learning*, 63(2), 296–329. <https://doi.org/10.1111/lang.12001>

Student-Made versus Pre-Made Word Cards for EFL Vocabulary Learning: Learning Outcomes and Learner Perceptions

Darrell Wilkinson (Tampere University).

Using word cards for deliberate vocabulary learning is widely supported by theoretical underpinnings, such as convenience, and affordance for active and passive recall (Nation, 2013). Furthermore, earlier research (Mondria & Mondria-de Vries, 1994; Waring, 1997) and a growing body of recent studies support word cards (Nakata, 2020; Reynolds & Shih, 2019; Wilkinson, 2017). However, whether students should make the cards or study from pre-made cards remains unknown (Lei & Reynolds, 2022). Given the abundance of pre-made cards (mainly digital) and considering the time-consuming nature of making the cards along with the lack of intrinsic motivation for the process (Reynolds & Shih, 2019; Wilkinson, 2017), this is an important issue.

Therefore, the mixed-methods research presented aims to address this gap by comparing the learning outcomes associated with (a) making word cards before repeatedly studying from them with (b) spending the same amount of time-on-task studying from pre-made cards.

The experiment adopted a counter-balanced pretest-posttest design with immediate and delayed active and passive recall tests. 85 pre-intermediate university students studied 15 words under each condition, with 15 additional control items used instead of control group. Outcomes were compared using repeated-measures MANOVA and ANOVA. Also, qualitative survey, interview, and observation data was analysed to examine students' perceptions of each method.

Findings indicated both methods were effective for developing word knowledge, but in the long-term, there was no statistically significant difference. Immediate post-tests indicated (when time-on-task and other confounding variables are controlled) pre-made cards were superior for both measures, suggesting for short-term memory, additional repetitions are more important than extra involvement load (Laufer & Hulstijn, 2001). However, after eight weeks of repeated study under natural learning conditions, this did not hold true. Qualitative data indicated students generally preferred pre-made cards, and motivation and learner preferences played a greater role than the card method.

In conclusion, the results provide further support for word card usage, but indicate the choice between self-made or pre-made should be based on learner preferences.

Use of L1 phonotactics in initial foreign-language speech segmentation

Sophia Wulfert (English Linguistics, TU Braunschweig), Katie Von Holzen (English Linguistics, TU Braunschweig), Marie Schnieders (English Linguistics, TU Braunschweig) and Holger Hopp (English Linguistics, TU Braunschweig).

One of the initial key skills in foreign language (FL) learning is to extract individual words from foreign speech. Phonotactics provides sublexical cues to the segmentation of the speech stream, but is language-specific (e.g., Mattys & Jusczyk, 2001; Weber & Cutler, 2006). We investigate the degree to which children initially apply the phonotactics of their first language (L1) to segmentation in a FL. In particular, we examine how differences in L1 knowledge and (meta-)linguistic skills among children affect the application of L1 phonotactics. In this talk, we explore how German-speaking children use phonotactic cues from their L1 when encountering English speech before they begin English instruction. In a listening task, they heard English sentences followed by an isolated probe pseudoword (either contained in the previous sentence or not) and indicated via button press whether they had heard the probe word in the sentence as a target. Target pseudowords occurred in the sentences either in a context which provided a clear L1-German phonotactic cue to a word boundary (e.g., lv; dal_vouchen) or an ambiguous cue (kv; dack_vouchen). Probes were either short (vouchen), or long (kvouchen/lvouchen) pseudoword candidates. If children use L1 phonotactic cues, they should easily detect pseudowords in clear boundary contexts. We therefore hypothesize that detection rates for targets will be highest for short probe words in clear boundary contexts (e.g., l_vouchen) and lowest for long probe words that contain a clear boundary cue within them (e.g., lvouchen). Words in ambiguous boundary cue contexts ((k)vouchen) are expected to show intermediate detection rates. In the talk, we will report results from 96 children aged 6 to 9 to be collected by July. So far, we have analyzed results from 27 6- to 7-year-olds. We examined the interaction of Boundary Type (clear, ambiguous) and Length (short, long) on target accuracy using a general-linear mixed model. The main effect of Length was significant ($\beta = -0.23$, $SE = 0.11$, $p = .03$), indicating that short words were detected more accurately than long words. Target accuracy was greatest for short target words with clear boundary cues (EMM = 0.54; Figure 1). This pattern demonstrates a tendency for young children to apply L1 phonotactic cues to segment speech in a foreign language. In our final, fully powered sample, we will examine how individual differences in age, phonological awareness and L1 knowledge affect the segmentation of FL speech. On top of providing insights into initial FL learning abilities, these results can inform educational policy-decisions about the optimal time for starting early FL instruction.

Direct and Indirect Contributions of Three Aspects of Morphological Knowledge to L2 Reading Comprehension

Junko Yamashita (Nagoya University) and Kunihiro Kusanagi (Prefectural University of Hiroshima).

Research into morphological knowledge (MK) as a component of reading comprehension (RC) started later than other lower-level components such as phonological or orthographic skills. However, the rapid growth of research has reached a consensus on the importance of MK for reading development. What we need now are finer-grained approaches examining different aspects of MK and the mechanism through which MK supports RC. Several theoretical frameworks give suggestions for such studies. The Reading Systems Framework (Perfetti & Stafura, 2014) postulates direct and indirect (via lexicon/vocabulary) pathways through which MK contributes to RC. The Morphological Pathways Framework (Levesque et al., 2021), which is based on the RSF, added the role of morphological decoding that involves morpho-orthographic and morpho-semantic analysis. Morphological decoding also represents an indirect pathway from MK to RC through the lexicon. Regarding the different aspects of MK, though a consensus has yet to arrive about the dimensionality of MK (Goodwin et al., 2017; 2021), L2 vocabulary researchers constructed a test (Word Part Levels Test) that measures three aspects of MK: Form (orthographic form), Meaning, and Use (syntactic properties, i.e., parts of speech) (Sasao & Webb, 2017). Its computerized adaptive testing version (CAT-WPLT) is also available (Mizumoto et al., 2019). Building upon these theoretical and methodological foundations, the present study examined how three aspects of MK contribute to RC by testing direct and indirect pathways that link MK to RC with vocabulary knowledge as a mediator. The participants were 211 L1-Japanese university students. Their English MK was measured by the CAT-WPLT. RC ability and vocabulary knowledge (breadth) were measured by the VELC (Visualizing English Language Competency) test, a standardized test of English-language skills for Japanese university students (Shizuka & Mochizuki, 2014). After checking the psychometric properties of the tests and correlations among variables, a series of path analyses was employed to find the best model to capture the relationship between RC and four predictors (Form, Meaning, Use, and Vocabulary). The final model showed a good fit to the data, $\chi^2(2) = 1.648$, $p = .439$, GFI = .997, CFI = 1.000, RMSEA = .000 [.000 - .129], and 25% of the variance in the RC test score was explained. Three direct effects on RC were significant: Vocabulary ($\beta = .144$), Meaning (.215), and Use (.317), and two indirect effects on RC via Vocabulary were significant: Meaning (.024) and Use (.022), which resulted in total effects of Meaning and Use being .239 and .339. In contrast, there was no significant effect of Form either on vocabulary or RC. In sum, although the form aspect of MK did not show any effect, semantic and syntactic aspects demonstrated influences on RC, both directly and indirectly, with the larger total impact of the syntactic aspect. Overall, this study endorsed the criticality of a finer-grained approach shedding light on what and how MK supports RC. The findings are discussed, considering the literature and potential limitations of this study, to deepen our understanding of the role of MK in L2 RC and suggest future directions.

Enhancing L2 Pronunciation through Embodied and Non-Embodied Shadowing: The Role of Speech Latency

Ting Yao (Universitat Pompeu Fabra), Patrick Louis Rohrer (Universitat Pompeu Fabra), Sharon Gutierrez Metelli (Universitat Pompeu Fabra), Valeria Tamburini (Universitat Pompeu Fabra), Bianca Regina Held (Universitat Pompeu Fabra) and Pilar Prieto (Universitat Pompeu Fabra).

Previous studies have reported that the shadowing technique, which involves immediately vocalizing speech stimuli, (see Lambert, 1992, p. 266) may be an effective tool to improve L2 English learners' pronunciation (e.g., Foote & McDonough, 2017; Sugiarto et al. 2020). However, little is known about the effectiveness of embodied shadowing (i.e., the immediate vocal and physical reproduction of speech and accompanying rhythmic hand gesture stimuli) and the role of the individual speech latency during training (i.e., the time-lapse between the model speech and the learners' reproduction, see Oki, 2010). The present study aimed to examine the effect of 6 sessions of embodied shadowing training on improving L2 English pronunciation (in terms of comprehensibility, accentedness, and fluency) in comparison with non-embodied shadowing training. Fifty-four learners of English were randomly assigned to either an embodied or non-embodied shadowing group. The embodied group reproduced the model's beat gestures while shadowing speech. The other group only shadowed the speech without seeing or having any specific instruction regarding gesture production. Additionally, the speech latency (or temporal distance) between the onsets of pitch-accented syllables produced by the native speaker and the corresponding learner's onsets during the training was coded with Praat (Boersma & Weenink, 2022). Pronunciation was assessed through oral reading tasks in the pre-test and post-test. Results revealed that, in contrast to the embodied group, the non-embodied group significantly improved their speech accentedness and fluency scores (see Figure 1) after training. Further analyses showed that speech latency during training was a key predictor of the learners' gains in comprehensibility and fluency (see Figure 2) in both conditions. The results suggest that the value of the shadowing technique for pronunciation learning may depend directly on the learners' shadowing performance during training, with shorter latencies associated with greater improvement. The null effects observed in the embodied shadowing training could be attributed to the increase in cognitive load caused by the incorporation of complex visual information, which might distract learners' attention from stimuli pronunciation. Overall, this study provides pedagogical implications, suggesting the gradual introduction of both non-embodied and embodied shadowing techniques over a longer period that allows learners to reduce speech latencies and further promote pronunciation learning.

References [1] Boersma, P. & Weenink, D. (2022). Praat: doing phonetics by computer [Computer program]. Version 6.2.12, retrieved 17 April 2022 from <http://www.praat.org/> [2] Foote, J. A., & McDonough, K. (2017). Using shadowing with mobile technology to improve L2 pronunciation. *Journal of Second Language Pronunciation*, 3(1), 34-56. [3] Lambert S (1992) Shadowing. *Meta* 37(2): 263–73. [4] Sugiarto, R., Prihantoro, P., & Edy, S. (2020). The Impact of Shadowing Technique on Tertiary Students' English Pronunciation. *Linguists: Journal of Linguistics and Language Teaching*, 6(1), 114-125.

The Asymmetry Hypothesis in Motion Event Descriptions: Evidence from the Learners of English, Hungarian and Japanese

Yuko Yoshinari (Gifu University), Kiyoko Eguchi (Osaka University), Miho Mano (Osaka University) and Yo Matsumoto (National Institute for Japanese Language and Linguistics).

This study examines the “Asymmetry Hypothesis” (Yoshinari et al. 2021) concerning the difficulty in learning the path-encoding patterns in motion event descriptions, for which two major typological types are known (Talmy 1991). The difficulty of learning a different typological type has been pointed out (e.g. Cadierno & Robinson 2009), but Yoshinari et al. (2021) further argue that one type of language is intrinsically more difficult than the other for the speakers of a different type: languages that often encode path in the main verb (the head) (so-called verb-framed/head path-coding (HPC) languages; e.g. Japanese) are more difficult to learn than languages that encode path in “satellites” and other “head-external” elements (e.g. adpositions, particles) (HExtPC languages; e.g. English and Hungarian). This is based on the observation that so-called HPC languages in fact use the mixed patterns of using the main verb and/or head-external elements to describe path depending on the particular events described, presenting difficulty of differentiation for learners, while HExtPC languages consistently use head-external elements for various events (although particular elements used differ among languages). To test this asymmetry hypothesis, we analyzed experimentally elicited data on three L1 (Japanese, English, and Hungarian) and four B1-level learner groups: English used by Japanese learners (E-L2(j)), Hungarian used by Japanese learners (H-L2(j)), and Japanese by English- and Hungarian-speaking learners (J-L2(e), J-L2(h)). The hypothesis predicts that responses in E-L2(j) and H-L2(j) are more L1-like than J-L2(e) and J-L2(h). The participants verbally described various clips of various motion events and responses for 24 clips representing the TO.IN path in self-motion (e.g. running inside), caused-motion (e.g. kicking inside), and visual motion events (looking inside) were analyzed. We first identified preferred L1 path-encoding patterns covering more than 80% of the responses for each clip. We then analyzed the L2 responses to find the ratios of the use of those “L1-like expressions” for each clip. The overall results match the predictions: J-L2(e) and J-L2(h) show significantly fewer L1-like expressions (19.4% and 20.6%, respectively) than E-L2(j) and H-L2(j) (39.4% and 28.1%, respectively). The tendency is especially clear in visual motion, where the L1-like patterns are different from the L1 patterns for self- and caused motion. However, a closer look at each clip indicates that there are certain cases where the prediction does not hold: H-L2(j) speakers used fewer L1-like expressions for certain caused-motion scenes than J-L2(h) speakers. For those clips, H-L2(j) speakers fail to employ the L1 pattern of using both a “preverb” (preverbal particle) and a case marker. Learning to use an unfamiliar preverb is the cause of the difficulty. These results suggest that the asymmetry hypothesis needs to be modified. The relative easiness of HExtPC languages for the speakers of HPC languages can be counterbalanced by the unfamiliarity of head-external elements.

Cadierno, T., & Robinson, P. 2009. Language typology, task complexity and the development of L2 lexicalization patterns. *ARCL* 7, 245-276. Talmy, L. 1991. Path to realization. *BLS* 17, 480-519. Yoshinari, Y. et al. 2021. *Idoohyoogenno ruikeironto dainigengosyuutoku* [Typology of motion-event descriptions and second-language acquisition]. Kurosio.

Cross-language orthographic associations in proficient Chinese-Japanese bilinguals in a lexical decision task with cross-language priming

Xuehan Zhao (Department of Linguistics, Graduate school of Arts and Letters, Tohoku University, Japan), Kexin Xiong (Frontier Research Institute for Interdisciplinary Sciences, Tohoku University, Japan) and Sachiko Kiyama (Department of Linguistics, Graduate school of Arts and Letters, Tohoku University, Japan).

The Revised Hierarchical Model (Kroll & Stewart, 1994) of the bilingual mental lexicon posits that proficient bilinguals can access concepts directly from the lexical form of the second language (L2) without first language (L1) mediation. It remains controversial whether this L1-L2 association depends on the degree of orthographic overlap between languages. In Japanese, certain words are written in logographic characters like traditional Chinese characters (CC: Kanji/Hanzi), some in syllabic Japanese characters (JC: Kana), and some in both CC and JC. For Chinese-Japanese bilinguals, JC words are reported to be more difficult to acquire than CC words due to the less orthographic L1-L2 association. It is hypothesized that even proficient L2-Japanese speakers whose L1 is Chinese access concepts of L2-Japanese words mediated by L1-Chinese when the words are written in JC, while they access concepts directly from L2-lexical representations in CC. We conducted a lexical judgment task with implicit cross-language priming with proficient Chinese-Japanese bilinguals (N = 23) by manipulating the orthographic forms of L2-Japanese prime words. We compared the reaction time of the participants' lexical decision of L1-Chinese target words based on four types of L2-Japanese priming words with various semantical and orthographical similarities to L1-Chinese: (1) Japanese-origin words (JOW; Wago) with JC (without semantic or orthographic similarity), (2) JOW with CC (without semantic similarity but high orthographic similarity), (3) JOW with a combination of JC and CC (without semantic similarity and low orthographic similarity), and (4) Chinese-origin words (COW; Kango) with CC (with high semantic and orthographic similarity). The L1-Chinese target words comprised translation equivalents of the L2-Japanese priming words (primed words), non-equivalents of the L2 priming words (unprimed words) as a control, and non-words as fillers. The participants judged the lexicality of L1-Chinese target words following the L2 primed words faster than the unprimed words ($p < 0.05$). Notably, the cross-language priming effect decreased for primed words with no semantic/orthographic similarity (i.e., JOW-JC; $p < 0.05$). In contrast to the Revised Hierarchical Model, our results suggest that proficient Chinese-Japanese bilinguals' access to concepts of L2-Japanese words is mediated by L1-Chinese lexical representations. Their strong reliance on L1 orthography when processing L2 words may facilitate L2 vocabulary learning, but it may prevent the proficient use of L2 words that do not share L1 orthography. This study demonstrated that orthographic similarities play positive and negative roles in Chinese-Japanese (L1-L2) associations.

Gestures in naturalistic classroom teaching and their impact on students' evaluations of teachers

Jing Zhou (Auckland University of Technology) and Yan Gu (University College London).

Gestures are widely adopted in language teaching and learning (Gullberg, 2006; Stam, 2012). Previous studies found that seeing teachers' gestures has implications for students' memory and learning efficiency (Tellier, 2008; Cook et al., 2008) and students' who gesture more were assessed to have a higher L2 proficiency in an audiovisual assessment than in an audio-alone assessment (Gullberg, 1998). However, it is unclear about (1) how teachers make use of gestures in naturalistic classroom teaching and (2) how teachers' gestures affect students' evaluations of teachers in addition to learning. To fill these research gaps, we first studied three teachers' use of different types of gestures in naturalistic English foreign language classroom teaching (Study1) and then examined how teachers' gesturing in the context of L2 vocabulary teaching would affect students' impressions of teachers as well as their learning outcome (Study2). In Study1 we analysed 135 minutes of video recordings of classroom teaching and found that teachers tend to beat and point when teaching difficult or important information, and teachers produced different gestures (i.e., interactive gestures, emblematic gestures) when facing different situations (i.e., disfluency, classroom silence) in L2 teaching. In Study2 we extracted all moments when teachers were teaching a vocabulary (79 clips), where some had no gestures and some had representational gestures, or beat gestures, or points, or self-adaptors. About 300 students would do a perception study to rate the extent to which they think the teacher in a clip explains the vocabulary well. In addition, there would be a quick test on students' recognition of the vocabulary after the rating task. There would be no mention of gestures to students at all. We would analyse how different types of gestures may predict students' evaluations of teachers and their recognition test outcomes. As perception and learning are multimodal, we would extract the prosodic (e.g., pitch, speaking rate) and linguistic features (e.g., sentence types, number of words) of teaching each vocabulary and put them in the regression model as control variables. The results would show how teachers' gestures can impact students' impression and learning after controlling for pitch and question sentence type. This study hopes to provide an in-depth understanding of how producing specific gestures of different types and frequencies will influence students' evaluations of the teachers and their teaching quality, as well as its actual impact on students' learning outcomes in the naturalistic classroom setting. Furthermore, this will serve as a foundation for theoretical and applied research on incorporating education and training programs aimed at raising educators' awareness of the pedagogical importance of gestures for them to provide high-quality verbal and nonverbal input.

A mixed methods study of learning L2 English modality through online concept-based language instruction

Jingyuan Zhuang (The Pennsylvania State University).

Expressing modality or “the speaker’s assessment of, or attitudes towards, the potentiality of a state of affairs” (Radden & Dirven, 2007, p. 233) is central to communication, but causes difficulty even for advanced second language (L2) learners (Tyler, 2012). Grounded in sociocultural theory (Lantolf & Poehner, 2014; Vygotsky, 1978) and cognitive linguistics (Sweetser, 1990; Talmy, 1988), whose synergy is demonstrably facilitative of L2 development, this study examined the effectiveness of online concept-based language instruction (C-BLI) on the L2 development of English modality, using a mixed methods intervention design (Creswell & Plano Clark, 2018). Seventy-eight L2 English learners were randomly assigned to a C-BLI group, a rule-based instruction group that differed with C-BLI in developmental framework but maintained the linguistic analysis of modals, or a control group receiving no instruction. Learners’ conceptual understanding and language use were measured before, immediately after, and one week after online learning, using language analysis, multiple choice, and written production tasks. Results showed that the C-BLI group made statistically significant and large gains in their conceptual understanding over time, and also demonstrated a clear advantage over the other two groups. The gains in the language use scores, however, were less sizeable. In order to obtain further insights into these quantitative results and to better understand the effects of online C-BLI, six learners from the C-BLI group were systematically selected using cluster analysis to participate in the longitudinal follow-up phase. Learners were asked to keep regular journal entries noting observed or experienced modality and thoughts on speaker intention in context. One-on-one journal discussion meetings helped learners internalize the concepts as a systematic and coherent orienting basis to interpret their observations and plan their own L2 use. Implications of the findings will be discussed, and limitations of the study will be pointed out.

References Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). Sage. Lantolf, J. P., & Poehner, M. E. (2014). *Sociocultural theory and the pedagogical imperative in L2 education: Vygotskian praxis and the research/practice divide*. Routledge. Radden, G., & Dirven, R. (2007). *Cognitive English grammar*. John Benjamins. Sweetser, E. (1990). *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure*. Cambridge University Press. Talmy, L. (1988). Force dynamics in language and cognition. *Cognitive Science*, 12, 49–100. Tyler, A. (2012). *Cognitive linguistics and second language learning: Theoretical basics and experimental evidence*. Routledge. Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Colloquium

Bilingual Cognition and Advances in Multi-competence Perspectives

Organisers: Chise Kasai, Motofumi Sumiya, Takahiko Koike, Takaaki Yoshimoto, Hideki Maki and Norihiro Sadato

Discussant: Jean-Marc Dewaele

Integration of language-specific categories in bilingual cognition: Evidence from motion event categorisation in Arabic L2 users of English

Panos Athanasopoulos (Lancaster University) and Amal Albureikan (Qassim University).

The current paper contributes to our understanding of Cook's integration continuum (Cook, 2002) by reporting a study on motion event categorisation in L1 speakers of Modern Standard Arabic (MSA) who are very advanced L2 users of English. According to Talmy's (2000) typology, MSA is a verb-framed language (V-language), where path of motion is expressed as a semantic component of a motion verb, and manner of motion is optionally expressed as a satellite. This contrasts with English, a satellite-framed language (S-language), where path is expressed in satellites, and manner is part of the verb semantics. For instance, the boy ran into the house, is translated into MSA as:

(1) جرى البيت الولد دَخَلَ Entered the boy the house running 'He entered the house running.'

Previous research has employed the so-called triads matching task, where participants have to match a target clip depicting motion with one of two alternate clips resembling the target in either the manner or path of motion. Speakers of V-languages tend to match on the basis of common path, and speakers of S-languages tend to match on the basis of common manner (Gennari et al., 2002). These effects are abolished under a concurrent verbal interference task that disrupts processing in the phonological loop, dampening the online feedback between stimuli and their corresponding linguistic representations (Papafragou & Selimis, 2010).

We employed the afore-mentioned motion triads matching task in monolingual speakers of MSA and MSA-English bilinguals given task instructions in MSA in Saudi Arabia, and in English monolinguals and MSA-English bilinguals given task instructions in English in the UK. Results showed the expected categorisation bias towards path in MSA monolinguals and towards manner in English monolinguals. Overall, bilinguals differed significantly from both monolingual groups regardless of language context, displaying unique categorisation behaviour that was in-between the two monolingual groups. However, findings also revealed a statistically robust, albeit numerically small, language context effect, such that those bilinguals tested in an English context made more manner preferences than those tested in a MSA context. These findings were further corroborated in a subsequent experiment with verbal interference in MSA and in English, in a repeated measures design. Selective disruption of one of the bilingual's languages led to categorisation patterns resembling the uninterrupted language.

Our findings are in line with previous studies showing effects of language experimental context (Athanasopoulos et al., 2015; Bylund & Athanasopoulos, 2017; Emmorey, 2019; Kersten et al., 2010; Lai et al., 2013), but qualify these findings in an important way. While cognitive categorisation can appear to be language-selective, as indexed by effects of language context and verbal interference, the overall categorisation patterns do not resemble those of monolinguals of either language, showing evidence in favour of partial integration of language-specific categorical representations. We thus provide evidence for a representational continuum that may be optimised for functional integration, but can be modulated as a function of situational cognitive demands, reflective of the flexible, dynamic nature of the multi-competent mind.

Cognitive states of bilingual speakers

Chise Kasai, Motofumi Sumiya, Takahiko Koike, Takaaki Yoshimoto, Hideki Maki and Norihiro Sadato (Gifu University, Hamamatsu University School of Medicine, National Institute for Physiological Sciences, Research Organization of Science and Technology, Ritsumeikan University)

A number of studies done by Cook and his colleagues have shown that bilinguals are a unique group of individuals who employ different cognitive states [1]. The purpose of the current study is to examine bilinguals from a neurological point of view by investigating their brain regions when processing second language grammar. Japanese particles, which are known as one of the most difficult grammatical aspects to acquire, were used as experimental stimuli. Highly proficient non-native (L2) Japanese users were tested, and their performance was compared to that of native (L1) speakers of Japanese.

Although a growing body of research has investigated the similarities or differences in linguistic processing between native (L1) and non-native (L2) speakers, the issue still remains controversial. Some studies have shown that the acquisition of grammar by late L2 users is typically less successful and produces less uniform, and perhaps even fundamentally different grammatical systems from L1 acquisition. These studies argue that the rule systems developed by late L2 users do not necessarily conform to the principles that constrain native grammar learners. Conversely, evidence from several experimental studies indicates that late L2 users can achieve native-like processing.

Two principal hypotheses have been expounded regarding L2 grammatical processing of sentence comprehension. Clahsen and Felser's [2, 3] influential shallow structure hypothesis postulates that L2 users adopt "shallow" parsing with reduced sensitivity to grammatical information; thus, a different parsing process occurs from L1. The second hypothesis assumes that L1 and L2 parsing processing is similar and explains the differences therein, in terms of inefficient lexical access routines or an increased burden on capacity-limited cognitive resources such as working memory in L2 [4]. By specifying working memory function as memory retrieval, Cunnings [5] argued that primary source of L1 and L2 differences (in the grammatical process) lies in the ability to retrieve information-constructed processing from memory,

To investigate, a functional MRI study was conducted with L1 Japanese speakers and highly proficient L2 Japanese users requiring oral production of grammatically correct Japanese particles. As a control, participants read out a visually designated mora indicated by underlining. Particle selection by L1 and L2 groups commonly activated the bilateral inferior frontal gyrus/insula, pre-supplementary motor area, left caudate, middle temporal gyrus, and right cerebellum, which constitutes the core linguistic production system. In contrast, the left inferior frontal sulcus, known as the neural substrate of verbal working memory, showed more prominent activation in L2 than in L1. These findings suggest that cognitive workload is involved when bilinguals process and produce L2.

References

- Cook, V., & Bassetti, B. (Eds.). (2011). *Language and bilingual cognition*. Psychology Press.
- Clahsen, H., & Felser, C. How native-like is non-native language processing? *Trends Cogn. Sci.* 10, 564-570 (2006a).
- Clahsen, H., & Felser, C. Grammatical processing in language learners. *Appl. Psycholinguist.* 27, 3-42 (2006b).
- Hopp, H. Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua* 120, 901-931 (2010).
- Cunnings, I. Parsing and working memory in bilingual sentence processing. *Bilingualism* 20, 659-678 (2017).

Effects of L1 and L2 on Categorisation of Body Parts from the Perspective of Multi-competence

Miho Sasaki (Keio University) and Yoshiko Murahata (Kochi University).

Recent categorisation studies have shown that languages segment parts of the body differently and that even if two languages have equivalent body part terms they cover a slightly different portion of the body (Enfield, Majid, & van Staden, 2006; Majid, 2010). Then the question arises whether conceptual representation of body parts in bilinguals is affected by their L1 or L2, or both. Little research on categorisation of body parts has been done so far in the field of bilingual cognition. Japanese terms for body parts such as *te* 'hand,' *ashi* 'foot,' *atama* 'head,' and *senaka* 'back,' and their English equivalents do not cover the same areas (Kasai, et al. 2012). The current study uses a body colouring task (van Staden & Majid, 2006) and investigates from the multi-competence perspective (Cook & Li, 2016) specifically whether Japanese speakers would shift the categorisation of body parts because of learning and using English as an L2. The test materials were prepared in Japanese and English, and the study included 131 Japanese users of English (JUEs) and 18 English L1 users. The JUEs were asked to colour each of the eight named body parts in Japanese (*te*, *ude*, *atama*, *kao*, *ashi* (脚) and *ashi* (足), *senaka*, *koshi*) and they were tested again 3-6 weeks later in English (hand, arm, head, face, leg, foot, back, lower-back). They were divided into four groups according to their English proficiency and experiences. The colouring was scored and analysed in smaller segments of the tested body parts. The qualitative analysis observes effects of proficiency as well as individual variation of colouring patterns including some deviated patterns in bilinguals, suggesting conceptual integration. Further statistical analysis shows that the JUEs categorise the areas such as foot and leg differently in the two languages. The effects of proficiency are also observed in the categorisation of head, foot and lower-back in English. The usefulness of qualitative analysis of bilingual cognition data and important factors in groupings of L2 users/bilinguals are discussed along with the overall results.

References: Cook, V. J. and Li, W. (Eds.) (2016). *The Cambridge Handbook of Linguistic Multi-competence*. Cambridge, UK: Cambridge University Press. Enfield, N. J., Majid, A., & Staden, M. (2006). Cross-linguistic categorization of the body: Introduction. *Language Sciences*, 28, 137-147. Kasai, C., Murahata, Y., Sasaki, M., Takahashi, J., Hattori, N., Iwai, A., Minoura, M. & Cook, V. J. (2012). Categorisation of body parts by Japanese learners of English. The 22nd Annual Conference of European Second Language Association. Poster presentation. Poznan, Poland. Majid, A. (2010). Words for parts of the body. In B. C. Malt & P. Wolff (Eds.) *Words and the mind: How words capture human experience* (pp. 57-71). Oxford, UK: Oxford University Press. van Staden, M. & Majid, A. (2006) Body colouring task. *Language Sciences*, 28, 158-161.

Poster presentations

How effective is intentional vocabulary learning using word cards?

Sally Alghamdi (Albaha University) and Jeanine Treffers-Daller (University of Reading).

Research shows that intentional vocabulary learning is more effective than incidental vocabulary learning (Elgort, 2011; M. STRATTON, 2022; Nation, 2013). Moreover, intentional learning using word cards is one of the most effective techniques in vocabulary learning (Nakata, 2019) as it can help commit large number of vocabulary items to long term memory in a short period of time (Dunlosky, Rawson, & Marsh, 2013; Fitzpatrick, Al-Qarni, & Meara, 2008). On the other hand, incidental exposure to vocabulary items in graded readers is viewed as the main means of vocabulary retention in some research (Krashen, 2004; Pitts, White, & Krashen, 1989). In the current study, seventy-five L1 low intermediate Arabic learners of English were assigned to three groups where they either learned a set of 45 intentionally using word cards or learned them incidentally using graded readers and a control group (CG) that did not receive any vocabulary guidance. The proportion of target words retained by Word Card Group (WCG) was compared to the proportion of target words retained by the Graded Readers Group (GRG) and both were compared with the CG performance. Four test sessions, three immediate post-test and one delayed post-test, were administered to measure vocabulary receptive knowledge with a Lexical Decision Task (LDT), and productive vocabulary knowledge was measured with a Gap-Fill Test (GFT). Analyses with a mixed factorial ANOVA revealed that the WCG outperformed other groups in all sessions in receptive and productive knowledge of target words. We also investigated the effect of word characteristics (morphological complexity and word length) and frequency of exposure on retention for Arab ESL learners. Multiple regression analyses showed that the frequency of exposure to the target words predicted the accuracy scores on the LDT and the scores on the GFT for the WCG, and the number of syllables in a target word predicted the GFT scores for all the groups combined. We conclude by formulating implications for intentional and incidental vocabulary learning.

References:

- Dunlosky, J., Rawson, K. A., & Marsh, E. J. (2013). Nathan MJ willingham DT improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Association for Psychological Science*, 14(1), 4-58.
- Elgort, I. (2011). Deliberate learning and vocabulary acquisition in a second language. *Language Learning*, 61(2), 367-413.
- Fitzpatrick, T., Al-Qarni, I., & Meara, P. (2008). Intensive vocabulary learning: A case study. *Language Learning Journal*, 36(2), 239-248.
- Krashen, S. D. (2004). *The power of reading: Insights from the research: Insights from the research ABC-CLIO*.
- M. STRATTON, J. (2022). Intentional and incidental vocabulary learning: The role of historical linguistics in the second language classroom. *The Modern Language Journal*.
- Nakata, T. (2019). Learning words with flash cards and word cards. *The Routledge Handbook of Vocabulary Studies*, , 304-319.
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge: Cambridge University Press.
- Pitts, M., White, H., & Krashen, S. (1989). Acquiring the second language vocabulary through reading: A replication of the clockwork orange study using second language acquirers.

Language awareness in primary school EFL lessons: teachers' cognitions and practices

Monika Bader (Western Norway University of Applied Sciences), Christine Möller-Omrani (Western Norway University of Applied Sciences) and Simon Borg (Western Norway University of Applied Sciences).

Language awareness (LA), often defined as 'explicit knowledge about language, and conscious perception and sensitivity in language learning, language teaching and language use' (ALA, n.d.), plays an important role in foreign language learning and teaching (e.g. Svalberg, 2007; Frijns et. al. 2018), as well as in national curricula, including in Norway, which is the context for the research we present here. How effectively LA is promoted in the classroom is influenced by teachers' cognitions (Borg, 2019) and pedagogical expertise and this study (which was part of a larger project examining the development of metalinguistic awareness in primary school pupils), examined three questions: 1. To what extent are English lessons for Grade 3 (mean age 8) characterized by pedagogical practices that promote the development of pupils' LA? 2. Where LA is promoted, (a) what pedagogical practices do teachers adopt and (b) which areas of language are targeted? 3. What cognitions – beliefs, knowledge, understandings and perceptions – related to LA but also to L2 teaching and learning more generally, to young learners and to the educational context (including the curriculum) shape teachers' pedagogical practices in relation to LA?

The results presented come from case studies of four teachers of English whose practices and cognitions were explored qualitatively through one background interview (to learn about the teachers' context and their general views about language awareness), two lesson observations (which were studied for evidence of practices that promote LA) and one post-observation interview (where classroom episodes relevant to LA were discussed with the teacher).

Overall, English lessons were characterized by modest levels of attention to LA. Lessons were typified by series of language activities which pupils completed, typically with limited explicit attention to language. In discussing their lessons, teachers demonstrated uncertain understandings of what LA meant and of how to interpret references to LA in the curriculum. While teachers were generally positive about the value of LA in the learning of English, various factors combined to limit the extent to which LA was promoted in their lessons. These included their beliefs about young learners and the process of L2 learning, as well as teachers' lack of theoretical and pedagogical knowledge relevant to LA.

On the basis of these conclusions, suggestions for the kinds of professional development support that would enable teachers to promote LA more effectively in primary English classrooms will be made.

References

- Association for Language Awareness (ALA). (n.d.). At https://www.languageawareness.org/?page_id=48
- Borg, S. (2019). Language teacher cognition: Perspectives and debates. In X. Gao (Ed.), *Second handbook of English language teaching* (pp. 1149-1170). Springer International Publishing.
- Frijns, C., Sierens, S., Van Avermaet, P., Sercu, L., & Van Gorp, K. (2018). Serving policy or people? Towards an evidence-based, coherent concept of language awareness for all learners. In C. Hélot, C.Frijns, K.V.Gorp, & S. Sierens (eds.), *Language Awareness in Multilingual Classrooms in Europe* (pp. 87-116). De Gruyter Mouton.
- Svalberg, A. (2007). Language awareness and language learning. *Language Teaching*, 40(4), 287-308. doi:10.1017/S0261444807004491

Acquisition of generic meanings in L3 Italian

Eleonora Boglioni (University of Southampton) and Roumyana Slabakova (University of Southampton).

Genericity contrasts. English speakers with Spanish as a second language learning Italian can potentially rely on the L1 and L2 grammars to parse the L3 input. Spanish and Italian share lexical and structural similarities, e.g., in the expression of generic meanings. Generic descriptions can refer to all the members of a species (kind reading: Tigers are becoming extinct) or instances of a species engaging in habitual behaviour (characterizing reading: Lions roar) (Krifka et al., 1995). Importantly, the syntactic realization of both generic meanings varies crosslinguistically. While English bare plural subjects can have kind/characterizing readings (Chierchia, 1998), Spanish and Italian generic subjects require definite articles (Longobardi, 1994). Hence, Spanish and Italian pattern similarly in this respect. Nonetheless, Spanish and Italian pattern differently in allowing bare singular objects with number neutral interpretation in a restricted verb class (e.g., *Tengo coche*, *I have car), where the object refers to one or more car(s) (Espinal, 2010). To express a similar but not exactly the same meaning, English and Italian require overt determiners. Thus, for each property, two of the languages investigated pattern together, to the exclusion of the third one. L3 acquisition predictions. Our research investigates crosslinguistic influence (CLI) effects from the background languages to the L3, testing the effects of structural similarity and L3 proficiency. On objects, English–Italian parallels could be beneficial, while negative CLI from Spanish is expected. On subjects, positive influence from Spanish is predicted. Methodology. We tested 60 adult L3 Italian learners with English and Spanish alternatively as L1/L2. To address the role of input, exposure and use, we recruited 30 L1 English speakers in England and 30 L1 Spanish speakers in Spain. These settings could favour English CLI on objects and Spanish CLI on subjects. Comprehension and use of generics were assessed in the L2 and L3 by means of an Acceptability Judgment Task, a Form-to-Meaning Task, and an Elicited Oral Production Task. Results. We ran linear mixed models (R packages *lme4*, *glmer*) on the L3 Italian judgements and oral data (Tables 2–4). On the AJT, significant differences ($p < .001$) were found on the contrasts between acceptable and unacceptable kind-referring subjects, and number neutral objects (Table 2), in interaction with L3 proficiency. This suggests property-by-property CLI with facilitative effects from Spanish on subjects and from English on objects. In oral production, we observed a main effect for L3 proficiency (Table 4). Individual results (Table 5) point to negative CLI from English (on subjects) and Spanish (on objects). Theoretical proposal. Additional language acquisition involves reorganising the known form–meaning mappings. Taken together, our experimental results point to Hegemony of the Form. In the case of universal semantic distinctions where all languages express the same meanings, although with different morphosyntactic means (e.g., kind/characterizing genericity), the forms exert strong crosslinguistic influence. Where a different meaning exists in one of the background languages (the Spanish number neutral meaning), it does not exert such a strong crosslinguistic influence. We will discuss implications of this proposal in other additional language acquisition scenarios. (500 words)

Internal validity of the new LLAMA (v.3) aptitude tests

Lars Bokander (Jönköping University), Vivienne Rogers (Swansea University), Paul Meara (Swansea University) and Brian Rogers (Swansea University).

The LLAMA aptitude tests have undergone significant revision since their first iteration (Meara, 2005). The revised version 3 has been available on the Lognostics website since 2021. In this paper we will update the internal validity analysis from Bokander & Bylund (2020) with data from 640 test-takers of the new version.

We analysed the LLAMA v3 scores with respect to item performance and internal consistency. Item performance was evaluated with classical test theory and Rasch model fit. Internal consistency is influenced by dimensionality, i.e., whether the test measures one or more constructs. Bokander & Bylund (2020) pointed out issues with dimensionality in some LLAMA v.1 subtests. We address this by reporting coefficient omega along with the traditional Chronbach's alpha. Omega is based on a factor analysis of the items in a test and has less strict statistical assumptions than coefficient alpha, for example, with regards to unidimensionality (Dunn et al., 2014).

The results indicate several improvements in the new LLAMA tests. Overall, items on average performed better than in LLAMA v1 with respect to Rasch model fit and discrimination between participants. Improved item performance in turn reflects improvements in internal consistency. LLAMA B has not notably changed, and again produced reliable scores. LLAMA D has undergone important changes in how the items are presented to test-takers. This has increased the internal consistency of the scores significantly compared to version 1. Coefficient omega for LLAMA D indicates an internal consistency comparable to that typically found with instruments in SLA research (Plonsky & Derrick, 2016). LLAMA E scores produced high values for both alpha and omega. This improvement is due to a greater score variance and the more normal distribution of scores obtained with the new version. LLAMA F has undergone major changes in item format and scoring procedure. We found a significantly higher internal consistency in our data set than what has previously been reported, with both alpha and omega performing well. Dimensionality related issues seem less pronounced with the new version, although some questions remain about its unidimensionality.

An implication for language aptitude researchers is that the new LLAMA tests should be able to produce better correlations with other SLA measures, because correlations depend on reliably obtained scores from all instruments in a study. If there actually exists a relationship between two variables, it may still go undetected when unreliable instruments are employed. The risk of this occurring in future aptitude research has now decreased.

References

- Bokander, L., & Bylund, E. (2020). Probing the internal validity of the LLAMA language aptitude tests. *Language Learning*, 70(1), 11-47.
- Dunn, T. J., Baguley, T., & Brunnsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, 105(3), 399-412.
- Meara, P. (2005). LLAMA language aptitude tests: The manual. Lognostics.
- Plonsky, L., & Derrick, D. J. (2016). A meta-analysis of reliability coefficients in second language research. *The Modern Language Journal*, 100(2), 538-553.

Interactive didactics: enhancing L2 speaking complexity and accuracy through a blended drama approach

Simona Bora (University of Essex (UK) / Free University of Bozen (IT)).

A steadily increasing number of educators in different contexts have been focusing their work on ways in which dramatic approaches can support foreign language learning. Drama is “communication between people” (Via, 1987: 10) and “inextricable part of all social interactions” (Russel di Napoli 2003:17) and therefore offers potentially a very rich interactive context for enhancing students’ L2 oral skills. Although drama in language teaching has gained recognition recently, there have been no quantitative studies conducted to date which investigate the gains made by students learning languages through a blended drama approach in a mandatory high-school curriculum (Bora 2022) in terms of complexity and accuracy. This paper sheds light on findings from a longitudinal (20 weeks) mixed-method approach study conducted with high-school intermediate level of language proficiency Italian students. Participants were exposed to a blended drama approach: theatre texts combined with drama activities followed by a stage performance. A control group was taught through a traditional approach instead. Quantitative data has been collected through the implementation of an oral pre-test, a mid-test and a post-test with three tasks: oral proficiency interview, story retelling and guided role-play. The impact of this approach is measured in terms of students’ gains in oral accuracy and complexity when tasks are taken together or separately. Results lend support to previous hypotheses of the effectiveness of a blended drama approach to enhancing L2 speaking skills compared to a traditional one. Implications for teaching practice are further discussed.

References - Bora, S. F (2022). Drama Pedagogy: Speaking Accuracy and Complexity through Contemporary Theatrical Texts and Performance in Foreign Language Learning. *Research in Drama Education: The Journal of Applied Theatre and Performance* - Housen, Alex, Folkert Kuiken, Ineke Vedder (2012), *Dimensions of L2 Performance and Proficiency. Complexity, Accuracy and Fluency in SLA*, Amsterdam: John Benjamins. - Palloti, Gabriele (2015), “CAF: defining, refining and differentiating constructs”, *Applied Linguistics*, 30(40), 590-601.

Cognitive engagement and glossing: effects of L1, L2 and intercomprehensible glosses on vocabulary learning. An eye-tracking study.

Ilaria Borro (Università degli Studi di Bergamo).

Studies of the effectiveness of glossing in L2 reading hypothesized the amount of cognitive engagement with the glossed words to be a relevant factor for vocabulary learning (Boers 2022). Different strategies have been implemented to manipulate cognitive engagement, such as L2 rather than L1 glosses (Kim et al 2020), multiple-choice glosses (e.g. Yoshii 2013), and glosses providing only one of the meanings of polysemic items (Boers 2000). Findings are currently mixed, albeit encouraging. Notably, none of the existing studies employed real-time measurements of cognitive engagement, which has been assumed rather than assessed. The present project aims to contribute filling this gap. One-hundred L2 Italian learners will be assigned to three experimental groups, each exposed to a version of a reading text with glosses requiring a growing amount of cognitive engagement: (i) L1 glosses; (ii) L2 glosses; (iii) glosses in a romance language, unknown but intercomprehensible to the learners due to a degree of semantic transparency with Italian (Blanche-Benveniste et al 1979). A control group will read the text with no glosses. The reading behavior and cognitive engagement will be assessed through eye-tracking, while the creation of explicit and implicit knowledge of vocabulary will be measured, respectively, with a set of pencil-and-paper tests and a priming protocol test. Immediate and delayed post-test results will be triangulated with eye-tracking data and retrospective verbal report outcomes, to value the participants' awareness at the point of learning. Expected outcomes include a larger cognitive engagement for L2 and intercomprehensible glosses as compared to L1 glosses, as well as a significant effect of cognitive engagement on vocabulary learning and retain. The project contributes with empirical evidence to the debate about the role of attention and cognitive engagement for the creation of implicit and explicit knowledge (Long 2017). Moreover, it presents immediate pedagogical implications for a more effective use of glosses in instructed SLA. Finally, it provides data regarding the possible use of intercomprehension as a pedagogic tool for language learning.

References Blanche-Benveniste, C., Borel, B., Deulofeu, J., Durand, J., Giacomi, A., Loufrani, C., Meziane, B., & Pazery, N. (1979). Des grilles pour le français parlé, « Recherches sur le français parlé », 2, 163-206. Boers, F. (2000). Enhancing metaphoric awareness in specialised reading. *English for Specific Purposes*, 19(2), 137–147. Boers, F. (2022). Glossing and vocabulary learning. *Language Teaching* (2022), 55, 1–23. Kim, H. S., Lee, J. H., & Lee, H. (2020). The relative effects of L1 and L2 glosses on L2 learning: A meta-analysis. *Language Teaching Research*, Online First. Long, M. (2017). Geopolitics, methodological issues, and some major research questions. *ISLA*, 1(1), 7- 44. Yoshii, M. (2013). Effects of gloss types on vocabulary learning through reading: Comparison of single and multiple gloss types. Special issue on learner-computer interaction in language education: A Festschrift in honor of Robert Fisher, *CALICO Journal*, 203–229.

Pragmatic awareness and proficiency: Are highly proficient learners more pragmatically aware?

Athenea Botey (Universitat de Barcelona) and Júlia Barón (Universitat de Barcelona).

In the field of interlanguage pragmatics (ILP), the relationship between foreign/second language (L2) proficiency and pragmatic competence has long been debated. Research on the interface between linguistic and pragmatic competence in an L2 has yielded results that point to the correlation of learners' proficiency level and pragmatic development (i.e., Bardovi-Harlig and Dörnyei, 1998; Cook and Liddicoat, 2002; Schauer, 2006) as well as the correlation of grammatical and pragmatic development (i.e., Hoffman-Hicks, 1992; Håkansson and Norrby, 2005; Celaya and Barón, 2015). Nonetheless, a number of studies have found that pragmatic competence and proficiency do not correlate and are in fact independent (i.e., Niezgodá and Röver, 2001; Matsumura, 2003). Most of the research regarding this interface has tested their participants' productive knowledge of pragmatics. In addition, speech acts such as requests, apologies, suggestions and refusals have been amongst the most frequently investigated in the aforementioned studies. That is why the present study focuses on compliment responses (CRs) and on perceptive knowledge of pragmatics. Therefore, this research aims to investigate the interface between proficiency and ILP by focusing on the pragmatic awareness of the speech act of CRs.

The participants were 72 in total, of which 8 were native speakers (NSs) of English (mean age= 19.5) and 64 were English as a foreign language (EFL) learners (mean age= 16.6) with Spanish/Catalan as their first language (L1). In order to assess their perceptive pragmatic knowledge of CRs, a pragmatic awareness video elicitation task (PAVET) was administered. When carrying out the PAVET test, participants had to rate the appropriateness of 15 CRs using a Likert-scale that went from 1 (inappropriate) to 6 (very appropriate). The audio-visual compliment-compliment response sequences were extracted from the TV series *Gilmore Girls*, chosen for its display of various degrees of social distance and power. Besides completing the PAVET, four participants agreed to carry out semi-structured post hoc interviews. The EFL learners were divided into groups of high and low proficiency according to their vocabulary sizes (V_YesNo v1.1 [lognostics.co.uk]) to analyse if and in what ways their responses to the task differed. Their ratings were then compared to a NS benchmark to obtain a sameness score.

Results indicated a moderate inverse correlation between proficiency level and near-nativeness. Findings, on the one hand, suggest that the L1 of the learners plays an important role in their L2 pragmatic awareness. On the other hand, they point to the direction that proficiency alone is not a determinative factor in order to acquire pragmatic knowledge as there are several other factors that can influence pragmatic development, such as individual differences, the non-linearity of pragmatic development, the use of compensatory nonverbal strategies or exposure to limited pragmatic input in the classroom context. All in all, given the results, it could be implied that pragmatic development and L2 language proficiency are of an independent nature.

From paper-based to computer-based integrated reading-to-write: Evidence for delivery mode effects at the CEFR B1 level

Tineke Brunfaut (Lancaster University), Luke Harding (Lancaster University) and Aaron O. Batty (Keio University).

Computer-based tasks are increasingly used in the teaching and assessment of second language (L2) writing, reflecting large-scale changes in day-to-day writing modes, technological developments, increased computer accessibility, administrative efficiency, and—most recently—online delivery due to the Covid19 pandemic. While some writing tasks are conceptualised as computer-based from inception, others have shifted mode from their original paper-based format (either to replace the latter or be used in parallel mode). Research has shown, however, that mode of delivery may have an effect on the L2 writing process and product, and on writers' perceptions (see Chan, 2018). In the last two decades, the teaching and assessment of writing has furthermore broadened from independent, writing-only tasks to include integrated tasks such as reading-to-write or listening-to-write, reflecting forms of authentic language use. While a sizable body of research on integrated-writing tasks now exists (Plakans, 2015), little is known yet about how these types of tasks operate comparatively in different delivery modes. However, a study by Authors (year) looking into the effect of delivery mode on a suite of independent and integrated writing tasks across CEFR B1-C1 proficiency levels found that while delivery mode had no discernible effect on scores at the B2 and C1 level, learners completing the B1-level tasks scored slightly, but statistically significantly, lower on the computer-based mode than they did on the paper-based mode.

This presentation reports on a follow-up study, aiming to gain more detailed insights into the observed delivery mode effect of the integrated reading-to-write task at the B1 level. To this end, the discourse characteristics of students' written samples on the B1 tasks (both independent and integrated ones) were investigated. 106 English-L2 learners based in three European countries each completed one version of the B1 writing-only task and one of the B1 reading-to-write task in the paper-based mode, and another version of each task in the computer-based mode. The tasks were counterbalanced for order of delivery mode and for task version. To establish the discourse characteristics of the written performances, students' writing samples were analysed using 28 automated discourse measures (using various software tools) and 8 rating scales developed for fine-grained judgements by two experts. The analyses covered the areas of reading-text source use, task fulfilment, textual organisation and structure, and language control. Comparative statistics (Wilcoxon Signed Rank Tests) showed that effects on discourse features were primarily associated with the integrated reading-to-write tasks, rather than the writing-only tasks, thus triangulating the previous score-based findings Authors (year). Specifically, learners gained small advantages in the paper-based mode regarding the selection of relevant content from source materials, and on measures of fluency and accuracy. Implications will be discussed concerning the potential impact of user-interface features and spelling/proofreading challenges for lower proficiency learners (B1) in the screen-based completion of integrated tasks.

References Authors (Year). Chan, S. (Ed.) (2018). The comparability of paper-based and computer-based writing: Process and performance [Special Issue]. *Assessing Writing*, 36. Plakans, L. (2015). Integrated second language writing assessment: Why? What? How? *Linguistics and Language Compass*, 9(4), 159-167.

Exploring effects of early extramural English exposure on university students' current L2 vocabulary

Nicole Busby (Norwegian University of Science and Technology (NTNU9)).

This exploratory study investigated university students' reported exposure to English through informal out-of-class (extramural) activities at an early age and how this related to their current vocabulary size in L2 English. Studies have shown that extramural activities can be an important predictor of L2 vocabulary knowledge and that the amount of time spent on these activities is very important (e.g. Sundqvist, 2009; Peters, 2018). Research into effects of extramural exposure on language acquisition has tended to focus on looking for relationships between L2 proficiency and extramural activities that participants engage in at the time of the study, but less is known about the relationship between language proficiency and activities earlier in life which could have contributed to L2 acquisition. To investigate this, students at a Norwegian university (N = 34) were asked to complete a survey comprising questions about their early extramural English activities and questions from the Vocabulary Size Test (VST: Nation & Beglar, 2007). Participants were asked to report the earliest extramural activity they felt made an important contribution to their current knowledge of English and the extramural activity they considered most important (if these were different), as well as the age at which they engaged in these activities. Participants' mean English vocabulary size, as measured by the VST, was 11,832 words. All participants were able to report specific activities that they felt were important to their acquisition of English. The most frequently reported categories for participants' earliest English activities were watching TV shows (24%) and reading books (26%). Activities that participants reported as the most important were more varied, with the most common being TV shows (21%), social media (15%) and online videos (15%). Regression analysis found that the age of reported earliest extramural exposure was a significant predictor for L2 vocabulary size, but that the current age of participants was not a significant predictor of vocabulary size. This is in line with previous research showing that age of onset is an important predictor of L2 acquisition (e.g. Hyltenstam, 1992), but it was interesting to see that this appears to be more important in this sample than the overall time available for language input (as indicated by participant age). Although this approach of asking participants for self-reports of early memories clearly has limitations in terms of accuracy of reporting, results from this exploratory study suggest that investigating early exposure to informal language learning activities could be an important avenue for future research.

Hyltenstam, K. (1992). Non-native features of near-native speakers: On the ultimate attainment of childhood L2 learners. *Advances in Psychology*, 83, 351-368. Nation, I.S.P. & Beglar, D. (2007) A vocabulary size test. *The Language Teacher*, 31(7), 9-13. Peters, E. (2018). The effect of out-of-class exposure to English language media on learners' vocabulary knowledge. *ITL-International Journal of Applied Linguistics*, 169(1), 142-168. Sundqvist, P. (2009). *Extramural English matters: Out-of-school English and its impact on Swedish ninth graders' oral proficiency and vocabulary* (Doctoral dissertation, Karlstad University).

Pragmatic Self-Concepts of Multilingual Children and Adolescents in Germany

Maria Busch (Martin Luther University of Halle-Wittenberg), Luca Plachy (Martin Luther University of Halle-Wittenberg), Matthias Ballod (Martin Luther University of Halle-Wittenberg) and Stephan Sallat (Martin Luther University of Halle-Wittenberg).

Growing diversity in global contexts leads to higher awareness of linguistic and cultural diversity and increasing research in multilingual language development and foreign & second language acquisition (FSLA). Although psychological factors of FSLA such as motivation and learners attitudes have been largely studied in mostly adult learners, the subjective perspectives of children and adolescents have not been given intense focus.

This is where so-called self-concepts, which are subjective views on oneself about one's own characteristics, competencies and challenges, come into play (Mercer, 2011). Self-concepts are not only related to motivation and learning processes, they are also connected with communicative interaction. Pragmatic language is a highly relevant aspect for multilingual language development as well (Félix-Brasdefer, 2017; Kecskes, 2018). While verbal and social self-concepts have been studied widely, it is unclear how self-concepts concerning pragmatic language have been taken into account so far. The research project "MehrSelbst" (funded by the German Federal Ministry of Education and Research, project duration 10/2022-09/2025) addresses this research gap and investigates research questions on how multilingual children and adolescents in Germany express their self-concepts through self-descriptions and self-evaluations regarding their own pragmatic language. This study therefore explores the perspectives of children and adolescents on their own pragmatic competences in a qualitative research design and triangulates them to their structural language and the evaluation of pragmatics by familiar caregivers and educators. Therefore, this proposed poster aims to present the theoretical background (results of a scoping review) of multilingual pragmatic language development, the interface with self-concept research and methodological considerations.

References: Félix-Brasdefer, J. César (2017): Interlanguage Pragmatics. In: Yan Huang (Hg.): The Oxford handbook of pragmatics. Oxford, United Kingdom: Oxford University Press (Oxford handbooks in linguistics), p. 416–434.

Kecskes, I. (2018). Intercultural Pragmatics. In F. Liedtke & A. Tuchen (Hrsg.), Handbuch Pragmatik (p. 140–148). J.B. Metzler.

Mercer, S. (2011). Language learner self-concept: Complexity, continuity and change. *System*, 39(3), 335–346. <https://doi.org/10.1016/j.system.2011.07.006>

A Systematic Review of the Construct Validity of the Academic Pearson Test of English

Phat Cao (University of Nottingham Malaysia), Csaba Zoltan Szabo (University of Nottingham Malaysia) and Barry Lee Reynolds (University of Macau).

Global-scale language proficiency tests, such as the Academic Pearson Test of English (PTEA), are often judged by the extent to which they accurately measure the underlying theoretical concept they claim to measure, termed construct validity (Chapelle, 2021; Hughes & Hughes, 2020). Construct validity, in turn, yields trust and confidence among various stakeholders. Despite a plethora of studies investigating different aspects of construct validity of the PTEA (e.g., De Jong & Zheng, 2016; Zheng & Mohammadi, 2013), the results have been mixed. One reason for this is the variation in conceptualising test constructs and their characteristics, which often lead to discrepancies in how the data are collected, analysed, and interpreted in the existing studies. Using a systematic review methodology, this study explored and evaluated the features of PTEA construct validity studies in terms of research objectives, construct characteristics and methodologies employed. Results from the review of 35 studies reveal that a variety of aspects of construct validity have been examined using both qualitative and quantitative designs, with trait-related constructs (i.e., gender factor or cognitive demand) being the predominant research focus and quantitative approaches predominantly employed. The findings additionally highlight discrepancies in the conceptualisation of constructs, resulting in the adopted theoretical frameworks and methods for construct validation as well as a tendency to overlook enabling skills (i.e., grammar, oral fluency, pronunciation, spelling, vocabulary, and written discourse). To contribute to the scholarship on construct validity of the PTEA, recommendations are made for broadening the research objectives and employing a mixed-methods approach in lieu of only quantitative or qualitative analyses. Furthermore, the theoretical frameworks employed for test validation purposes were considered. Implications are also discussed as they are related to construct validity and validation research of the PTEA as well as language testing.

Exploring the Relationship Between Teachers' and Students' Emotional Intelligence and Emotional Vocabulary

Allen Chee (University of Nottingham Malaysia), Csaba Szabo (University of Nottingham Malaysia) and Sharimila Ambrose (University of Nottingham Malaysia).

Research in the affective domain indicates that learners with higher Emotional Intelligence (EI) are more likely to succeed academically due to their ability to regulate emotions, build social relationships, understand, and express their own and others' emotions (Maccann et al., 2020). Meta-studies have shown that there is a significant relationship between EI, learners' academic achievement and English language abilities (Akpur, 2020; Sánchez-Álvarez et al., 2020). However, learners' ability to recognise, process, and express their own and others' emotions may be impaired by the lack of requisite emotional vocabulary (EV) knowledge (Dewaele, 2015). EV was found to play a vital role in learners' emotional development and that language was a key element in shaping the development of these concepts (Nook et al., 2020). However, learners must first acquire awareness and knowledge of EV in order to comprehend and verbalise their emotions (Hoemann et al., 2019). Despite the importance of EI and EV in language learning, the relationship between EI and EV is largely unexplored, especially in Malaysia. This study aimed to investigate the relationship between the four factors of learners' EI (well-being, self-control, emotionality, sociability) and EV. The first study was conducted with 46 foundation students at a private university in Malaysia. Participants completed an EI questionnaire (TEIQue-SF) and a label generation task as a measure of productive EV knowledge. The results showed a positive and significant correlation between EI and EV. The emotionality factor had the strongest correlation with EV, while language background had no effect. The second study aimed to confirm and expand these results by controlling for language background, dominance, and proficiency in a larger cohort of participants. These findings confirmed the significant relationship between EI and EV. This highlights the importance for curriculum developers and practitioners to support students' development of EI and EV when developing teaching materials and best teaching practices.

References Akpur, U. (2020). A Systematic Review and Meta-Analysis on the Relationship Between Emotional Intelligence and Academic Achievement. *Educational Sciences: Theory & Practice*, 20(4), 51–64. <https://doi.org/10.12738/JESTP.2020.4.004>

Dewaele, J.M. (2015). On emotions in foreign language learning and use. *The Language Teacher*, 39(3), 13–15.

Hoemann, K., Xu, F., & Barrett, L. F. (2019). Emotion words, emotion concepts, and emotional development in children: A constructionist hypothesis. *Developmental Psychology*, 55(9), 1830–1849.

<https://doi.org/10.1037/dev0000686>

Maccann, C., Jiang, Y., Brown, L. E. R., Double, K. S., Bucich, M., & Minbashian, A. (2020). Emotional Intelligence Predicts Academic Performance: A Meta-Analysis. *Psychological Bulletin*, 146(2), 150–186.

<https://doi.org/10.1037/bul0000219.supp>

Nook, E. C., Stavish, C. M., Sasse, S. F., Lambert, H. K., Mair, P., McLaughlin, K. A., & Somerville, L. H. (2020). Charting the development of emotion comprehension and abstraction from childhood to adulthood using observer-rated and linguistic measures. *Emotion (Washington, D.C.)*, 20(5), 773–792.

<https://doi.org/10.1037/EMO0000609>

Sánchez-Álvarez, N., Berrios Martos, M. P., & Extremera, N. (2020). A Meta-Analysis of the Relationship Between Emotional Intelligence and Academic Performance in Secondary Education: A Multi-Stream Comparison. *Frontiers in Psychology*, 11, 1517. <https://doi.org/10.3389/FPSYG.2020.01517/BIBTEX>

Exploring Linguistic Relativity: The Effect of the French Grammatical Gender System on Bilingual Adults' Perception of Objects

Zhuohan Chen (Department of Education, university of Oxford).

A growing body of literature on linguistic relativity (the hypothesis that language influences thoughts) has been focusing on whether there is an effect of the grammatical gender system on perception (Bassetti & Nicoladis, 2016; Lambelet, 2016; Samuel et al., 2019). Evidence suggests that there is an effect of the gender system on perception of both monolinguals and bilinguals (Bassetti & Nicoladis, 2016). However, the findings are potentially limited due to certain methodological limitations (Samuel et al., 2019), for example, the lack of L2 proficiency testing (Bassetti, 2007). Thus, attempting to apply a more rigorous methodology, this study aims to assess the potential effect of the French grammatical gender system on French speakers' and learners' perceptions of object gender. The study involved 140 participants, divided into four groups (N = 35 per group): English monolinguals, French monolinguals, English-French (English-dominant) bilinguals, and French-English (French-dominant) bilinguals. An online experiment was distributed to participants, including a background information questionnaire, English and French vocabulary tests, and a voice distribution task. Quantitative data were analysed using multi-level modelling, relevant regression analyses, and t-tests. Follow-up open-ended question data were coded and analysed using chi-squared tests. The results supported linguistic relativity: The French grammatical gender system affected the perception of French monolinguals and English-dominant bilinguals. The effect of French on French-dominant bilinguals was not reduced by the acquisition of English and seemed to be independent of L2 proficiency. Additional findings included the potential tendency of the French to introduce gender biases (stereotypical association between objects and gender, resulting from the grammatical gender system). The present study adds supporting evidence to the relativity debate by attempting to apply rigour by utilising a pilot study, a robust sample size, pertinent control items, L2 proficiency testing, and advanced data analysis tools. Pedagogically, findings highlight the need to emphasise the discrepancies between students' potential preconceived perceptions and the rules of the L2 grammar when teaching an L2 with a different grammatical gender system to L1. The study also shed new light on whether gendered languages draw out potential gender stereotypes among bilinguals.

Bassetti, B. (2007). Bilingualism and thought: Grammatical gender and concepts of objects in Italian-German bilingual children. *The International Journal of Bilingualism: Cross-Disciplinary, Cross-Linguistic Studies of Language Behaviour*, 11(3), 251-273. doi:10.1177/13670069070110030101

Bassetti, B., & Nicoladis, E. (2016). Research on grammatical gender and thought in early and emergent bilinguals. *The International Journal of Bilingualism: Cross-Disciplinary, Cross-Linguistic Studies of Language Behaviour*, 20(1), 3-16. doi:10.1177/1367006915576824

Lambelet, A. (2016). Second grammatical gender system and grammatical gender-linked connotations in adult emergent bilinguals with French as a second language. *The International Journal of Bilingualism: Cross-Disciplinary, Cross-Linguistic Studies of Language Behavior*, 20(1), 62-75. doi:10.1177/1367006915576832

Samuel, S., Cole, G., & Eacott, M. J. (2019). Grammatical gender and linguistic relativity: A systematic review. *Psychonomic Bulletin and Review*, 26(6), 1767-1786. doi:10.3758/s13423-019-01652-3

Saliency in Second Language Acquisition: A Systematic Review

Saioa Cipitria (Vrije Universiteit Brussel), Georgia Knell (Vrije Universiteit Brussel), Esli Struys (Vrije Universiteit Brussel), Ludovic De Cuypere (Vrije Universiteit Brussel) and Alex Housen (Vrije Universiteit Brussel).

In the field of second language acquisition (SLA), saliency refers to the extent to which a linguistic element stands out from its context, and, hence, the likelihood of it being picked up by the learner's perceptual apparatus. The degree of saliency of a particular linguistic feature is believed to drive attention, essential for input (i.e., available stimuli) to become intake (i.e., noticed stimuli) and thus to acquire (i.e., cognitively register) it (Schmidt, 1990).

The concept of saliency in SLA has gained in prominence since the meta-analysis by Goldschneider and DeKeyser (2001), who identify relevant factors for the order of morpheme acquisition in English (e.g., phonological/orthographic substance, frequency, etc.). These factors comprise different aspects of saliency. Ellis (2016) classifies saliency as experiential (i.e., related to language experience), psycholinguistic (i.e., pertaining to the linguistic context and the learner), and perceptual (i.e., concerning the form). Empirical research on the role of saliency in L2 learning remains scarce, in particular concerning the manipulation of saliency as an independent variable (yet see contributions to Gass et al., 2018). This systematic review therefore explores the conceptualization and operationalization of saliency by reviewing research that directly explores and manipulates saliency in the context of L2 language acquisition by considering different types of saliency (i.e., experiential, psycholinguistic, and perceptual) and how they affect language processing. We extracted 993 references from two databases (Web of Science and Scopus), of which 88 studies were accepted for more detailed annotation. Other coding criteria include saliency type and the linguistic aspect under study (morphosyntax, lexicon, phonology, etc.). Based on cognitive theories in both SLA and cognitive psychology, we expect a positive correlation to emerge from our review of empirical studies between saliency, attention and ultimate acquisition. We further consider how this correlation is mediated by the various types of saliency. Finally, we discuss general methodological trends, strengths and shortcomings in research on saliency in SLA research as they emerge from our survey.

L'influence translinguistique dans les productions narratives des francophones en italien L2 et l'utilisation des particules de portée additives "anche" et "ancora"

Maria Rosaria D'Angelo (Doctorante, UMR7023-SFL/Université de Paris 8.- Paris.).

Notre étude porte sur des productions de récits d'images réalisées par des apprenants francophones de l'italien ayant trois niveaux interlinguistiques différents. Pour le recueil de données, nous avons utilisé une tâche narrative permettant d'analyser le marquage des relations additives (Benazzo & col. 2004). Nous disposons de quinze récits par niveau d'apprenant (débutant, intermédiaire et avancé) et de deux groupes de contrôle italien et français. Nous avons analysé les différences et/ou les ressemblances interlinguistiques entre l'italien et le français qui sont variables selon le niveau d'acquisition, sur la mise en œuvre en italien L2 d'adverbes dits « particules de portée » tels que « anche » et « ancora » dont les propriétés, qui jouent un rôle crucial du point de vue communicatif, concernent plusieurs niveaux linguistiques : sémantique, syntaxique et discursif. « Anche » et « ancora », considérés comme des particules additives créant des relations d'addition, sont abordés habituellement dans le cadre de la sémantique formelle (König 1991). Cependant, des études empiriques sur l'acquisition des phénomènes de portée au sein de l'approche fonctionnelle ont mis en évidence la nécessité de prendre en compte des unités contextuelles plus large que la phrase. Elles analysent ces phénomènes de portée dans des variétés initiées par Dimroth et Klein (1996) et ont contribué à éclairer le fonctionnement de ce type d'items en relation avec le discours dans lequel ils sont produits. Nos résultats montrent que le français LM influence les productions des apprenants en italien surtout au niveau de la structure syntaxique même si la signification de base de la particule n'est pas affectée. Nous avons finalement constaté des stratégies interlinguistiques de transfert aussi bien que des stratégies propres au système de l'interlangue dans les domaines de la syntaxe des particules, ce qui amène à des placements anormaux. On remarque notamment des phénomènes de transfert dans les contextes de portée à distance et à gauche, donc dans des contextes qui créent des problèmes de distribution, qui produisent évidemment des situations de détresse expressive. Par conséquent le transfert semble représenter un support en cas de difficulté, car il exprime la référence instinctive à des solutions connues pour résoudre cette difficulté même au niveau non débutants.

Bibliographie Benazzo, S., Dimroth, C., Perdue, C., & Watorek, M. (2004). Le rôle des particules additives dans la construction de la cohésion discursive en langue maternelle et en langue étrangère. *Langages* 155, pp. 76-105 Dimroth, C., & Klein, W. (1996). Les particules focales dans les variétés d'apprentissage : un cadre d'analyse et quelques exemples. *Zeitschrift für Literaturwissenschaft und Linguistik*, 104, 73-114. König, E. (1991). *The meaning of focus particles: a comparative perspective*. London: Routledge.

Secondary 1 pupils' spontaneous EFL oral acquisition: the advantage of modelization

Marie-Ange Dat (Nantes Université), Delphine Guedat-Bittighoffer (Université d'Angers) and Rebecca Starkey-Perret (Nantes Université).

The field of second language (L2) acquisition in guided settings regularly raises the issue of oral fluency. While a broad scientific consensus identifies automation as the key to L2 oral proficiency (DeKeyser, 2009; Paradis, 2009), the question of how to achieve this remains unanswered. From recommendations of using explicit knowledge (DeKeyser, 2003) to proposals to build skills in the classroom from a modeled language that draws on learners' internal grammar (Germain, 2018; Ellis & Wulff, 2020) by appealing to implicit memory and the process of proceduralization (Paradis, 2009), a wide range of hypotheses remains open. In our study context (2019-2022), which focuses on 179 French speakers beginning to learn English as a foreign language during the first year of middle school (11.3 years old), we wanted to measure their level of acquisition by comparing (T1 first month of the school year, T2 9 months later) two teaching methods during one school year. The first one represents prototypical foreign language teaching, i.e. an eclectic set composed of the action-oriented approach (CEFR) and explicit elements of the L2. The second is the Neurolinguistic Approach (also called NLA, Germain, 2018) developed in Canada. The primary principles of the approach are : modeling, implicit learning, primacy of oral and authentic situations of interlocution. We ask ourselves whether learning spontaneous speaking based on modelization and implicit memorisation in the context of extensive teaching (4h per week) is more effective in terms of spontaneous interaction than L2 more explicit approaches. The 179 pupils were in 8 classes: 4 experimental classes teaching with NLA, and 4 prototypical control classes. We assessed each pupil's ability to interact orally using the Oral Proficiency Interview (OPI) scale (Macfarlane & Montsion, 2016), the purpose of which is to assess speaking skills in real-life situations. A repeated measures ANOVA analysis shows that highly significantly ($p < 0.001$) pupils in classes that use modelization and implicit learning double their spontaneous interlocution scores compared to more prototypical classes that use explicit learning pathways.

References : DeKeyser, R. (2003). Implicit and explicit learning. In C. Doughty & M. Long (Eds.), *Handbook of second language acquisition* (pp. 313–48). Oxford, England: Blackwell. DeKeyser, R. (2009). Cognitive-psychological processes in second language learning. In M. Long & C. Doughty (Eds.), *Handbook of second language teaching* (pp. 119–38). Oxford, England: Blackwell. Ellis, N. C. & Wulff, S. (2020). Usage-based approaches to L2 acquisition. In VanPatten, B., Keating, G. D., & Wulff, S. (Eds.), *Theories in Second Language Acquisition: An introduction*. (pp. 63-82). New York & London: Routledge. Germain, C. (2018). *The Neurolinguistic Approach (NLA) for Learning and Teaching Foreign Languages – Theory and Practice*. Cambridge Scholars Publishing. MacFarlane, D., & Montsion, L. (2016). *Comparaison d'échelles de compétence orale*. Ministère de l'Éducation nationale et du Développement de la Petite Enfance du Nouveau-Brunswick. Paradis, M. (2009). *Declarative and Procedural Determinants of Second Languages*. Amsterdam: John Benjamin.

How do learner-internal and learner-external individual differences affect adolescent learners' L2 English speaking development? - A dense longitudinal study.

Vanessa De Wilde (Ghent University).

Researchers have investigated the role of individual differences in second language learning and found that learners' development is influenced by both internal differences (e.g. cognitive differences) and external differences (e.g. length of instruction) (Dörnyei, 2014; Muñoz, 2012). Many of these studies investigated second language learning in adults. Recently, some studies have also looked into the role of individual difference variables in language learning in young learners (e.g. De Wilde & Goriot, 2022; Leona et al., 2021) but only few studies have adopted a longitudinal approach (e.g. Pfenninger, 2022). In the present study we aim to analyze the role of several internal differences (such as cognitive differences and motivation) and external differences (such as out-of-school exposure and instruction) in L2 English speaking development in adolescent learners. In order to be able to capture real-life speaking development, we conducted a dense longitudinal study with frequent measurements of L2 speaking skills. Learners in the first year of secondary school (11 to 13 years old, n=48) did a speaking task every school week from September to May (25 weeks). The study took place in Flanders and the Netherlands. The participants came from three class groups which were different in terms of instruction: one group had received L2 English instruction in primary school, the second group had just started with formal L2 English lessons and the last group did not have any formal English lessons yet. At the start of the study the participants did multiple tasks looking into individual differences (prior knowledge of English, Dutch vocabulary knowledge, reasoning skills, working memory capacity, motivation, and out-of-school exposure). Generalized additive mixed models (GAMMs) were used to model (non-linear) learning trajectories over time and to investigate the role of internal and external differences in this development. Results show that learners' speaking scores are influenced by time and prior L2 English knowledge. Initial differences in learners' prior knowledge seem to be the main predictor of L2 English speaking skills. Detailed results and implications of the study will be discussed during the presentation and suggestions will be done for future studies.

References

- De Wilde, V., & Goriot, C. (Eds.). (2022). *Second Language Learning Before Adulthood: Individual Differences in Children and Adolescents* (Vol. 65). Walter de Gruyter GmbH & Co KG.
- Dörnyei, Z. (2014). *The psychology of the language learner: Individual differences in second language acquisition*. Routledge.
- Leona, N. L., van Koert, M. J., van der Molen, M. W., Rispen, J. E., Tijms, J., & Snellings, P. (2021). Explaining individual differences in young English language learners' vocabulary knowledge: The role of Extramural English Exposure and motivation. *System*, 96, 102402.
- Muñoz, C. (Ed.). (2012). *Intensive exposure experiences in second language learning*. *Multilingual Matters*.
- Pfenninger, S. E. (2022). Emergent bilinguals in a digital world: A dynamic analysis of long-term L2 development in (pre) primary school children. *International Review of Applied Linguistics in Language Teaching*, 60(1), 41-66.

Investigating native language attrition of late Spanish/English bilinguals in the UK: the case of aspectual marking in Spanish

Laura Dominguez (University of Southampton), E Jamieson (University of Southampton), Glyn Hicks (University of Southampton) and Monika S. Schmid (University of York).

We investigate possible attrition (i.e. changes in the native grammar) of 30 late Spanish-English bilinguals who have settled in the UK for over 15 years. Specifically, we explore whether aspectual marking is a vulnerable area for the native grammar of these speakers using a multi-method approach to data collection and analysis. We chose this grammatical property as it is predicted to undergo attrition by the Attrition via Acquisition (AvA) model (Hicks & Domínguez 2020) as Spanish and English share the same syntactic/semantic features, but they map differently onto morphosyntactic forms. That is, Spanish speakers do not need to acquire a new feature/structure when using these verb forms in English, but they need to determine in which contexts each aspectual form is grammatical or ungrammatical. Attrition has also been attested (e.g. Cuza (2010)) for this construction for a group of Spanish-English bilinguals in the USA.

As Table 1 shows, Spanish and English each have two ways to express an event in the present (whether the event is ongoing or habitual) but they differ in the specific contexts in which each of these forms can be used. In English, the simple present form is used for generic/habitual actions, while the periphrastic *be+ing* form is used in the progressive/ongoing aspect. In Spanish, there is an overlap in the aspectual interpretations of the two verb forms: the corresponding *estar+ndo* form is also used in the progressive/ongoing aspect, but the simple present can be used for both generic/habitual actions and the progressive/ongoing aspect.

English Spanish Ongoing Now she is running. Ahora ella está corriendo. Ahora ella corre.

Habitual Every day she runs. Todos los días ella corre.

Table 1. Distribution of verbal forms to express ongoing and habitual actions in English and Spanish

We predicted that attrition, if attested, would manifest in a preference for the *estar+ndo* form over the simple present in ongoing situations as this is the only form available in English in this context.

We elicit three types of data using a self-paced reading task (online), an acceptability judgment task (offline), and sociolinguistic interviews (oral). These results were compared with similar data from a group of monolingual Spanish speakers who also took part in the study. Modelling the results of the experimental tasks shows no overall attrition at group level. However, two individuals show an English-like pattern in the judgment task. The oral production data also show that, again at an individual level, some speakers may undergo attrition of their aspectual system as a result of the influence of English manifested as a preference for the periphrastic construction in representing ongoing aspectual interpretations in the L1. These results support the idea that, in general, morphosyntactic features of an L1 are robust even after extensive exposure to an L2; however, attrition is possible for some speakers at least when the underlying features are shared.

References:

Cuza, A. (2010). On the L1 attrition of the Spanish present tense. *Hispania*, 256-272.

Hicks, G., & Domínguez, L. (2020). A model for L1 grammatical attrition. *Second Language Research*, 36(2), 143-165

Investigating sensitivity to partial rules in English wh-questions: Japanese vs. Vietnamese L2 learners

Nigel Duffield (Konan University), Ayumi Matsuo (Kobe College) and Trang Phan (University of Languages & International Studies, Vietnam National University - Hanoi).

Our project is concerned with L2 learners' sensitivity to exceptions to the general rules of (short) wh-movement and complementation in English (Chomsky 1957, et seq). (In contrast to the attention paid to constraints on long-distance movement – e.g., L1: Guasti, 1994; L2: Martohardjono, (1993), White & Juffs, 1998 – exceptions to short-movement have been neglected hitherto). The specific focus is on (1) Wh-interactions with finiteness (She knew [why she should eat more vegetables]. ~ She knew [*why to eat more vegetables]; (2) Wh-interactions with negative adverbials (“?What₁ under no circumstances₂ should₃ you ... ~ ??Under no circumstances₂ what₁ should₃ you... ~*Under no circumstances₂ should₃ what₁ you ... ~ ?What₁ should₃ under no circumstances₂ you ... give to a dog.”; cf. Haegeman 2012); (3) Wh-interactions in ‘Sluice-Stranding contexts’ (“She asked to attend to something, but I can’t remember what/*what to/??to what?”; Culicover 1998); (4) Raising vs. Control complements (“Amy is *uncertain/certain to lock up. Amy is uncertain/?certain how to lock up.”; cf. Yoshimura et al., 2016). This phase of the project compared Vietnamese and Japanese L2 learners of English: though both L1s are wh-in situ languages, structural commonalities between Vietnamese and English predict an advantage for Vietnamese learners in conditions that probe knowledge of the “left periphery”.

Method. The current experiment used two formats for eliciting judgments: a pen-and-paper judgment task (P-AJT); and an online version, involving the same materials (O-AJT). Each participant was presented minimal contrast sentences, differing only in the critical contrast, where the former are designated ‘grammatically acceptable’, the latter ‘marginal’ or ‘grammatically unacceptable’ in the literature. Participants were asked to rate the acceptability of each trial on a Likert scale from -2 (‘unacceptable’) to +2 (‘perfectly natural’).

Items. The stimulus items comprised 348 sentences across four conditions, plus 16 control items (8 minimal pairs, clear grammaticality contrasts). Using a modified Latin Square, six randomized-trial subsets were prepared, each comprising 58+16 items.

Participants. 24 adult native-speakers and 166 L2 learners participated in the study. However, a high percentage of the L2 learners failed to distinguish acceptable from unacceptable items in the control condition [5]. Results from the remaining participants {56 VN; 29 JA} were entered into the main analysis.

Results. Preliminary descriptive results from the first two conditions (plus control condition) bear out our predictions: on each of the specific contrasts, Vietnamese L2 learners’ judgments pattern with those of native-speakers, albeit in attenuated form, whereas Japanese L2 learners either fail to distinguish between grammatical and ungrammatical sentences, or else diverge sharply from both groups in preferring a clearly marked option (#2-213 order). Full results and inferential statistics will be presented at conference.

Selected References Culicover, P. (1999). *Syntactic Nuts: hard cases, syntactic theory, ...* Oxford: OUP. Duffield, N. (2018). *Reflections on Psycholinguistic Theories...* Cambridge: CUP. Haegeman, L. (2012). *Adverbial Clauses, Main Clause Phenomena...* Oxford: OUP. White, L., & Juffs, A. (1998). Constraints on wh-movement In S. Flynn et al. Yoshimura, N., Nakayama, M., et al. (2016). ‘Control and Raising constructions in early L2 English Acquisition’, *Second Language*, 15. (pp. 53-76).

What Do Elicited Imitation Tasks for Young EFL Learners Measure?

Akiko Eguchi (Ritsumeikan University) and Remi Murao (Nagoya University).

Elicited imitation (EI), a testing technique that makes the participants listen to a series of sentences and repeat them verbatim, has been widely used to assess a second language (L2) from various perspectives (e.g., Yan et al., 2016 for a review). The underlying assumption is that EI tasks require learners to process and reconstruct the stimulus sentences with their internal grammar and are applicable to a range of proficient learners by adjusting the stimuli. If this assumption is evidenced, EI can be a promising method to examine scant L2 grammatical knowledge in young learners as repeating phrases and sentences is a familiar activity in classrooms. Compared with numerous L2 adult studies, however, research on using EI with young L2 learners is scarce, and is limited to testing the effects of particular pedagogical interventions by pre-post comparison (e.g., Campfield & Murphy, 2014). Therefore, it is unclear what exactly EI tasks for young learners measure and to what extent primary memory—which is beneficial to rote repetition and differentiated from speech-processing memory—affects EI performance.

This study clarified what EI tasks measure for young L2 learners by focusing on three variables: syntactic knowledge, vocabulary knowledge, and primary memory. We created 24 sentences for EI tasks using familiar words and various syntactic structures (e.g., “SVO,” “WH do SVO”) that appeared in their textbooks but adjusted to longer sentences to avoid rote repetition ($M = 9.4$ syllables). EI performance was assessed using a 6-point-scale scoring, assuming that L2 sentential knowledge develops from a single word to a syntactic phrase to a sentence.

A total of 138 Japanese elementary students (Grade 4 and 5: aged 9–11), who were learning English as a foreign language (EFL), participated in four tasks: (1) EI tasks, (2) untimed sentence correction tasks (SCT) (e.g., *I soccer play. *What color you like?) for syntactic knowledge, (3) vocabulary size test (VST) (Sato, 2019), and (4) digit span task (DST) for primary memory capacity. To determine the extent to which variance in EI scores could be explained by the individual variables, a linear mixed effects model was performed. The results of the final best-fitting model revealed a significant main effect of DST (estimate = 1.72, $t = 2.93$, $p < .01$) and the interaction between SCT and VST (estimate = 7.21, $t = 4.12$, $p < .001$). This result suggests that EI can successfully measure young L2 learners’ internal grammar integrated with vocabulary knowledge while controlling for the potential confound of rote repetition.

Campfield, D., & Murphy, V. A. (2014). Elicited imitation in search of the influence of linguistic rhythm on child L2 acquisition. *System*, 41, 207–219. Sato, T. (2019). Developing a vocabulary size test for Japanese elementary school students. *JES Journal*, 19, 130–146. Yan, X., Maeda, Y., Lv, J., & Ginther, A. (2016). Elicited imitation as a measure of second language proficiency: A narrative review and meta-analysis. *Language Testing*, 33, 497–528.

Acquisition of quantifier-negation scope and grammatical morphemes by Chinese learners of Korean

Shuo Feng (Peking University), Meiling Yu (Peking University), Nai'An Ren (Peking University) and Kailun Zhang (Peking University).

The present study aims to explore the acquisition of universally quantified objects under negation and grammatical morphemes in Korean by Chinese learners. In English, (1) has the not>all reading as the dominant reading. Chinese allows negation and quantifier NP each to precede or follow the other as in (2a-2b), with different scope interpretations. Korean, a SOV language, allows the quantifier NP to precede negation and the preferred reading for (3a) is the all>not reading. However, when the morpheme changes from the object case marker -(l)ul in (3a) to the contrastive focus marker -(n)un in (3b), the preferred reading for (3b) becomes the not>all reading (Park & Dubinsky, 2020). Previous L2 studies found that L2 learners were under strong L1 influence in acquiring quantifier-negation scope and Korean morphemes were challenging to acquire (e.g., Ahn, 2015; Chung & Shin, 2022). However, previous work has focused on a limited set of structures (e.g., quantified subjects) and little research has been carried out with participants with L1s other than English.

(1) English: Tom didn't send all the letter. (not>all) (2) Chinese: a.

Tangmu meiyou ji suoyou de xin. (not>all) Tom not-have send all Poss letter
b.

Tangmu suoyou de xin dou meiyou ji. (all>not) Tom all Poss letter all not-

have send (3) Korean: a. Tom-ka motun phyenci-lul an ponay-ss-ta. (all>not) Tom-

Subj all letter-Do not send-Pst-Decl b. Tom-ka motun phyenci-nun an ponay-ss-

ta. (not>all) Tom-Subj all letter-Cf not send-Pst-Decl

Method: Forty-four Chinese learners of Korean and 30 native Korean speakers completed a self-paced reading task (SPRT), a Korean proficiency task and an acceptability judgment task (AJT). Take the SPRT as an example. Each trial in the SPRT consisted of four steps (see Figure 1). Different scope readings were manipulated in Step 2 where the pictures showed how many objects were affected by an action (either 3 out of 5 or none). Test stimuli were constructed in a 2 x 2 design with Context (2 levels: partitioned set and full set) and Morpheme (2 levels: -(l)ul and -(n)un). Four counterbalanced presentation lists were created using a Latin square block design. The task contained 16 targets, 10 controls and 30 fillers.

Findings & Conclusion: NNSs were native-like in interpreting quantifier scope in negated sentences with -(l)ul. NNSs' Korean proficiency played an important role in interpreting quantifier scope, but the influence of proficiency in acquiring -(n)un as a contrastive focus marker was restricted. Although NNSs were not native-like in interpreting sentences with the two morphemes in the offline task, they showed native-like processing pattern in the SPRT. The divergence of NNSs' performance between the SPRT and the AJT is further discussed in light of NNSs' less robust encoding of morphological information and increasing processing load in the AJT.

References: Ahn, H. (2015). *Second Language Acquisition of Korean Case by Learners with Different First Languages*. Doctoral dissertation, University of Washington. Chung, E. S., & Shin, J. A. (2022). Native and second language processing of quantifier scope ambiguity. *Second Language Research*. Jiang, N., Novokshanova, E., Masuda, K., & Wang, X. (2011). Morphological congruency and the acquisition of L2 morphemes. *Language Learning*, 61(3), 940-967. Park, K., & Dubinsky, S. (2020). The effects of focus on scope relations between quantifiers and negation in Korean. *Proceedings of the Linguistic Society of America*, 5(1), 100-106.

Figure 1. Sample display in the partitioned set context in the SPRT (see the attached pdf.)

Priming structural similarity in L1-L2 interactions: The use of Present Perfect vs. Simple Past by German learners of English

Regina Grund (University of Würzburg) and Barış Kabak (University of Würzburg).

The contrast between the English Present Perfect (PP) vs. Simple Past (SP) often poses a challenge for German learners of English, due to not only fuzzy boundaries between the two, but also the presence of structurally similar albeit not always functionally equivalent constructions in German: Perfekt (er hat gekocht 'he has cooked') and Präteritum (er koch-te 'he cook-ed'). While Perfekt is the most commonly used past reference in spoken German, its contextual equivalent in English may correspond to either PP or SP. Although such overlaps might predict an overuse of PP in SP contexts, recent research has shown the opposite: German learners underuse PP in comparison to English monolinguals, possibly due to PP being more complex than SP (e.g., Fuchs et al. 2016). Such corpus-based studies however mainly focused on frequencies and did not investigate to what extent the L1-L2 overlap may influence target-like suppliance in L2 English.

Here, we investigated potential priming effects of the German Perfekt and Präteritum on the use of PP and SP in L2 English. L1 German university students of English (n=24) were given a written gap-filling task with 24 translanguaging exchanges, each containing a German turn followed by an English response. The latter included a gap requiring the tense that was either structurally similar to the primed tense in German (prime-response match: Perfekt-PP or Präteritum-SP) or dissimilar (prime-response mismatch: Perfekt-SP or Präteritum-PP).

Mismatch (Perfekt-SP) example: • Was hast du denn so gemacht? [Perfekt primed] • Last week, I ____ (to make) a cake. [SP expected]

While an increase in the correct suppliance due to syntactic copying was predicted in the match conditions, syntactic copying was expected to provoke incorrect suppliance in the mismatch conditions. Thirteen (near-)native lecturers of English grammar served as controls to verify whether the English turns were PP- or SP-lenient.

Preliminary results revealed that the error rates were 10% for the match vs. 28% for the mismatch condition. When the German prompt matched the expected English tense, there were more errors in the Präteritum-SP (13%) than in Perfekt-PP (6%) context, suggesting facilitative priming effects for Perfekt. When the prime and the expected tense mismatched, however, 33% of answers in Präteritum-PP were erroneous while the error rate in Perfekt-SP was 23%. Error analyses revealed that while the majority of errors in Präteritum-PP were unrelated to priming (56%), the most frequent error in Perfekt-SP was the incorrect suppliance of PP (76%), implying Perfekt's unfavorable effect.

We suggest that a primed Perfekt strongly activates PP in English and is more likely to induce CLI effects than Präteritum. As such, corpus-based SLA studies should be complemented with experimental approaches tapping possible L1-L2 interactions in order to offer a more comprehensive account of equivalence classification in L2 grammar learning, with broader teaching implications where L1 priming can be leveraged for dynamic noticing and feedback.

Fuchs, R., et al. (2016). The present perfect in learner Englishes: A corpus-based case study on L1 German intermediate and advanced speech and writing. V. Werner et al. (Eds.), *Re-assessing the present perfect*. Mouton.

The interpretation of verbal moods in Spanish: A close replication of Kanwit and Geeslin (2014)

Aarnes Gudmestad (Virginia Tech), Amanda Edmonds (Université Côte d'Azur), Carlos Henderson (University of Lund), Christina Lindqvist (University of Gothenburg) and Şebnem Yalçın (Bogazici University).

Studies on sociolinguistic variation in a second language (L2) demonstrate that learners develop similar sensitivity to the linguistic and extralinguistic factors that influence variable patterns in their input (Geeslin & Long, 2014). This research has largely focused on learner production or on the selection of variable forms through preference tasks, meaning that little is known about how learners develop the ability to interpret variable structures. In the first study of its kind, Kanwit and Geeslin (2014) investigated the interpretation of a variable structure by L2 learners, focusing on the interpretation of verbal moods (the subjunctive and indicative) in Spanish adverbial clauses. They analyzed L2 development using cross-sectional data and found distinct patterns among lower- and upper-level learners. Crucially, Kanwit and Geeslin (2014) provide researchers with a methodological model for investigating the interpretation of variable forms. Given that research on the interpretation of sociolinguistic variation is in early stages, we do not yet know whether Kanwit and Geeslin's (2014) findings generalize to other learner populations. One approach to verifying research findings is to replicate key studies (Porte & McManus, 2019). We thus conducted the first close replication of Kanwit and Geeslin (2014), modifying one variable—the participants' first language (L1). We collected data from learners of Spanish with three different L1s. The L1s chosen—French, Swedish, and Turkish—are typologically distinct and have yet to receive much attention in variationist second language acquisition (SLA). This means that investigating these understudied language combinations allows us to explore the generalizability of previous results.

We collected data from 40 participants in each of the three L1 groups. All participants were enrolled in undergraduate-level Spanish courses. We used Kanwit and Geeslin's (2014) interpretation task and grammar test and designed our own background questionnaire, as it was crucial for us to obtain detailed information on the participants' language background. Other than adding a new independent variable (participant's L1), we followed Kanwit and Geeslin's (2014) data coding and analysis procedures. The dependent variable was the response for each item on the interpretation task. There were seven independent variables: mood, verbal morphology regularity, clause order, adverb/conjunction, proficiency (measured by the grammar-test score), course level, and participant. We analyzed the data quantitatively and found that learners' interpretation of verbal moods in Spanish differed as a function of their L1. Our study advances knowledge in SLA in three ways. First, we respond to calls for replication studies in SLA (Porte & McManus, 2019). Second, we contribute to the expansion of variationist SLA by conducting much-needed work on the interpretation of variable forms. Third, our study impacts variationist SLA and SLA more broadly by expanding the L1-L2 combinations investigated.

References Geeslin, K. L., & Long, A. Y. (2014). *Sociolinguistics and second language acquisition: Learning to use language in context*. Routledge. Kanwit, M., & Geeslin, K. L. (2014). The interpretation of Spanish subjunctive and indicative forms in adverbial clauses: A cross-sectional study. *Studies in Second Language Acquisition*, 36, 487-533. Porte, G., & McManus, K. (2019). *Doing replication research in applied linguistics*. Routledge.

Investigating the linguistic and social effects of the first year of schooling on the grammar of child heritage speakers: focus on Polish heritage children in the UK

Anna Hart (University of Southampton), Laura Dominguez (University of Southampton) and E Jamieson (University of Southampton).

Three main accounts have been proposed to explain heritage speaker acquisition: incomplete acquisition (Montrul, 2008; Polinsky, 2006; Silva-Corvalan, 2003), attrition (Polinsky, 2011) and parental input effects (Pascual y Cabo, 2018; Montrul & Sanchez-Walker, 2013; Pires & Rothman, 2009). Few studies have focused on children heritage speakers (Laskowski, 2009; Montrul 2016) although this group of speakers experience a major shift in their linguistic input and social context around the age of 4: they start school in the majority language. In this study we focus on the acquisition of Case by a group of 30 Polish heritage speaker children (HSC) in England in order to try to tease apart these three possible accounts. Starting school is a crucial event in the life of HSC: it signals the exposure to the majority language and a significant change in their linguistic input; it also signals a change in their social environment. Hence, we investigate whether starting school in the majority language is a key event with significant consequences for heritage language acquisition, i.e. whether the heritage children's grammatical system is complete or not before they enter mainstream education in the majority language, and/or whether any changes can be observed after one year of schooling. We focus on a group of 30 young Polish heritage children in the UK and investigate their development of Case marking during their first year in primary school. Polish marks Case on nouns with specific morphological forms in a consistent and uniform fashion which is not found in English. Previous studies have shown that adult Polish heritage speakers have problems with particular Case markings (e.g. locative) (Kozmińska, 2015; Wolski-Moskoff, 2019), but not others (i.e. nominative). We tested the productive command of these forms using a narrative retelling task and appropriate grammatical knowledge using an acceptability judgement task. Sociograms were used to investigate the heritage children's social networks. The children were tested twice, at the start and at the end of their first year in primary school. The Bilingual Language Experience Calculator (Unsworth, 2013) was used to calculate how much input (quantity and quality) they receive. Out of the three tested cases, (nominative, genitive and locative), nominative remains 'intact' throughout the first year. There is evidence of attrition for genitive, as 20% of the children show lower accuracy with this case that was 100% accurate at the start. Locative is the most challenging case as the mean accuracy rate is already lower at the beginning of the school year than for the other two cases (92.23%), but it decreases further as the year progresses to 80.54% which points to both incomplete acquisition and attrition for this case. These results reveal that attrition of certain morphosyntactic forms does occur after only 9 months in school but that this change, which is only attested in some cases and for some children, is modulated by both linguistic and social factors.

Crosslinguistic influence and proficiency in L2 and L3 knowledge of aspect in Japanese

Akbar Nadjar Hendra (University of York), Leah Roberts (University of York) and Heather Marsden (University of York).

This study investigates prior-language influence in the acquisition of aspectual interpretation. We compare L1-Indonesian L2-English L3-Japanese speakers with L1-English L2-Japanese. Aspect is grammaticized in Japanese and English, whereas it is marked lexically in Indonesian. Previous research (Roberts & Liska, 2013, 2019) showed that acquisition of grammaticized aspect in L2 was facilitated by grammaticized aspect in the L1. Further, Gabriele (2009) showed that L1-English L2-Japanese speakers could acquire interpretive properties of Japanese aspect that differ from English. The present study expands the focus to L3 speakers, investigating the effect of an L1 with non-grammaticized aspect and L2 with grammaticized aspect on L3 acquisition of aspect.

The study focuses on interpretation of the Japanese “-te iru” verbal morphology. The canonical interpretation is the progressive (e.g., “tabe” [eat] + “te iru” = is eating). However, with achievement verbs “-te iru” has a resultative interpretation, unlike English “be+ing” (1). In further contrast to English, “-te iru” is incompatible with a futurate meaning, and hence infelicitous with futurate adverbs (2) where the simple non-past is required. Indonesian expresses the different aspectual meanings with distinct pre-verbal adverbials.

1. Hikouki-ga kuukoo-ni tsui-te iru Plane-NOM airport-at arrive-TE IRU “The plane has arrived (*is arriving) at the airport.”

2. Hikouki-ga sugu-ni shuppatsu suru / *shi-te iru Plane-NOM soon departure do / do-TE IRU “The plane is departing soon.”

If L2/L3 acquisition of Japanese “-te iru” is influenced by English, then the semantics of “be+ing” may (i) obstruct acquisition of the resultative interpretation, and (ii) lead to non-target-like acceptance of futurate interpretations. However, if influence arises from Indonesian, while acquisition of the licit meanings of “-te iru” could be hindered, a futurate interpretation should not be posited.

Twenty-nine native Japanese, 20 L1-English L2-Japanese and 22 L1-Indonesian L2-English L3-Japanese speakers participated in the experiment. The L2/L3 speakers’ Japanese proficiency was intermediate–advanced, with similar proficiency test score distributions in each group. Participants completed two tests of knowledge of “-te iru”. An adapted replication of Gabriele’s (2009) sentence compatibility task included accomplishment verbs in the “-te iru” form following contexts that depicted completed events (i.e., grammatical resultative), or ongoing events (ungrammatical). An acceptability judgement task (AJT) compared grammatical resultative “-te iru” with ungrammatical “-te iru” after futurate adverbs. Ordinal mixed-effects models of the ratings for each task showed that both non-native groups differentiated between grammatical and ungrammatical interpretations, but the magnitude of differentiation was significantly smaller than in the native group ($p < .001$). Notably, the L2/L3 ratings on ungrammatical conditions were higher than in the native group. Follow-up models on the L2/L3 groups with proficiency score as a predictor showed a weak interaction in the AJT of proficiency with condition (furate v. resultative) and group ($b = -0.065$; $p = .076$), due to lower ratings for the futurate in the L3 group than in the L2 group, as proficiency increased. We argue, drawing on models of L3 acquisition, that this relatively more target-like behaviour on the futurate by the higher proficiency L3 speakers suggests an effect of L1 Indonesian influence during L3 development.

Cross-linguistic structural priming as a mechanism of cross-linguistic influence: Asymmetrical effects of L1 activation and inhibition

Carrie Jackson (The Pennsylvania State University) and Holger Hopp (TU Braunschweig).

Cross-linguistic influence (CLI) is a key characteristic of bilingualism; however, the mechanisms of CLI are still not clear. Recently, several proposals to grammatical CLI have argued that cross-linguistic priming is a key mechanism of CLI in child bilinguals (Nicoladis, 2012; Serratrice, 2016, 2022). We examine the scope of priming as a mechanism of cross-linguistic influence (CLI) in bilinguals by aiming to boost CLI through priming in early and late L2 learners. In two cross-linguistic structural priming studies with 32 less-proficient adolescent (Study 1) and 60 more highly proficient adult German-English learners (Study 2), we assess whether cross-linguistic priming enhances CLI in English for well-formed (adverbial fronting; 1a), dispreferred (TP-LP orders; 1b) and ungrammatical structures (verb-second; 1c; ungrammatical verb raising; 1d). All word orders on the left in (1a-d) are well-formed in German. (1) a. Last week the man was/ate in the garden. vs The man was/ate...(adverbial fronting) b. #The man was/ate last week in the garden. vs in the garden last week. (TP-LP) c. *Last week was/ate the man in the garden. vs Last week the man was/ate ... (V2) d. *The man ate often meat. vs The man often ate ... (Verb raising) Participants were primed with the German word orders (left in (1a-c)) in German, and were then prompted to produce a sentence in English. We measured which of the two respective word orders in (1a-c) they produced in English. L2 learners in both studies showed CLI from German in their English sentence production in a baseline task (Figures 1&2). For grammatical L1-L2 word orders in L2 English, i.e. fronting in (1a) in the priming task, logistic mixed effects modelling shows that less-proficient learners in Study 1 also exhibited short-term cross-linguistic priming, which extended to longer-term priming among the more proficient learners in Study 2 in a posttest. However, there was no evidence that cross-linguistic priming increased the use of dispreferred (1b, TP-LP orders) or ungrammatical L1-based word orders in L2 English (1c, V2) in either study. Rather, as seen in Figures 1&2, the overall production of these word orders decreased from baseline via the priming task to the posttest. The proportion of ungrammatical word orders not contained in the priming task (1d, Verb raising) did not change from baseline to posttest. Together, these results suggest that, while cross-linguistic priming leads learners to increase the use of shared, grammatical structures, it leads to the inhibition of non-shared, ungrammatical structures in L2 production. We conclude that cross-linguistic priming has asymmetrical effects on CLI of grammatical and ungrammatical L1-based structures in the L2 in that the Interlanguage grammar modulates activation and inhibition of the L1. References Nicoladis, E. (2012). Cross-linguistic influence in French–English bilingual children's possessive constructions. BLC Serratrice, L. (2016). Cross-linguistic influence, cross-linguistic priming and the nature of shared syntactic structures. LAB Serratrice, L. (2022). What can syntactic priming tell us about crosslinguistic influence? In *Syntactic priming in language acquisition: representations, mechanisms and applications*. Benjamins.

It's about time: Exploring refugee L2 learners' narratives of time through the lens of agency

Julia Jakob (Faculty of Education, University of Cambridge).

**

"We wrote a mock exam [in German class today]. And it was very easy. I asked the teacher if I could go home. Because attendance is mandatory. [...] She said yes [this time]. [...] But [normally] it's a waste of time. I have to wait until the others are done. And I can't do anything."

**

Abud's expressed frustration over his lack of agency to fill his time in ways he considers useful or rewarding is illustrative of the experiences of many refugees. While narratives of time have been explored in sociology and forced migration studies, they remain underresearched in SLA. This is despite data from forced migration studies showing the potential relevance of this phenomenon for how refugee learners feel able to invest in L2 learning.

Drawing on qualitative data from semi-structured interviews conducted during five months of fieldwork with ten young men who live in Austria as refugees and engage in formal learning of German as an additional language, this paper will discuss participants' narratives of time that are associated with their experience of learning German as an additional language.

In particular, the paper builds on Schlote's (1996) work in the field of critical sociology of time to interrogate how investment in L2 learning is affected by:

- the concession of control over one's own time to other actors, such as teachers and the state, in the L2 classroom, the broader Austrian refugee L2 education system, and beyond - the expressed inability to plan for an often uncertain future in Austria, e.g. due to long waiting times for asylum, L2 courses, or a non-permanent legal status of protection - the experiences of (partial) autonomy in structuring one's own time under consideration of both the above and individual affordances and challenges, with a particular focus on unstructured time in L2 courses, independent L2 study, and opportunities for L2 use

The subjective experience of time intersects with L2 investment as conceptualized by Darwin & Norton (2015) on several levels. Firstly, narratives of time need to be understood within systemic patterns of control that form a structure that may, to an extent, be challenged and resisted. Secondly, self-positioning and positioning by others interact with narratives of time through notions of productivity and the pace of L2 learning. Thirdly, an unclear future may render the potential benefits of L2 learning uncertain.

The findings from this small-scale study indicate the need for future research into L2 learners' narratives of time in (forced) migration contexts. Due to the relevance for issues of L2 learners' investment, a better understanding of the subjective experiences of time can inform future critical evaluation of integration policymaking and L2 course design and pedagogy.

References:

Darvin, R., & Norton, B. (2015). Identity and a Model of Investment in Applied Linguistics. *Annual Review of Applied Linguistics*, 35, 36–56. <https://doi.org/10.1017/S0267190514000191>

Schlote, A. (1996). *Widersprüche sozialer Zeit: Zeitorganisation im Alltag zwischen Herrschaft und Freiheit*. VS Verlag für Sozialwissenschaften.

Can adults learn L2 grammar after prolonged exposure under incidental learning conditions?

Panagiotis Kenanidis (Friedrich-Alexander-Universität Erlangen-Nürnberg), Ewa Dabrowska (FAU Erlangen), Miquel Llompart (Universitat Pompeu Fabra) and Diana Pili-Moss (Leuphana University of Lüneburg).

While late second language (L2) learning is assumed to be largely explicit, there is some evidence that adults are able to acquire the grammar of novel languages under incidental learning conditions (Rebuschat et al., 2021; Ruiz et al., 2018). In two experiments, we revisit the question of whether adults can learn grammar incidentally and investigate whether word order and morphology are susceptible to incidental learning to the same degree. Additionally, we investigate the extent to which learners' first language (L1) background and metalinguistic awareness of the target grammatical structures can modulate L2 grammar learning outcomes.

In experiment 1, forty-one monolingual speakers of English took part in a five-session online study, during which they were exposed to Kepingalo, an artificial language that consisted of nouns, verbs and adjectives and had case marking and variable word order: a canonical Subject-Object-Verb order and a non-canonical Object-Subject-Verb order (1). First, participants were explicitly trained on the nouns of the language. This was followed by a two-alternative forced-choice task (2AFCT) consisting of two blocks. In the first block, participants received vocabulary training while being incidentally exposed to grammar. Two videos, each showing two aliens performing an action, were presented, accompanied by a sentence corresponding to one of them (Figure 1). The videos differed in 1) one of the aliens, or 2) the color of an alien or, 3) the action performed. Block 2 served as a grammatical comprehension test. Here, the two videos differed in that the agent/patient roles were reversed, and no feedback was given. The 2AFCT was repeated in the first four sessions. In session 5, grammatical knowledge was further assessed through a final grammatical comprehension test and a grammaticality judgment task (GJT) including both word order and case marking violations. Despite extensive exposure to input, and although performance on vocabulary increased significantly across sessions, learners' grammatical comprehension showed little improvement over time, and this was limited to Subject-Object-Verb sentences only. Furthermore, participants were significantly better at detecting word order than case marking violations in the grammaticality judgment task (Figure 2).

Experiment 2 further increased the amount of incidental exposure to the artificial language by providing two additional sessions and was conducted with native speakers of German. Moreover, testing in the last session was followed by the administration of a post-test awareness questionnaire. Leveraging their prior experience with a morphologically richer L1, participants showed better learning of both word order and case marking compared to their L1 English counterparts, yet continued to display substantial difficulties with the latter (Figure 2). Furthermore, grammar learning was found to be contingent on the learners' level of awareness.

Taken together, the results of the two experiments underscore adult learners' difficulty with case marking and point towards the presence of a threshold in incidental L2 grammar learning, which appears to be tightly linked to prior L1 experience (Ellis, 2006; MacWhinney, 2005). In addition, our findings continue to highlight the facilitative role of awareness on L2 learning outcomes (Andringa, 2020; Schmidt, 2012).

Testing the interface of implicit and explicit L2 grammar knowledge and their reciprocal relationship: A one-year longitudinal study

Kathy Kim (Boston University).

In recent years, SLA researchers have suggested that the association between explicit (conscious) and implicit (unconscious) second language (L2) knowledge is reciprocal; that is, awareness (or explicit knowledge) not only facilitates the development of implicit knowledge but can also be a product of implicit knowledge (Ahn, 2020; Andringa, 2020; Kim & Godfroid, 2023). In this project, I build on this line of research on the interface question and explore the extent to which the relational strengths between explicit and implicit L2 knowledge change over the course of one year. Do explicit and implicit knowledge remain reciprocal equally after 6 months and 12 months of using the L2 in an immersion context? I use an autoregressive cross-lagged model (ACLM) to address these questions.

One hundred forty-nine international students at an English-medium university completed five linguistic tests that measured their explicit and implicit knowledge of L2 English. The untimed written grammaticality judgment test (GJT) and metalinguistic knowledge test (MKT) served as measures of explicit L2 English knowledge. The timed written GJT, oral production, and elicited imitation were administered as implicit L2 English knowledge measures. These measures were provided at three timepoints over a year (T1: January–February, T2: April–May, T3: November–December).

Across the three waves, accuracy scores improved in most measures, with MKT exhibiting the steepest improvement (40% increase in accuracy from T1 to T3) and oral production showing the least (4% increase in accuracy from T1 to T3). The results of two-factor CFA models—that were performed separately for each timepoint data—suggested that all measures meaningfully explained their corresponding constructs at each timepoint. Building on these CFA models, I am currently evaluating the results of an autoregressive cross-lagged model to explore the causal relationship between the constructs. Preliminary ACLM findings suggest that the strongest predictor of both implicit AND explicit knowledge at time 3 may be the previous level of implicit knowledge and not explicit knowledge. Results suggest that the development of advanced linguistic competence is composed of dynamic interaction between explicit and implicit knowledge with implicit knowledge playing a key role.

A longitudinal exploration of the effects of singing on L2 beginner pronunciation

Vita Kogan (University College London) and Nadezda Bragina (Queen Mary University).

It is not uncommon for adult L2 learners to have accented production even at the late stages of L2 acquisition when all other linguistic systems are mastered (Reiterer et al., 2011). Previous research has shown a connection between musical training and improved L2 pronunciation (Christiner & Reiterer, 2017; Nardo & Reiterer, 2009; Seither-Preisler et al, 2014). However, most of the studies conducted on the topic have focused on laboratory settings and cross-sectional design. Thus, little is known about the effect of regular singing in the instructional SLA, especially with lesser-taught non-Romance languages such as Russian.

This longitudinal study follows beginner learners of Russian (n=11) for one year as they undergo their typical university-level language course. In contrast to the previous year's group of students (controls), this cohort had 30 minutes of in-class singing weekly as part of their academic syllabus. The production data were collected every week for 24 weeks (12 weeks of Semester 1 and 12 weeks of Semester 2) and constituted both reading and spontaneous speech. The control group did not undertake any training and participated in regular language classes following their normal academic routine. The pronunciation was evaluated: 1) holistically with 6 naïve ratters submitting their impressions of the accentedness on a 5-level Likert scale, and 2) acoustically in Praat with vowels, consonants (voice onset time and palatalization), intonation contour/F0, and lexical stress examined for all 24 production samples for each participant (216 speech samples). The holistic measures did not differ significantly between the treatment and control groups but linear hierarchical regression analysis demonstrated a significant difference in vowel production ($p=.007$) and lexical stress ($p>.0001$). As vowels are typically more sustained or extended during singing in comparison to speaking, learners who sing regularly might have more time to process language-specific acoustic cues. Previous research has also shown that music training is positively related to the acquisition of prosody (e.g., Torppa et al., 2020). These findings suggest that signing should be integrated into the L2 curriculum as a relatively simple and enjoyable way to improve some aspects of L2 production.

References

- Christiner, M., Rüdigger, S., & Reiterer, S. M. (2018). Sing Chinese and tap Tagalog? Predicting individual differences in musical and phonetic aptitude using language families differing by sound-typology. *International Journal of Multilingualism*, 15(4), 455-471.
- Nardo, D., & Reiterer, S. (2009). Musicality and phonetic language aptitude. *Language talent and brain activity*, 213-256.
- Reiterer, S. M., Hu, X., Erb, M., Rota, G., Nardo, D., Grodd, W., ... & Ackermann, H. (2011). Individual differences in audio-vocal speech imitation aptitude in late bilinguals: functional neuroimaging and brain morphology. *Frontiers in psychology*, 2, 271.
- Seither-Preisler, A., Parncutt, R., & Schneider, P. (2014). Size and synchronization of auditory cortex promotes musical, literacy, and attentional skills in children. *Journal of Neuroscience*, 34(33), 10937-10949.
- Torppa, R., Faulkner, A., Laasonen, M., Lipsanen, J., & Sammler, D. (2020). Links of prosodic stress perception and musical activities to language skills of children with cochlear implants and normal hearing. *Ear and Hearing*, 41(2), 395-410.

The role of working memory and creativity in written task performance

Judit Kormos (Lancaster University) and Shungo Suzuki (Waseda University).

Although the role of creativity in second language (L2) learning is relevant both theoretically and pedagogically, creativity is an under-researched area in the field of second language acquisition. Writing in another language requires creativity in knowledge creation and transformation as well as the careful orchestration and coordination of cognitive processes. Therefore, creativity can be expected to play a role in content planning, idea generation and lexical selection processes of L2 writing. Recent research in cognitive psychology also suggests that there is a link between divergent thinking (a key facet of creativity) and working memory capacity (Frith et al., 2021). As L2 writing is a cognitively demanding task that draws on existing L2 knowledge and skills, working memory (WM) resources can also exert a substantial impact on the quality of the written product. Nonetheless, research on the role of WM in L2 writing is scarce and has mostly involved either young L2 learners (Michel et al., 2019) or university students (e.g., Mavrou, 2020; Vasylets & Marin, 2021) and has investigated a limited type of writing tasks. Our project is novel in jointly examining the role of creativity and WM in a large sample of Hungarian teenage learners of English using a narrative and argumentative writing task.

In our research, 100 Hungarian secondary school students aged between 16-18 at B1-B2 level of language proficiency wrote a story based on six unrelated pictures and an argumentative essay. Participants completed a forward and a backward digit span task and an emotional stroop task and a recently standardized and validated creativity test for the Hungarian context (Fáy et al., 2021) which allowed us to examine not only the role of creative fluency and elaboration but also facets of originality. The accuracy, lexical sophistication, diversity and syntactic complexity of the students' writing was analysed with automated natural language processing tools: TAALES (Kyle & Crossley, 2015); TAALED (Kyle, Crossley & Jarvis, 2021); TAASSC (Kyle, 2016); GAMET (Crossley et al., 2019). The length of the text in words and sentences was the measure of productivity. A member of the team that developed, validated and standardized the creativity test scored the creativity tests using standardized procedures. The impact of creativity and working memory on the lexical diversity, sophistication, syntactic complexity, accuracy and length of the argumentative and narrative texts was examined using structural equation modelling.

The results reveal a statistically significant relationship between WM and facets of creative originality as well as a significant, albeit weak, role of WM and creativity in the accuracy, syntactic complexity and lexical diversity of learners' performance in both tasks. The findings also suggest that creative originality is associated with a wider range of performance measures in the narrative than in the argumentative task. The findings of our research can inform L2 writing pedagogy and yield insights into what areas of L2 writing students with lower working memory capacity and creativity might need additional support.

Une étude exploratoire sur l'impact de la profondeur orthographique et de la complexité morphologique de la L1 sur le traitement morphographique en français L2

Pierre Largy (Université Toulouse-Jean Jaurès), Vera Serrau (Université Toulouse-Jean Jaurès) and Cecilia Gunnarsson-Largy (Université Toulouse-Jean Jaurès).

Le degré de profondeur orthographique d'une langue (distance entre orthographe et prononciation) semble solliciter des modalités différentes de traitement orthographique (Katz & Frost, 1992). Dans ce sens, les locuteurs d'une L1 à orthographe opaque comme le français ou l'anglais devraient traiter les mots majoritairement dans leur structure visuo-orthographique (accès direct aux mots), alors que les locuteurs d'une langue à orthographe transparente comme l'italien ou l'espagnol privilégieraient le traitement des unités sub-lexicales (assemblage entre graphèmes et phonèmes ou entre morphèmes). Des études récentes portant sur l'impact de l'orthographe sur la performance en L2 montrent que les locuteurs d'une L1 à orthographe transparente s'appuient davantage sur la forme écrite des mots (orthographe) et sur la voie sub-lexicale en L2, ce qui influence positivement leur performance orthographique en L2 (van Daal and Wass, 2017) mais négativement leur performance phonologique en L2 (prononciation superflue des lettres silencieuses dans des mots de la L2 : Bassetti & Atkinson, 2015). Or, d'une part, l'impact de la profondeur orthographique de la L1 sur le traitement orthographique en L2 semble avoir été rarement étudié. D'autre part, la plupart des études sur ce sujet se focalisent sur le traitement de mots simples et ne prennent pas en compte l'interaction entre profondeur orthographique et complexité morphologique. La présente étude vise à observer si des effets translinguistiques de transfert cognitif attribuables à la profondeur orthographique et la complexité morphologique (richesse de la morphologie flexionnelle) de la L1 émergent lors du traitement morphographique (= traitement orthographique des mots fléchis) en français L2. Nous nous focalisons sur la performance morphographique dans des verbes fléchis homophoniques vs. hétérophoniques (arrive/arrivent vs. part/partent) en tâche de rappel écrit de phrases dictées en Français L2, auprès de deux groupes de 48 apprenants experts (B2/C1) ayant une L1 à orthographe transparente et à morphologie complexe (L1-TR+MC), l'espagnol ou l'italien, vs. une L1 à orthographe opaque et à morphologie simple (L1-OP+MS), l'anglais. Les résultats de notre étude renvoient à un transfert cognitif, lors du traitement écrit des mots fléchis en Français L2, des mécanismes cognitifs développés en fonction de la profondeur orthographique et de la complexité morphologique de la L1. En effet, la performance morphographique des apprenants italophones/hispanophones est significativement supérieure à celle des apprenants anglophones. De plus, la réussite des apprenants italophones/hispanophones ne varie pas en fonction de l'audibilité de l'accord verbal, alors que celle des apprenants anglophones est affectée négativement par l'homophonie verbale. Nous interprétons ces tendances par le transfert à la L2 de l'appui sur l'orthographe et du recours majoritaire au traitement morphologique auprès des L1-TR+MC. Ce transfert favoriserait la performance morphographique en L2, indépendamment de la disponibilité d'indices audibles de discrimination morphologique. En revanche, le transfert à la L2 de l'appui sur l'oral et du recours majoritaire au traitement lexical auprès des L1-OP+MS complexifierait la performance morphographique en L2, tout particulièrement en absence d'indices audibles de discrimination morphologique (homophonie).

Effect of first language lexicalisation on second language lexical inferencing and acquisition: A study of French-speaking learners of Chinese as a foreign language

Enhao Léger-Zheng (Laboratoire de NeuroPsychoLinguistique, Université de Toulouse Jean Jaurès) and Olga Théophanous (Laboratoire de NeuroPsychoLinguistique, Université de Toulouse Jean Jaurès).

Lexical inferencing is an important reading and vocabulary learning strategy in second language (L2) acquisition. However, few studies have investigated the effect of lexicalisation, i.e. the existence of an attested lexical equivalent for an L2 word in learners' native language (L1), on these processes. Paribakht (2005) observed that Iranian learners of English succeed more readily in inferring the meaning of English words lexicalised in Persian than non-lexicalised ones. Paribakht and Tréville (2007), comparing Iranian and French learners of English, found that both groups obtained better inferencing scores for lexicalised words. However French-speaking learners had generally better performances than Iranians in inferring English words. The similarity L1-L2 (French-English) had thus a positive effect on lexical inferencing. The aim of the present study is two-fold. First, it focuses on difficulties French-speaking learners may encounter during the inferencing of unknown Chinese words, lexicalised and non-lexicalised in French. Second, it focuses on the subsequent retention of these inferred words. 33 intermediate level learners of Chinese were asked to infer via the Think-aloud protocol the meaning of 20 Chinese words (10 lexicalised and 10 non-lexicalised in French) integrated in two texts written in Chinese. We examined and classified 654 inferences based on the knowledge sources classification by Paribakht (2005), to which we added four specific sources: form, meaning and pronunciation of component character(s) in a Chinese word, as well as the relationship between component characters in a Chinese word. Contrary to the studies on English L2 mentioned above, our results show that French-speaking learners of Chinese had better inferencing score for non-lexicalised words. Known characters comprised in an unknown Chinese word are likely to help learners understand the meaning of this word partially, if not totally. However, learners obtained more totally correct answers for lexicalised words than non-lexicalised ones. As for subsequent retention of inferred words, according to the results of the Vocabulary Knowledge Scale administered three times (once pre-inference and twice post-inference, 2 and 4 weeks later), analysed with the Linear Mixed Model, learners had significantly more difficulties in retaining the meaning of a non-lexicalised word in both short and long terms ($p=0.01$). Besides, our results reveal that learners' cognitive investment in inferred words via dictionary consulting during the post-inference phase may have positive impact on lexical gains for these words, but only in terms of lexical form, not meaning. This result, contrary to that of Laufer and Rozovski-Roitblat's study (2015) on English L2, is probably due to the particularity of Chinese characters and lexis.

LAUFER, B. & ROZOVSKI-ROITBLAT, B. (2015). Retention of new words: Quantity of encounters, quality of task, and degree of knowledge. *Language Teaching Research*, 19, 687-711.

PARIBAKHT, T. (2005). The influence of first language lexicalization on second language lexical inferencing: A study of Farsi-speaking learners of English as a foreign language. *Language Learning*, 55(4), 701-748.

PARIBAKHT, T. & TRÉVILLE, M.-C. (2007). L'inférence lexicale chez des locuteurs de français et des locuteurs de persan lors de la lecture de textes anglais : Effet de la lexicalisation en première langue. *CMLR*, 63(3), 399-428.

A corpus-based approach to 'singular they' in L2 writing

Yi Liu (University of Sheffield) and Thomas Hammond (University of Sheffield).

This study uses a learner corpus to investigate what features of the antecedent determine use of 'singular they' in naturalistic L2 writing. Background: In English, 'singular they' is used with grammatically singular antecedents where gender is indeterminate (Paterson 2011). It can be co-referenced with both (in)definite specific (a) and non-specific (b) antecedents:

(a) I gave a/the student top marks, they were happy ([+/-definite, +specific]) (b) If I gave a/the student top marks, they would be happy ([+/-definite, -specific])

Antecedents of 'singular they' can also vary regarding notional number (e.g., someone [-plural] vs. everyone [+plural]). Previous L1-English studies found 'singular they' occurs mostly with indefinite antecedents that are non-specific and notionally plural (e.g., any student, everybody) (Paterson 2014). Similar to English, Mandarin Chinese does not have nominal inflection for gender but third-person pronouns do carry gender information when written. Ta (他: he/gender neutral; 她: she) is used for singular referents, and the plural marker -men (们) can be attached to indicate plurality with grammatically plural referents: 他们 (ta-men: masculine/gender neutral) and 她们 (ta-men: feminine). Previous L2-English studies have largely focused on elicited use of 'singular they' by learners of highly gendered L1s (e.g., Russian and Italian), showing potential L1 influence (Stormbom 2022). Compared to those languages, Mandarin Chinese is more similar to English in terms of gender marking, thus, research is needed to test whether this similarity effects learners' use of 'singular they'. Research Question: Is Chinese learners' use of singular they conditioned by the specificity, definiteness, and notional number of its antecedent, and how does this compare to previous L1/L2 studies? Methodology: Using the TECCL corpus, we examine the distribution of 'singular they' in Chinese university students' essays (n=417), analysing the form and properties of their antecedents. We found 1000 tokens of 'they' in total and only 7.5% (n=75) were 'singular they'. Results: The antecedents of 'singular they' manifest in four forms: possessive determiners (e.g., our school), definite NPs (e.g., the government), indefinite pronouns (e.g., anyone), and indefinite NPs (e.g., a family). The distributions regarding the features are shown below: [+definite] 46% [+specific] 17% [+plural] 32% [-indefinite] 54% [-specific] 83% [-plural] 68% Conclusions: When using 'singular they', there appears to be no preference for the features [+/-definite] but a preference for non-specific, notionally singular antecedents, which differs from previous L1 and L2 studies. This suggests that Chinese learners do not treat 'singular they' as a direct equivalent of ta-men, but instead conceptualise this as a separate pronominal form from plural they. We discuss this in relation to Homonymy Theory (Whitley 1978; Paterson 2014) and L1 influence in the use of L2 epicene pronouns (singular they and generic he) more generally.

Selected references Paterson, L. (2014) British pronoun use, prescription, and processing: Linguistic and social influences affecting 'they' and 'he'. London: Palgrave Macmillan. Stormbom, C. (2022) Singular they in English as a foreign language. *Applied Linguistics Review*, vol.13, no.5, pp.873-897.

Intervening with debates: EFL students re-engaging in an Exploratory Practice classroom

Elizabeth Machin (None).

The present study offers insights on engagement in an English as a Foreign Language (EFL) classroom through implementation of the Exploratory Practice Principles at a university in Spain. The account sits within the third strand of the L2 Motivational Self System (L2MSS) (Dörnyei, 2009), the recently re-defined 'L2 learning experience' (Dörnyei, 2019), and within the broader domain of the construct of engagement (e.g., Seligman, 2011). The need for understanding evolved out of disaffection: the puzzle of student rejection of classroom communication work in pairs / threes. Student-led debates (a 'Potentially Exploitable Pedagogic Activity' (PEPA)) were explored as a device for restoring engagement. Teacher and students together examined the apparent success of the debates within a classroom discussion and perception questionnaires were individually completed by the students. The debate scripts prepared by groups of students, comments gathered by the students during the classroom discussion and individual freestyle commentary within the questionnaires were analysed using thematic analysis (Braun & Clarke, 2006, 2013). Six themes emerged from the discussion and freestyle commentary: (1) Positive nerves (2) We liked it (3) Respect (4) We learnt new things (5) A challenge and (6) A good form. The present study finds that the students demonstrated motivated L2 behaviour in the group writing of their debate scripts and appeared engaged whilst debating. The students were moved to improvise during the debate, with the data suggesting the form of a debate provided welcome scaffolding.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*.

<https://doi.org/10.1191/1478088706qp063oa> Braun, V., & Clarke, V. (2013). *Successful qualitative research: A practical guide for beginners*. Sage.

Dörnyei, Z. (2009). The L2 Motivational Self System. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, Language Identity and the L2 Self* (pp. 9–42). *Multilingual Matters*.

<https://doi.org/10.21832/9781847691293-003> Dörnyei, Z. (2019). Towards a better understanding of the L2 Learning Experience, the Cinderella of the L2 Motivational Self System. *Studies in Second Language Learning and Teaching*, 9(1), 19–30. <https://doi.org/10.14746/ssl1t.2019.9.1.2>

Seligman, M. E. P. (2011). *Flourish: A new understanding of happiness and well-being – and how to achieve them*. Nicholas Brealey.

Exploring vocabulary attrition through network models: Possibilities, findings and potential

Paul M. Meara (Swansea University) and Imma Miralpeix (Universitat de Barcelona).

The mental lexicon has often been conceptualized as a network where words are interconnected (Aitchison, 1987). However, research so far has only rarely looked at the possible implications of a networked vocabulary, nor has it explored in depth how network structures can give insight into processes such as language attrition (Meara, 2004). Understanding how language is lost is a concern for psychologists, applied linguists and second language acquisition researchers (Ecke, 2004; Schmid & Köpke, 2019).

This presentation makes use of simulations (i.e., computer models of vocabulary-like networks) using Cellular Automata (Kauffman, 1993), which are minimally organised networks where each word is linked to a small number of other words, and where words change their activity status depending on the inputs they receive from other words. These models can provide useful analogues for real vocabularies and exhibit some emergent properties which are relevant to understand the mental lexicon (Meara & Miralpeix, 2022). In these networks, each word is linked to two other words and are activated ('on') or deactivated ('off') in response to the input it receives from its two links. Some words are easily activated, while other words are more difficult to activate.

Different ways of modelling attrition in a vocabulary network are explored, focusing on the results of the simulations and their implications. For example, it is interesting to note that vocabularies do not get smaller if we simply turn some words 'off'. In the case of complete deactivation of a network, it is possible for it to reconstruct itself, either partially or completely, if the right kind of stimulus is provided. The simulations also suggest that we need to distinguish between attrition events and vocabulary loss events. Not all attrition events lead to immediate vocabulary loss, but an accumulation of minimal attrition events typically and quickly leads to a catastrophic collapse of activity in a network. The work reported here suggests that vocabulary attrition is a more complex phenomenon than it is usually assumed to be and that simulations, which have not been very much used in applied linguistics research, might offer a useful way of investigating it further.

References

- Aitchison, J. (1987) *Words in the Mind: An Introduction to the Mental Lexicon* (4th ed.). Blackwell.
- Ecke, P. (2004). Language attrition and theories of forgetting: A cross-disciplinary review. *International Journal of Bilingualism*, 8(3), 321–254. <https://doi.org/10.1177/1367006904008003090>
- Kauffman, S. A. (1993). *The Origins of Order. Self Organisation in Selection and Evolution*. Oxford University Press.
- Meara, P.M. (2004). Modelling vocabulary loss. *Applied Linguistics*, 25(2), 137–155. <https://doi.org/10.1093/applin/25.2.137>
- Meara, P.M. & Miralpeix, I. (2022). Vocabulary networks workshop 1: Introduction. *Vocabulary Learning and Instruction*, 11(1), 17–34. <https://doi.org/10.7820/vli.v11.1.meara1>
- Schmid, M.S. & Köpke, B. (2019). *The Oxford Handbook of Language Attrition*. Oxford University Press.

Early bilingual Papiamentu-Dutch reading development in a post-colonial context

Gil-Marie Mercelina (Radboud University), Eliane Segers (Radboud University), Ronald Severing (University of Curaçao) and Ludo Verhoeven (Radboud University).

Most children in a post-colonial context learn to read in an L2, de facto a foreign language. Recent developments in the Caribbean lead to the introduction of Creoles as language of instruction in the classroom, enabling children to learn to read in the L1 as well. In the Dutch Caribbean Leeward islands, children learn to read in either the L1 Papiamentu Creole or the L2 Dutch. In the present study, we examined the effect of initial language of reading instruction (L1 vs. L2), individual differences (kindergarten speech decoding, vocabulary, rapid naming, short-term memory, phonological awareness, letter knowledge, and oral comprehension), and cross-linguistic transfer on reading development of 289 children. We followed a longitudinal design (kindergarten to second grade) in a natural setting of initial reading instruction in L1 vs L2 using ANOVA and SEM analyses. Results showed kindergartners to perform better on all L1 Papiamentu precursor skills. Decoding performance depended on initial language of instruction. L1-instructed children's Papiamentu decoding was predicted by phonological awareness and letter knowledge. L2-instructed children's Dutch decoding was predicted by phonological awareness, letter knowledge, short-term memory, and speech decoding. Transfer effects were found from the language of decoding instruction to the other language with L1 precursors and decoding skill contributing to the prediction of L2 decoding but not the other way around. As for reading comprehension, instruction language was significant only for Dutch reading comprehension. Kindergarten Papiamentu vocabulary and short-term memory along with word decoding predicted Papiamentu reading comprehension. As for Dutch reading comprehension, along with word decoding, both groups relied on kindergarten vocabulary, oral comprehension, and short-term memory while Dutch-instructed children also relied on kindergarten letter knowledge. However, oral comprehension weighed more heavily for the L1-instructed children while word decoding weighed more heavily for the L2-instructed children. Regarding transfer, the L1-instructed group relied on kindergarten precursors while the Dutch instructed group relied on first-grade precursors to predict the other language. To conclude, both groups learned to decode regardless of instruction language, but results reveal the pivotal role and benefit of including the L1 in early bilingual reading development in a post-colonial context while early fostering of vocabulary and automatization of decoding can support later reading comprehension.

Is L2 grit related to absolute levels of language attainment?

Hitoshi Mikami (Chubu University).

Grit, mental stamina necessary to pursue long-term goals, has recently received the attention of researchers in the field of second language (L2) acquisition (SLA), for the reason that such mental stamina has particular relevance to a long and continuous process like L2 development. Within this trend, Teimouri et al. (2022) introduced the concept of L2-specific grit, and demonstrated the validity of their language-domain-specific measure of grit called the L2 grit scale. The validity of this new scale has also been supported by the results of other studies, and L2 grit is becoming a hot topic in SLA research. This study, however, argues that there is a limitation in the previous validation studies. That is, the relevance of L2 grit to absolute levels of language attainment has been left unclear due to the use of self-assessed proficiency. From an assessment perspective, it is crucial to clarify how learner characteristics are related to objective L2 proficiency as well as to self-perceived proficiency. In this context, the aim of the present study was to investigate how L2 grit is related to absolute language proficiency. This study was conducted in cooperation with the English department of a Japanese university. One hundred and six students took a proficiency test and then responded to a questionnaire. L2 proficiency was measured using a test called TOEIC, which is one of the most popular standardized language tests assessing non-native speakers' English proficiency. TOEIC estimates one's English proficiency based on the scores of two sections (i.e., reading and listening comprehension sections). Consequently, three types of proficiency scores were obtained from the test: the reading and listening scores (hereafter Reading and Listening), and the overall score (hereafter Overall Proficiency). To explore how grit is related to the proficiency measures, grit and L2 grit were measured using the localized versions of their respective scales. Partial correlation analysis was first carried out to confirm how L2 grit is related to the three proficiency measures (Reading, Listening, and Overall Proficiency) when compared to grit. Control variables were gender and year at university. Following this, a series of hierarchical regression analyses were performed to evaluate the incremental prediction of the test scores from grit when L2 grit was used as the first step regression model. The main findings can be summarized as follows. First, the results of the partial correlation analyses showed that L2 grit always has stronger correlations with the test scores than grit. Second, the results of the regression analyses confirmed that L2 grit was a consistent predictor of Reading, Listening, and Overall Proficiency. These results not only provide additional support for the validity of the L2 grit scale, but also demonstrate that gritty L2 learners are indeed likely to achieve a higher linguistic level in their target language. Reference Teimouri, Y., Plonsky, L., & Tabandeh, F. (2022). L2 grit: Passion and perseverance for second-language learning. *Language Teaching Research*, 26(5), 893–918. <https://doi.org/10.1177/1362168820921895>

Comprehension of prefaced disagreements in French by advanced L2 learners

Simone Morehed (Université de Fribourg).

This paper presents a study on the comprehension of disagreements by advanced L2 learners of French. The aim is to study at which moment in the sequential context they understand the upcoming disagreement, and through which (para)verbal resources. Conversation analysis studies describe disagreements as “non-preferred actions”, since they generally go against the social cohesion and thus are prefaced with hesitations, pauses and mitigators (Schegloff 2007). The preface signals the upcoming “problem” and delays and attenuates the disagreement. The preface’s importance is shown in psycholinguistic studies stating that L1 speakers start deciphering the incoming utterance from its start (Barthel et al. 2016). The above studies raise the questions on how L2 learners understand different types of disagreement markers in preface position: which (para)verbal resources do they attend to, and is the preface sufficient to identify the disagreement or is more context needed? What are the similarities and differences between L2 learners and L1 speakers of French? For L2 learners’ interactional and pragmatic competences, comprehension is crucial, without which the learner has difficulties interacting appropriately (Taguchi & Roever 2017). We analysed disagreements from spoken authentic interactions of corpora of L1 French with the method of conversation analysis. The disagreement markers were categorized and include different types, from monosyllabic hesitation markers like “ben”, to constructions like the partial agreement “oui mais” and the epistemic stance “je sais pas”. Based on the analyses, an experiment was designed and conducted with 200 adult advanced L2 learners of French (L1 German or Italian), and 100 L1 speakers. The experiment contained two tasks: an online questionnaire and a stimulated recall. The questionnaire included authentic spoken disagreement sequences in L1 French. The length of the initial disagreement utterance was altered to investigate where in the sequential context the disagreement was understood: from the first part of the preface to the whole utterance. The participants answered written questions reflecting their comprehension of the disagreement. In the stimulated recall, the participants motivated their answers, where we analyse which (para)verbal resources in the disagreement sequences they identified and attended to. This paper will present the results that are currently being analysed. The first results show that the L2 learners generally need more interactional context to understand the disagreement than the L1 speakers, for whom (part of) the preface often is sufficient. In the stimulated recalls, all three participant groups often mention verbal resources, whereas paraverbal resources, e.g. intonation, are mentioned less. The ongoing analyses will investigate in detail how the different markers are interpreted, and further compare the three participant groups. The results can give insights into L2 learners’ comprehension difficulties in authentic oral interactions, and specifically which disagreement markers pose problem or are interpreted differently from L1 speakers.

Barthel, M., Sauppe, S., Levinson, S. C., & Meyer, A. S. (2016). The timing of utterance planning in task-oriented dialogue: Evidence from a novel list-completion paradigm. *Frontiers in Psychology*, 7.

<https://doi.org/10.3389/fpsyg.2016.01858> Schegloff, E. A. (2007). *Sequence Organization in Interaction*. Cambridge: Cambridge University Press.

Taguchi, N., & Roever, C. (2017). *Second language pragmatics*. New York: Oxford University Press.

Do executive function capacities mediate noticing during face-to-face oral interaction?

Jonathan Moxon (Saga University).

Interaction is considered to be facilitative of L2 learning (Long, 1996) and a key role is played by corrective feedback in the processes of noticing (Schmidt, 1990) that mark the beginnings of interaction-driven learning. The importance of attention in the processes of noticing suggest noticing is mediated by individual differences in cognitive capacities, and previous research indicates that global working memory capacities may play a role. This research extends this viewpoint by focusing on the microprocesses of working memory supervision, namely executive function (Miyake et al., 2000). The role of three executive functions, namely shifting, updating, and inhibition during interactionally situated noticing were examined using two tests to measure each of these microprocesses. Three approaches to the measurement of noticing were used, namely stimulated recall, modified output, and a methodologically innovative approach which used a novel measure which combined these two data sets into a single variable where noticing was considered to have occurred when either participants reported noticing or noticing the gap during stimulated recalls, and/or produced modified output in the light of feedback. Modified output data was further informed by response time data (Moxon, 2020), where unusually longer response times were considered indicative of noticing regardless of uptake type and shorter response times were coded no noticing, again regardless of uptake type. The target structure investigated was past counterfactual conditional antecedent and main clause formulation. Simple linear regression analyses indicated that noticing as measured by the response time measure in if clause-focused episodes was predicted by shifting capacities, providing some support for the research hypothesis that greater cognitive flexibility in shifting between meaning and form-focused task sets results in a higher likelihood of noticing corrective feedback. In contrast, updating and inhibition were not found to be significantly associated with any of the noticing measures.

SOURCES Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of Second Language Acquisition* (pp. 413–468). Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49–100. Moxon, J. (2020). The efficacy of uptake response times as a measure of noticing of corrective feedback during oral interaction feedback episodes. *Kumamoto University Studies in Social and Cultural Sciences*, 18, 75–98. Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129–158.

Processing reduced speech in the L1 and L2: A combined eye-tracking and ERP study

Kimberley Mulder (Amsterdam Center for Language and Communication, University of Amsterdam), Sophie Brand (Radboud Teacher Academy, Nijmegen), Lou Boves (Center for Language Studies, Radboud University Nijmegen) and Mirjam Ernestus (Center for Language Studies, Radboud University Nijmegen).

In everyday speech, words are often reduced. Native listeners generally understand reduced forms effortlessly (e.g., Ernestus, Baayen & Schreuder, 2002), but learners can have problems understanding reduced forms (Ernestus, Dikmans & Giezenaar, 2017; Nouveau, 2012; Wong et al., 2015). We investigated the effect of reduction in a passive listening visual world task with French natives and Dutch learners of French. Participants listened to French sentences containing either a reduced or full form of a target noun with a schwa in the first syllable (e.g., /rkɛ̃/ vs. /rəkɛ̃/ for requin 'shark'). The targets were presented in the middle of sentences and were not predictable from the preceding context. The sentences ended with a noun phrase that was semantically related to the target. One second before the start of the sentences four line drawings were shown on the screen, one of which corresponded to the target word. At the end of the sentences a photograph of a scene was shown, and participants had to decide whether the scene depicted the content of the sentence. Eye movements and EEG were recorded simultaneously throughout the experiment. We used eye fixations to define alternative time-lock moments -in addition to the onset of the target words- for the analysis of the EEG signals around the target words. In addition, we developed novel combinations of fixations and saccades as additional predictors in Linear Mixed-Effects Models of the EEG signals. We found a stronger effect of reduction on phonetic processing and semantic integration in learners than in natives, but the effects are different from the N100/N400 and P600 effects found in previous research. Interestingly, we found significant effects of the eye movements and fixations in the one second preview of the four line drawings, as well as in a time window covering the target words. Also, the LME models with time-lock defined by a fixation on the target picture were different from the models with time-lock on target word onset. We model how the use of visual information affects the auditory processing.

Ernestus, M., Baayen, H., & Schreuder, R. (2002). The recognition of reduced word forms. *Brain and language*, 81, 162–173. Ernestus, M., Dikmans, M., & Giezenaar, G. (2017). Advanced second language learners experience difficulties processing reduced word pronunciation variants. *Dutch Journal of Applied Linguistics*, 6(1), 1-20. Nouveau, D. (2012). Limites perceptives de l'e caduc chez des apprenants néerlandophones. *Revue Canadienne de Linguistique Appliquée*, 15, 60-78. Wong, S. W. L., Mok, P. P. K., Chung, K. K.-H., Leung, V. W. H., Bishop, D. V. M., & Chow, B. W.-Y. (2015). Perception of native English reduced forms in Chinese learners: Its role in listening comprehension and its phonological correlates. *TESOL QUARTERLY*, 51, 7-31.

i-lex: an improved method of assessing L2 learner ability to see connections between words?

Ian Munby (Hokkai Gakuen University).

Knowing a word's associations is considered an aspect of word knowledge (Nation, 2001). It follows that L2 learner ability to see connections between words may improve with gains in vocabulary knowledge and proficiency. However, WA research in L2 has not produced a conclusive link between WAT performance and proficiency (Wolter, 2002) and existing WATs suffer from methodological weaknesses. For example, receptive WATs allow testees to guess word associations through elimination processes without knowing why the words are associated. With productive WATs, scoring responses with word association norms is problematic because results depend on the cue words and norms lists used (Schmitt, 1998). The objective of this study is to circumvent these problems by developing a new WAT inspired by Meara (1994). He suggests presenting learners with the three most common associates of a cue word and asking them to supply the missing word. Following this format, a new WAT called *i-lex* (Munby, 2017) was developed using sets of three cue words (CWs) chosen from the five most common associates to 50 target words (TWs) listed in the Edinburgh Associative Thesaurus, or EAT (Kiss et al, 1973).

In the first study (*i-lex v1*), on average, a group of 25 native speakers outperformed an experimental group comprising 98 Japanese learners of English. An unpaired t-test between the mean scores of the two groups yielded a significant difference of $t = 11.153$, $p < 0.0001$. Further, to probe a possible link between *i-lex* performance and vocabulary knowledge, non-native *i-lex* scores were compared with a kanji translation test (Webb, 2008). Pearson correlations among these scores are $r = .729$ (1-sided p value, significant at $p < 0.01$). In a follow-up study (*i-lex v2*), *i-lex* scores of 164 Japanese L2 users of English were compared with the same kanji translation test and an additional measure of vocabulary knowledge: Levels 1-5 of the New VLT (McLean and Kramer, 2015). This yielded significant positive correlations of $r = .729$ and $r = .827$ ($p < 0.01$) respectively, suggesting that the ability of this group to see links between highly frequent English words is related to their vocabulary knowledge. A comparison of *i-lex v2* initial scores and retest scores after 2 weeks yielded $r = .871$ ($p < 0.01$) indicating that the results are reliable.

References Kiss, G.R., Armstrong, C., & Milroy, R. (1973). An Associative Thesaurus of English. EP Microfilms, Wakefield. McLean, S., & Kramer, B. (2015). The creation of a new vocabulary levels test. *Shiken*, 19(2), 1–11. Meara, P.M. (1994) Word associations in Spanish. Available: www.lognostics.co.uk/vlibrary/meara1994.pdf Munby, I. D. (2017). *i-lex v1 and v2: An Improved Method of Assessing L2 Learner Ability to See Connections between Words?* *Vocabulary Learning and Instruction*, 6(1), 75-90. Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press. Schmitt, N. (1998). Quantifying word association responses: What is native-like? *System*, 26, 389-401. Webb, S. (2008). Receptive And Productive Vocabulary Sizes Of L2 Learners. *Studies in Second Language Acquisition*, 30, 01,79-95.

Insights into the Appropriate Use of Machine Translation in Foreign Language Education

Mikie Nishiyama (Postgraduate School, Tokyo Healthcare University) and Noriko Matsuda (Kindai University).

Because of the shift in many university classes to an online format due to the COVID pandemic, computer-mediated online educational environments will continue to be normal for EFL (English as a foreign language) learners. Under such circumstances, it might be better to incorporate the use of machine translation (MT) into foreign language (FL) education at universities. An overview of previous research shows that there are still few studies involving questionnaire surveys of EFL learners regarding the use of MT in FL learning (Briggs, 2018), and that the educational effects of using MT as a supporting tool for learners in FL learning require further investigation (Lee, 2020). Moreover, learners' attitudes to the use of MT are also unknown, and it is necessary to find out what they think about its use in language study (Gally, 2019). In order to propose a method to use MT in FL education, it is necessary to examine how learners use MT, how learners can effectively use MT for FL learning, and how instructors involved in FL education should appropriately respond to learners' use of MT. In light of the above, the present study conducted a questionnaire survey of 245 Japanese EFL learners regarding the use of MT in FL learning, such as translation sites and translation applications, to investigate: (1) how they use MT, (2) their awareness of language forms between the target language and their first language when using MT, and (3) their perceptions toward using MT. The participants' responses to the questionnaire survey were subjected to exploratory factor analysis, and four factors were identified: 1. first language use pre-edit, 2. English-Japanese comparison post-edit, 3. awareness, and 4. trust in MT. Also, based on the responses to the questionnaire items about (1), EFL learners' MT uses were classified into four clusters, which we refer to as the pre-edit only cluster, post-edit only cluster, both pre-edit and post-edit cluster, and neither pre-edit nor post-edit cluster. This study analyzes the relations between the four clusters of learners' use of MT and the questionnaire items of (2) and (3). Based on the results of the analysis, this study suggests the usefulness of using MT as a supporting tool in FL learning, and the perceptions that may be related to learners' use of MT. The appropriate use of MT in FL education is discussed.

References Briggs, N. (2018). Neural machine translation tools in the language learning classroom: Students' use, perceptions, and analyses. *The JALT CALL Journal*, 14(1), 3–24. Gally, T. (2019). The implication of machine translation for English education in Japan. *Language Teacher Education*, 6(2), 1–14. Lee, S.-M. (2020). The impact of using machine translation on EFL students' writing. *Computer Assisted Language Learning*, 33(3), 157–175.

A core metadata schema for L2 data

Magali Paquot (Université catholique de Louvain), Alexander König (CLARIN ERIC), Egon W. Stemle (Eurac Research) and Jennifer-Carmen Frey (Eurac Research).

The main objective of this presentation is to introduce a core metadata schema for L2 production data, more particularly learner corpora, which is the result of extensive collaboration between learner corpus compilers at two European institutions and a research data infrastructure expert and member of CLARIN's metadata taskforce.

The project stems from the recognition that one area that would benefit significantly from standardization is L2 data description, which includes metadata at the level of the dataset as a whole and metadata used to describe the individual learners and task types/registers the corpus is meant to represent. There are a number of reasons why this is important. First, standardized and well-structured metadata increases the findability and usability of existing learner corpora. Second, it should enhance the comparability of datasets and comparability of L2 studies, provided researchers agree on a common set of definitions. Extensive metadata that follow - at best - a standardized vocabulary, and have a strong focus on findability, accessibility, interoperability, and reusability (FAIR) are an essential aspect of FAIR research data (Wilkinson et al. 2016).

In continuation of Granger & Paquot (2017), our proposed metadata schema is divided into a number of different sections for Corpus metadata (itself divided into Administrative metadata (e.g. authors or license) and Corpus design metadata (e.g. date and place of collection or type of task)), Text metadata (fine-grained per-text information), Learner metadata (details about the learners, e.g. age, languages spoken), Annotation metadata (e.g. details about manual or automatic annotation), Annotator metadata (e.g. professional and language background), Transcriber metadata (e.g. native language or language repertoire) and Task metadata (e.g. instructions, time constraints). While basic information about learners (authors) and language samples (texts) are typically found as part of metadata associated with a learner corpus, other aspects such as those related to the annotation or transcription procedure or the specificities of a task are often found elsewhere (e.g. corpus manual) or are just absent from currently available learner corpora. Our proposal is to provide a systematic description of all these aspects as part of core metadata. A beta version of the core metadata schema was tested on a number of learner corpora representing a variety of learners and language samples. After a first presentation to the community at a domain-specific conference we are now in the process of revising our initial proposal based on the comments received and will release a stable version of the core metadata schema in Spring 2023.

References: Granger, S. & Paquot, M. (2017). Towards standardization of metadata for L2 corpora. Invited talk at the CLARIN workshop on Interoperability of Second Language Resources and Tools, 6-8 December 2017, University of Gothenburg, Sweden.

https://sweclarin.se/sites/sweclarin.se/files/event_atachements/Granger_Paquot_Metadata_G%C3%B6teborg_final.pdf Wilkinson, M. D. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, 3(1), 1-9.

Autism and Cognition in Bidialectalism

Kakia Petinou (Cyprus University of Technology) and Kyriakos Antoniou (Cyprus University of Technology).

Speakers of two languages (bilinguals) or dialects (bidialectals) often exhibit smaller vocabularies in each language separately compared to monolinguals. In contrast, some (controversial) evidence suggests that bilinguals enjoy benefits in aspects of non-verbal cognition (e.g., executive control) [1]. Moreover, Autism Spectrum Disorder (ASD) is often linked to language and other intellectual disabilities, with language being more adversely impacted [2]. Finally, work with ASD bilinguals has generally shown that bilingualism does not have any further detrimental effects on ASD individuals' cognition, with few studies even reporting positive bilingual cognitive effects [2]. Here, we focus on bidialectalism—on bidialectal experiences such as second-dialect proficiency, use and degree of dialect switching—and examine how its interaction with autistic traits affects fluid intelligence and second-dialect vocabulary in neurotypical young adults. Sixty-two bidialectals who spoke Cypriot Greek (native dialect) and Standard Greek (SG) as second dialect (mean age=20.5, SD=5.3) took the Autism-Spectrum (AQ) and Systemizing Quotient (SQ) [4] for autistic traits, a Language Background and Dialect Switching Questionnaire, a SG vocabulary test, the WASI [5] and Standard Progressive Matrices (SPM) intelligence tests [6]. We formed composite scores for related variables by averaging relevant z-transformed measures (in parentheses): Autism (AQ, SQ total) and IQ (WASI, SPM). An analysis on SG vocabulary, with Autism, SG use, Switching, the Autism by SG use and Autism by Switching interactions as predictors showed only a negative effect of Autism ($F(1, 42)=10.96$, $p<.01$). A similar analysis on IQ, that additionally included SG vocabulary (proficiency) and the Autism by SG vocabulary interaction as predictors, revealed a significant Autism by Switching interaction ($F(1, 39)=5.73$, $p<.05$). This indicates that (1) high autistic traits have a negative effect on IQ only at low levels of switching (Figure 1) and/or that (2) switching has a positive effect on IQ only at high autistic traits (Figure 2). The latter interpretation may be in line with claims that bilingualism has positive effects only in individuals who are not at the peak of cognition [1]. While this study is ongoing, our results suggest that bidialectalism does not interact with autistic traits in affecting verbal cognition. We discuss whether the interactive effect of switching and autistic traits on fluid intelligence represents a bilingual advantage.

[1] Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. *Psychological bulletin*, 143, 233. [2] Peristeri, E., Silleresi, S., & Tsimpli, I. M. (2022). Bilingualism effects on cognition in autistic children are not all-or-nothing: The role of socioeconomic status in intellectual skills in bilingual autistic children. *Autism*, 26, 2084-2097. [3] Wheelwright, S., Baron-Cohen, S., Goldenfeld, N., Delaney, J., Fine, D., Smith, R., ... & Wakabayashi, A. (2006). Predicting autism spectrum quotient (AQ) from the systemizing quotient-revised (SQ-R) and empathy quotient. *Brain research*, 1079, 47-56. [4] Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: The Psychological Corporation. [5] Raven, J., Raven, J. C., & Court, J. H. (2000). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 3: The Standard Progressive Matrices, Including the Parallel and Plus Versions*. San Antonio, TX: Harcourt Assessment.

Tracking the development in Japanese EFL learners' alignment activity and topic management in study abroad, virtual exchange, and language classroom

Sajjad Pouromid (Kansai University).

Interactional competence (IC), defined as the contingent and context-specific use of interactional practices to achieve joint actions by interlocutors in social interaction (Pekarek Doehler, 2019), has attracted a growing body of research. In the realm of second language acquisition, research has drawn on ethnomethodology and conversation analysis to elucidate how L2 learners develop the competencies to participate in social activities. There are studies with both longitudinal and cross-sectional designs with the aims of tracking developments in IC or documenting its facets. Such studies often focus on specific actions and the methods L2 speakers employ to accomplish those actions. Some aspects that have already been explored include turn-taking practices (Cekaite, 2007), opening tasks (Hellermann, 2008), telling stories and responding to them (Pekarek Doehler & Berger, 2018), as well as repairing troubles (Hellermann, 2011). Learning a foreign language is not a monolithic experience, however, as L2 is learnt in various learning environments. Against this backdrop, this presentation reports on the preliminary results of a longitudinal conversation analytic study comparing the development of IC in three groups of Japanese EFL learners drawing on analogous data collected at regular intervals from synchronous online discussions among members of each group. The first group comprised three participants studying abroad in the US and Canada over a one-year span, and there were three participants in the second group engaged in a language learning virtual exchange (VE) program between their home university and a partner university from the US. The third group also included three participants taking a general English course at their university without a study abroad or a VE component. The analysis focuses on how these novice L2 learners developed their linguistic repertoire and expanded their interactional inventories to perform conversational alignment activity and topic management. The findings suggests that the participants studying abroad in an English-speaking context performed alignment activities earlier than their peers in the other two groups. These participants were able to produce locally contingent turns and demonstrate their active listenership by displaying their understanding of the previous turns and orienting to them. They were also able to manage the emerging topics in their interactions and expand them to sustain the flow of the conversation.

Cekaite, A. (2007). A child's development of IC in a Swedish L2 classroom. *The Modern Language Journal*, 91(1), 45–62.

Hellermann, J. (2008). *Social actions for classroom language learning*. Clevedon: Multilingual Matters Ltd.

Hellermann, J. (2011). Members' methods, members' competencies: looking for evidence of language learning in longitudinal investigations of other-initiated repair. In Joan K. Hall, John Hellermann & Simona Pekarek Doehler (Eds.), *L2 IC and Development* (pp. 147–172). Clevedon: Multilingual Matters.

Pekarek Doehler, S., & Berger, E. (2018). L2 IC as increased ability for context-sensitive conduct: A longitudinal study of story-openings. *Applied Linguistics*, 39(4), 555-578. doi:10.1093/applin/amw021

Pekarek Doehler, S. (2019). On the nature and the development of L2 interactional competence: State of the art and implications for praxis. In R. M. Salaberry & S. Kunitz (Eds.), *Teaching and testing L2 IC* (pp. 25-59). New York: Routledge.

Adult learner motivation to learn Chinese in second and foreign language contexts

Jie Rao (PhD candidate at University of York) and Bimali Indrarathne (Associate Professor at University of York).

Abstract: The last two decades have witnessed an increasing interest among adults in many countries to learn Chinese as a second/foreign language. A few studies have already examined adult learner motivation to learn Chinese in different language contexts, for example, in the USA (e.g., Wen, 2022), the UK (e.g., Mayumi & Zheng, 2021), and in China (e.g., Li & Zhang, 2021). To investigate whether and how L2 motivation of Chinese as a second and foreign language differs in second and foreign language contexts, comparative study needs to be conducted with L2 learners from similar backgrounds, which enables us to better understand the influence of language context on L2 motivation. Therefore, this study aims to investigate the differences in the motivation of adult learners of Chinese in a foreign and second language context by using Dörnyei (2005, 2009)'s Motivational Self System which includes the Ideal L2 self, Ought to L2 self and L2 learning experience as its theoretical framework. Qualitative method was adopted in this study with semi-structured interviews. Ten participants were recruited from each context by convenient sampling (20 in total). They belonged to the UK, the USA, Mexico, Colombia, Zambia, Nigeria, Zimbabwe, and Pakistan. The interviews were approximately 50 minutes long. The questions were designed based on the three facets of Dörnyei's L2MSS (ideal L2 self, Ought-to L2 self and L2 learning experience), on which the coding themes were formed as well. Nvivo (software) was used in coding and analyzing the interview. The results revealed that the learning experience is the most powerful factor to boost learners' motivation in both contexts. For example, attitude towards the target language learning experience, teacher influence on motivation, and interaction with target communities. Both groups showed similar levels of Ideal L2 self, that is, interest in the language and culture, etc. A disparity between the two groups is that participants from the Chinese context posed a much stronger Ought-to L2 Self than those from the UK context. This can be attributed to their higher level of anxiety in speaking to native speakers of Chinese and stronger instrumental motivation such as seeking for better job opportunities in China.

References Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates. Dörnyei, Z. (2009). The L2 motivational self system. In Z. Dörnyei, & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 9–42). Bristol: Multilingual Matters. Li, M., & Zhang, L. (2021). Tibetan CSL learners' L2 motivational self system and L2 achievement. *System*, 97, 102436. Mayumi, K. & Zheng, Y. (2021). Becoming a speaker of multiple languages: an investigation into UK university students' motivation for learning Chinese. *The Language Learning Journal*. 1-15. Wen, X. (2022). Chinese language learning motivation: a study of individual-contextual interactions. *Journal of Multilingual and Multicultural Development*, 1-17.

Cross-linguistic pervasiveness of Agent-first in passive contexts and effects in L2 grammars

Isabel Repiso (Salzburg University) and Cyrille Granget (Université Toulouse - Jean Jaurès).

One of the principles governing utterance structure in initial stages, i.e. the so-called Basic Variety, is that the NP-referent with highest control comes first (Klein & Perdue 1997: 315). This semantic principle has not been tested in the grammar of more advanced L2 learners that have benefited from guided instruction. In the aim of addressing this gap, our study analyses utterance structure in productions where agent-first and patient-first are two options. Crucially, French and Spanish share a common repertoire of agent-first constructions, such as SVO with causative verbs (1a) and causative verbal periphrasis (1b), as well as a common repertoire of patient-first constructions, such as passives with 'être'/'ser' (2a), and verbal periphrasis 'se faire'/'hacerse' (2b). The former constructions mark the instantiation of the patient as object, whereas the latter mark it as subject.

1. a. Le bully l'a agressé / El bully lo agredió [The bully attacked him] b. Le garçon l'a fait trébucher / El chico lo hizo tropezar [The guy made him tripped]
2. a. Il a été agressé / Él fue agredido [He was assaulted] b. Il s'est fait gifler / Él se hizo abofetear [He got slapped]

We designed a visual stimulus in which 50% of the vignettes represented a human figure as agent, and 50% as patient with another animated entity; the latter were our target items. Participants were asked to retell what happened to the human figure. First, we carried out an offline-retelling task targeting the elicitation of intra-typological variation data across L1 French vs L1 Spanish (84 responses obtained from 21 speakers per group). Then, we replicated the task in L2 French with a group of 20 Spanish-speaking learners. Our L1 results show significant differences in the syntactic instantiation of patients: they are saliently marked as syntactic subjects in French and as direct objects in Spanish (e.g., *Il s'est fait agresser* vs. *El bully lo golpeó*). This result suggests that constructions in Spanish are stronger governed by the semantic principle Agent-first, even in contexts where a form of passive construction could be expected. Unlikely L1 grammars, the degree to which one argument "controls" the situation was not so clear-cut expressed in L2 French grammars. Learner Varieties exhibit greater frequencies of stative verbs that give to read the syntactic subject as a sort of middle voice's experiencer (e.g., *Il a peur du chien*). Beyond this developmental feature, L2 grammars prove to be Agent-first rooted. In all, our results suggest that L2 learners do not efficiently overcome the Agent-first pervasiveness of their L1 in passive contexts, suggesting that the Agent-first principle is not only linked to the learner's environment and competence but also to the typological differences between the native and the target language.

References Klein, W., Perdue, C. (1997). The Basic Variety. Or: Couldn't natural language be much simpler? *Second Language Research* 13, 301-347.

Learning syntactic variation in L2 sentence processing: The role of prediction error

Duygu Şafak (TU Braunschweig) and Holger Hopp (TU Braunschweig).

Learning and using syntactic variation gives rise to several challenges, as illustrated in the English dative alternation, since learners need to acquire the syntactic options available in an L2 and the different constraints on optionality. In English, most ditransitive verbs can take both double-object dative (DO) and prepositional-object dative (PO). However, this optionality is constrained by gradient selectional restrictions on the type of their arguments: DO-bias verbs, e.g., pay, tend to occur more frequently with DO, while PO-bias verbs, e.g., send, show a probabilistic tendency to prefer PO. Research on L2 predictive processing indicates that adult L2 learners are sensitive to these gradient verb constraints during real-time comprehension [1-2], which suggests that they can extract distributional probabilities of complementation from the L2 input.

In this study, we explore how L2 learners learn syntactic variation by testing whether and how L2 learners adapt their syntactic predictions when encountering unexpected input that signals shifts in syntactic variation, e.g., when DO-bias verbs occur with the PO structure rather than the DO structure. To this end, we tested adult L1 German intermediate to advanced L2 learners of English (n=48) in a priming experiment using visual world eye-tracking. Participants first read aloud written prime sentences crossing the factors verb bias (DO-bias vs. PO-bias) and structure type (DO vs. PO); see Figure 1. Subsequently, they listened to spoken target sentences while viewing visual displays with an agent referent (the tailor), a recipient referent (the model), and a theme referent (the dress) (Figure 1). The results of cluster-based permutation analyses showed PO-priming effects: L2 learners generated more expectations for the PO structure when they heard the target verb after PO (vs. DO) primes, as evidenced by a significantly lower proportion of looks to the recipient than to the theme in Figure 2a. The results also revealed surprisal effects of verb bias on PO-priming effects, such that priming effects were larger when the prime structure did not match the bias of the prime verb, i.e., after PO primes with DO-bias verbs than with PO-bias verbs (Figure 2b). However, there were no strong priming effects for the DO structure (Figure 3).

The overall pattern of results demonstrates that, like for L1 learners and users [3], prediction error drives implicit learning in an L2. We discuss the findings in the context of error-based implicit learning models [4]. [1] Şafak, & Hopp. (in press). Cross-linguistic differences in predicting L2 sentence structure: The use of categorical and gradient verb constraints. *SSLA*. [2] Wolk, Wolfer, Baumann, Hemforth, & Konieczny. (2011). Acquiring English dative verbs: proficiency effects in German L2 learners. In Carlson, Hölscher, & Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*. [3] Chen, Wang, & Hartsuiker. (2022). Error-based structure prediction in language comprehension: Evidence from verb bias effects in a visual-world structural priming paradigm for Mandarin Chinese. *J Exp Psychol Learn Mem Cogn*, 48, 60-71. [4] Chang, Dell, & Bock. (2006). Becoming syntactic. *Psychological Review*, 113, 234-272.

Does L1 orthographic depth influence L2 orthographic processing of inflected words?

Vera Serrau (Université Toulouse Jean Jaurès), Cecilia Gunnarsson-Largy (Université Toulouse-Jean Jaurès) and Pierre Largy (Université Toulouse-Jean Jaurès).

Orthographic depth, i.e. the extent to which an orthography deviates from the “one-grapheme-to-one-phoneme” principle, influences orthographic processing in expert readers/writers (Frost, 1994). In deep orthographies like English or French, L1 writers (Deep_L1) tend to access previously-acquired words as whole visual forms. Conversely, in shallow orthographies like Italian or Spanish, L1 writers (Shallow_L1) tend to sequentially parse sub-lexical units like phonemes and graphemes. Recent research shows that Shallow_L1 “upgrade” the sublexical route to morphological parsing when dealing with complex words (Görge et al., 2021). L1 orthographic depth seems likewise to influence L2 orthographic/phonological processing. Compared to Deep_L1, Shallow_L1 show a higher L2 spelling performance (van Daal & Wass, 2017) but more frequent L2 orthography-based phonological errors (e.g. silent letters erroneously-pronounced in L2 words: Bassetti & Atkinson, 2015). This suggests that Shallow_L1 transfer their reliance on orthography to the L2. Our previous study (Author, 2023) on orthographic processing of homophonic vs. heterophonic inflected verbs in L2 French (mange/mangent vs. boit/boivent) revealed that Shallow_L1 having a richly-inflected L1 (Italian/Spanish) rely more on L2 orthography-oriented morphological processing (= same performance for homophonic and heterophonic inflected words), whilst Deep_L1 having a poorly-inflected L1 (English) rely more on L2 phonology-oriented lexical processing (= lower performance while dealing with homophonic inflected words). This study investigates whether, during L2 morphographic processing (i.e. processing of written inflected word), L1-to-L2 transfer effects attributable to L1 orthographic depth emerge also when comparing L2 writers whose L1s have a similar degree of inflectional richness. We observe two groups of 60 expert L2 English writers (C1) whose L1 is Italian (IT-L1: Shallow_L1) vs. French (FR-L1: Deep_L1), two richly-inflected languages, dealing with L2 English morphology in a task of written-recall of dictated sentences. We focus on 36 homophonic (passed/past; won/one) vs. heterophonic (based/best; bought/boat) English word-pairs, containing regular (passed; based) vs. irregular (won; bought) inflected verbs. Controlling word-pairs audibility enables us to observe to what extent the compared groups rely on phonological cues during L2 morphographic processing, whereas controlling verb regularity enables us to observe how the compared groups deal with morphologically-decomposable words (effective sub-lexical processing) vs. morphologically-indecomposable words (ineffective sub-lexical processing). We generally expect L1 orthographic depth to influence L2 morphographic performance. More precisely, IT-L1 should be less affected by homophony than FR-L1, since Shallow_L1 tend to rely less on phonological cues. IT-L1 should likewise outperform FR-L1 when dealing with L2 word-pairs including regular verbs. Indeed, processing regularly-inflected words sublexically should help Shallow_L1 distinguish the verb (pass-ed) from its uninflected homophonic/heterophonic counterpart (past/pest). Conversely, FR-L1 should outperform IT-L1 when dealing with L2 heterophonic word-pairs including irregular verbs, since the distinction between the verb (bought) and its heterophonic counterpart (boat) should benefit from the reliance on phonological cues (dominant in Deep_L1), and not from the reliance on sub-lexical processing (dominant in Shallow_L1). The data collection was conducted via an ad hoc-built privacy-preserving keystroke-logging website (OnLog), and statistical analyses are currently being performed (results in progress).

A longitudinal study of preschool learners' L1-L2 vowel production

Šárka Šimáčková (Palacký University Olomouc), Václav Jonáš Podlipský (Palacký University Olomouc) and Monika Kučerová (Palacký University Olomouc).

Bilinguals, including L2 learners, represent speech sounds in one shared phonological system [1] which changes dynamically with experience [2]. Interactions of L1~L2 sound categories are bidirectional [3]. L1 sounds can drift toward L2 sounds because of immersion [2] or even few-week-long classroom learning [4]. They can also move away from similar L2 categories [5]. Most cross-language effects are reported for late(r) learners; young simultaneous bilinguals in bilingual communities keep equivalent sound categories separate [6]. Little is known about preschool classroom learners; longitudinal studies are also scarce. This longitudinal study focused on L1&L2 vowel production by 7 Moravian-Czech preschoolers (aged 3;9-5;8) attending weekly EFL classes (45-minute exposure to SSBE) for at least 10 months. We asked: (1) Are the learners' L1&L2 vowels separated acoustically? (2) Did production of L1 and/or L2 vowels change over time? During 3 months, the children attended 8 recording sessions, 2 in Czech (9 weeks apart) and 6 in English. They produced mono-/di-syllabic words in a picture-naming task. The target words (38 English, 24 Czech), selected according to familiarity and imageability, included the RP vowels /i,ɪ,ɛ,æ,ʌ/ and Moravian-Czech short /i,ɛ,a/; the remaining English and Czech monophthongs occurred in fillers.

Normalised vowel height, F1-F0, and retraction, F2-F0, (all in ERB) were modelled by two linear mixed-effects models with Vowel and Time and their interaction as the fixed effects (/æ/ in Time 1 as intercept), and Speaker (varying intercepts and slopes for Time) and Word (varying intercepts) as the random effects. In Time 1, English-/æ/ (typically merged with Czech-/ɛ/ in late(r) learners [7]) differed reliably from the nearest categories: both in height and retraction, it was in-between English/Czech-/ɛ/'s on the one hand, and Czech-/a/ and English-/ʌ/ on the other. Reliable shifts between Times were found in height for Czech-/ɪ/, which raised (unlike English-/ɪ/), for Czech-/ɛ/, which lowered (and started overlapping with English-/æ/), and for English-/ʌ/, which raised (and became more centralised, unlike Czech-/a/).

In summary, we found that a difficult L2 category, /æ/, was surprisingly well differentiated from the neighbouring vowels. The longitudinal changes suggest that for preschoolers even limited input can lead to measurable phonetic shifts to accommodate L1&L2 categories in the shared vowel space.

References [1] Flege & Bohn (2021). The revised speech learning model. In R. Wayland (Ed.), *Second language speech learning: Theoretical and empirical progress*, 3-83. CUP. [2] Sancier & Fowler (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *J. Phonetics*, 25(4), 421-436. [3] Kartushina et al. (2016). How and when does the second language influence the production of native speech sounds: A literature review. *Language Learning*, 66(S2), 155-186. [4] Chang (2012). Rapid and multifaceted effects of second-language learning on first-language speech production. *J. Phonetics*, 40(2), 249-268. [5] Flege (2007). Language contact in bilingualism: Phonetic system interactions. *Laboratory Phonology*, 9(353-381). [6] Guion (2003). The vowel systems of Quichua-Spanish bilinguals. *Phonetica*, 60(2), 98-128. [7] Šimáčková & Podlipský (2018). Production accuracy of L2 vowels: Phonological parsimony and phonetic flexibility. *Research in Language*, 16(2), 169-191.

The Instructed Learning of Form–Function Mappings of L2 English Generic NPs

Neal Snape (Gunma Prefectural Women's University), Helen Zhao (The University of Melbourne) and Menghan Wang (The University of Melbourne).

An interest in providing instruction to L2 learners about article semantics has been approached recently by adopting a generative approach to SLA (GenSLA). Our study, like Umeda et al. (2019), provides instruction on English genericity to upper-intermediate and advanced L1-Japanese L2-English learners. However, instruction was designed and conducted as an experimentalized online tutor (Zhao & MacWhinney, 2018) that implements random assignment, practice tasks, and immediate feedback. Participants were randomly divided into a treatment group (n = 20) and a control group (n = 17). Participants' TOEIC scores were used as the measure of L2 proficiency. A pre-test, training sessions and two post-tests were administered that measured knowledge of English generics based on Krifka et al.'s (1995) classification of kind NP-level and sentence-level generic sentences. Only the definite singular and bare plural can be used for NP-level generics such as a natural kind with a kind predicate like the dodo bird/dodo birds is/are extinct. The indefinite singular and bare plural are acceptable for sentence-level generics, e.g., an orange/oranges has/have lots of vitamin C. The treatment group were trained for generic usage through an online acceptability judgement task with lots of practice trials and metalinguistic feedback (in Japanese). The control group received a comparable amount of time of online English preposition training. The study was programmed in PsyToolKit by a research programmer. Linear mixed effects models were built to analyse group differences in treatment effects across time and moderated by L2 proficiency. The results from both groups are provided in figures 1-8 below. For the NP-level generics, the treatment group showed a significant growth in the use of the definite singular and bare plurals from pre-test to immediate post-test. Similarly, they significantly improved on the use of the indefinite singular and bare plurals for the sentence-level generics. All improvements were retained in the delayed post-test (two weeks apart). The control group's performance did not differ over time. L2 proficiency was found to influence treatment effects on the more difficult generic forms (i.e., definite singular for NP-level generics and indefinite singular for sentence-level generics). In GenSLA, some properties of language are deducible through exposure to primary linguistic data (PLD), as discussed in Rothman (2008). We believe that due to a lack of generic NPs in PLD, ongoing exposure to form-function mappings can help learners to restructure their interlanguage grammars with explicit knowledge becoming implicit knowledge over time.

References

Krifka, M. Pelletier et al. (1995). Genericity: an introduction. In G. Carlson & F. Pelletier (Eds.), *The generic book* (pp.1-125). Chicago: University of Chicago Press. Rothman, J. (2008). Aspect selection in adult L2 Spanish and the Competing Systems Hypothesis: When pedagogical and linguistic rules conflict. *Languages in Contrast*, 8, 74-106. Umeda, M. et al. (2019). The long-term effect of explicit instruction on learners' knowledge on English articles. *Language Teaching Research, Special Issue: Grammatical Meaning and the Second Language Classroom*, 23, 179-199. Zhao, H., & MacWhinney, B. (2018). The instructed learning of form–function mappings in the English article system. *The Modern Language Journal*, 102, 99-119.

Exploring receptivity to adjectival scales in L2 implicature derivation

Glenn Starr (University of Wisconsin-Milwaukee) and Emilie Destruel (University of Iowa).

Recent forays in experimental scalar implicature research show that native speaking participants derive implicatures far more often in scenarios where the stronger term on an adjectival scale represents a bounded endpoint (Van Tiel et al. 2016). With the statement “This math problem is difficult”, for example, a scale containing the words “difficult” and “impossible” is generated. The term “impossible” denotes a maximal value of the property in question, a feature embedded into the lexico-semantics of the word itself (Kennedy 2007). This promotes a fixed interpretation of the term that can be arrived at independent of context, thus making it clearly distinguishable from its weaker counterpart on the scale. Therefore, when a participant is asked, “Would you conclude that the problem is difficult but not impossible?” the answer will likely be “yes”. However, with the statement “This student is intelligent”, a hearer is much less likely to conclude that the speaker intended to mean “the student is intelligent but not brilliant”. This is because the scale < intelligent, brilliant > contains relative adjectives that better describe ranges rather than endpoints. Delineating a boundary between the two terms necessitates reference to a contextual domain in relation to the object(s) being described. Researchers argue that more effort is required to determine whether implicature calculation is necessary under these conditions because hearers have difficulty (1) identifying a borderline and (2) establishing clear standards which might warrant the derivation (Frazier et al. 2008).

Some research suggests that language learners do not process certain constructions as efficiently as native speakers due to increased processing costs in the L2. However, researchers have discovered that L2 learners can compensate for this with increased sensitivity to surface-level lexico-semantic cues during sentence processing (Clahsen & Felser 2018). Since the stronger terms on bounded or partially bounded scales contain built-in endpoints which facilitate the demarcation of boundaries between scalemates, this study investigates L2 receptivity to these properties. Native speakers and L2 participants will complete an inference task and decide whether the speaker meant to negate the stronger term. Target bounded/unbounded scales contain pairs of gradable adjectives only, an understudied part of speech in the L2 scalar implicature literature. We also consider whether this phenomenon applies cross-linguistically by testing both L1-French L2-English as well as L1-English L2-French learners. Results will further our understanding of scalar diversity in L2 scalar implicature processing. The experiments are currently underway, and we are eager to present our findings in August.

Clahsen, H., & Felser, C. (2018). Some notes on the shallow structure hypothesis. *Studies in Second Language Acquisition*, 40(3), 693-706.

Frazier, L., Clifton C., and Stolterfoht, B. (2008). Scale structure: processing minimum standard and maximum standard scalar adjectives. *Cognition* 106, 299–324. doi: 10.1016/j.cognition.2007.02.004

Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy*, 30(1), 1–45. <https://doi.org/10.1007/s10988-006-9008-0>

Van Tiel, B., Van Miltenburg, E., Zevakhina, N., & Geurts, B. (2016). Scalar Diversity. *Journal of Semantics*, 33(1), 137–175. <https://doi.org/10.1093/jos/ffu017>

The German Summary corpus (GerSumCo): A new resource for contrastive research into L2 German of advanced writers

Carola Strobl (University of Antwerp), Helena Wedig (University of Antwerp), Jim J.J. Ureel (University of Antwerp) and Tanja Mortelmans (University of Antwerp).

The German Summary corpus (GerSumCo) is a new corpus for contrastive research into German as a second (L2) vs. first language (L1). GerSumCo was created to investigate cohesion in academic L2 German writing produced by advanced learners. There are several corpora for the contrastive investigation of German learner language available, targeting diverse acquisition levels, text types and L1 backgrounds (e.g., KOLAS: Knorr & Andersen, 2017; Falko: Lüdeling et al., 2008). However, whereas summary writing is an interesting genre for the analysis of cohesion (as seen in Walter, 2007), the only existing corpus of summaries to date is the Falko summary subcorpus (Lüdeling et al., 2008). Preliminary analyses of the Falko summary L2 subcorpus revealed a high degree of patchwriting, i.e., students copy-pasting larger chunks of text from the original text. Since this creates a bias in the data, we decided to compile a new summary corpus. The specificity of our corpus is twofold: First, students created summaries from two different source texts, i.e., they needed to create their own coherent flow, which diminishes the problem of patchwriting. Second, all summaries were written based on the same source texts and under comparable conditions: All students had to write a summary of two popular scientific texts about a topic related to language variation in contemporary German (e.g., Kiezdeutsch, Mundartdebatte in der Schweiz). To date, GerSumCo consists of 89 summaries which were written by 42 L2 German students with diverse L1s and 47 L1 German students, with the corpus still growing. The texts were collected at several German Universities during the academic year of 2022-23. For a research project aimed at investigating cohesive strategies deployed by L1 and L2 German writers, the corpus was pre-processed and general linguistic information was added automatically (e.g. part-of-speech). The first analysis of the corpus focuses on connectives as a well-researched cohesive device in learner language. Manual corrections of an automatic pre-annotation via DimLex (Scheffler & Stede 2016; Stede 2002) were conducted by three trained annotators using guidelines based on the PDTB-3 scheme (Webber et al. 2019). The poster will introduce the corpus to the research community and first results of the contrastive analysis of connectives.

References

- Knorr, D., & Andresen, M. (2017). Commented Learner Corpus Academic Writing (KoLaS). Hamburger Zentrum für Sprachkorpora. <http://hdl.handle.net/11022/0000-0001-B732-8>.
- Lüdeling, A., Doolittle, S., Hirschmann, H., Schmidt, K. & Walter, M. (2008). Das Lernerkorpus Falko. *Deutsch als Fremdsprache*, 2, 67–73. <https://doi.org/10.37307/j.2198-2430.2008.02.02>
- Scheffler, T., & Stede, M. (2016). Adding semantic relations to a large-coverage connective lexicon of German. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 16)* (pp. 1008–1013). ELRA.
- Stede, M. (2002). DiMLex: A lexical approach to discourse markers. In A. Lenci & V. Di Tomaso (Eds.), *Exploring the lexicon: Theory and computation* (pp. 1–15). Edizioni dell'Orso.
- Walter, M. (2006). Hier wird die Wahl schwer, aber entscheidend. *Konnektorenkontraste im Deutschen. Österreichisches Jahrbuch Deutsch als Fremdsprache 2006*.
- Webber, B., Prasad, R., Lee, A., & Joshi, A. (2019). *The Penn Discourse Treebank 3.0 Annotation Manual*. University of Pennsylvania.

Multilingual experience results in early noticing and resolution of translation ambiguity in vocabulary learning

Lari-Valtteri Suhonen (University of Borås).

One difficulty that arises when learning new words in a new language is whether the meaning of new lexical items aligns with pre-existing conceptual categorizations/representations. Previous research (e.g., Eddington & Tokowicz, 2013) refers to the lack of such alignment across languages as translation ambiguity. Jiang (2002) has proposed that learners are largely dependent on metalinguistic awareness in translation ambiguity resolution. If pre-existing metalinguistic awareness is a factor in learning and using translation-ambiguous items, it should be the case that experienced language learners (i.e., multilinguals) will outperform language learners with less experience. This would be expected to be the case since they – as a part of their previous successful acquisition of additional languages – have experience with the phenomenon. While experienced language learners have been found to be overall better at vocabulary learning than inexperienced language learners (Kausihanskaya & Marian, 2009; Papagno & Vallar, 1995; Van Hell & Mahn, 1997), the present study aims to investigate whether there is an additional effect for specifically translation-ambiguous items. This study compared language learners with little experience in foreign language learning (L1 English, N = 30) and experienced language learners (L1 Swedish and L2 English, N = 30). The participants learned both translation ambiguous and non-translation ambiguous items using virtual flashcards with a picture and the target item. During learning, time spent on each virtual flashcard during learning was recorded as were response time and accuracy during assessment. The experienced language learners were indeed overall faster during learning (est. -617ms, $p < 0.01$), corroborating previous research. The difference between the groups decreased over the course of learning (est. 55ms, $p < 0.01$). For translation ambiguous items, no significant differences between the groups were found in terms of accuracy. Furthermore, both groups were affected by translation ambiguity during assessment in that they used significantly more time on translation ambiguous items (est. 834ms, $p < 0.01$). However, the experienced learners were comparatively less affected by translation ambiguity (est. -513ms, $p < 0.01$). It can be hypothesized that this is due to their pre-existing awareness of the phenomena as well as how to deal with it.

References Eddington, C., & Tokowicz, N. (2013). Examining English-German translation ambiguity using primed translation recognition. *Bilingualism: Language and Cognition*, 16, 442–457.

Jiang, N. (2002). Form-meaning mapping in vocabulary acquisition in a second language. *Studies in Second Language Acquisition*, 24(4), 617–637.

Kaushanskaya, M., & Marian, V. (2009). The bilingual advantage in novel word learning. *Psychonomic Bulletin & Review*, 16, 705–710.

Papagno, C., & Vallar, G. (1995). Verbal short-term memory and vocabulary learning in polyglots. *Quarterly Journal of Experimental Psychology*, 48A, 98–107.

van Hell, J. G., & Mahn, A. C. (1997). Keyword mnemonics versus rote rehearsal: Learning concrete and foreign words by experienced and inexperienced learners. *Language Learning*, 47, 507–546.

Acquisition of the “that”-trace effect by Japanese learners of English: Examination of the adverb effect and its implications for the theory of the anti-locality

Kasumi Takahashi (Graduate Student, University of Tsukuba) and Yuichi Ono (Faculty of Humanities and Social Sciences).

This paper discusses the L2 learner’s acquisition of the “that”-trace effect, in which the ungrammaticality originates from the sequence of “that+trace” like “*Whoi do you think [that ti met John]?” The well-known syntactic constraint responsible for this effect is “the anti-locality constraint (ALC)” on wh-movement, which prohibits movement that is too short (e.g., Brillman & Hirsch, 2016; Erlewine, 2020). This constraint further explains the so-called “adverb effect,” which is important evidence. The unacceptability is ameliorated by the intervention of adverbials between “that” and trace, as in “Whoi do you think [_ that fortunately [ti met John]]?” (Schippers, 2020).

ALC is assumed to be a universal constraint. Thus, examining L2 acquisition of “that”-trace effect might investigate the general issue of accessibility to universal knowledge. However, few studies in SLA have examined sensitivity to the adverb effect on the “that”-trace construction. A recent study by Kim and Goodall (2022) examined Korean and Spanish learners of English and concluded that ALC does not work in L2. However, neither the adverb effect nor presence/absence of “that” was not examined in their work, which is an academic gap in our research. Second, previous studies have not considered individual differences in acceptability (Coward & McDaniel, 2021). Accordingly, our research questions are: (1) Are Japanese intermediate learners of English (JLEs) sensitive to the adverb effect and the presence/absence of “that”? (2) Are there individual differences in sensitivity to ALC observed by JLEs?

Nineteen Japanese university students with intermediate proficiency levels (CEFR: B2-C1) participated in our study. The factors and conditions were set as follows: [\pm that (presence/absence)], [\pm adverbial (absence/presence(word/phrase/clause))], and [+adv_function (sentential/temporal)]. Four tokens were created per condition (n=112 items in total), along with filler items (n=96).

Main findings are as follows: (1) Main effect of [\pm that] condition was significant ($F(1,18)=8.650$, $p=.009^{**}$), and [–that] was rated higher for both [+adverbial] and [–adverbial], implying the sensitivity to [\pm that]. (2) Regarding the [+that] condition, there is no significant difference in the acceptability of [\pm adverb] ($F(1,18)=0.0451$, $p=.834$), indicating that no adverb effect was observed in this experiment. After cluster analysis for individual differences into two groups, we found a tendency that the upper-rating group (n=11) prefers [+that] systematically, contrary to the prediction of ALC, while the lower-rating groups (n=8) prefers [–that] systematically, in accordance with the prediction of ALC. Thus, we conclude that two separate groups gave a systematic complementary judgement in our intermediate level Japanese learners of English.

References Brillman, R., & Hirsch, A. (2016). An Anti-Locality Account of English Subject/Non-Subject Asymmetries, CLS 50. Coward, W., & McDaniel, D. (2021). The that-trace effect. Goodall G (ed.) The Cambridge handbook of experimental syntax. Cambridge: Cambridge University Press, pp. 258–77. Erlewine, M. Y. (2020). Anti-locality and subject extraction. *Glossa: A Journal of General Linguistics* 5: 84. Kim, B., & Goodall, G. (2022). The source of the that-trace effect: New evidence from L2 English. *Second Language Research*. 1-24. Schippers, A. (2020). COMP-trace revisited: an indirect dependency analysis. *ResearchGate*. 1-24.

Indirect effect of orthographic form on phonetic realisation in L2 German: A Corpus study of inflectional endings in spontaneous speech

Megumi Terada (MA Student).

Recent studies have demonstrated the effect of orthography on L2 speech production (Hayes-Harb & Barrios, 2021). However, most studies have been conducted in experimental settings and little is known about spontaneous L2 speech production, with the exception of Young-Scholten and Langer (2015). To address this issue, a corpus study was conducted on the German schwa in inflectional endings; despite its orthographic representation in the canonical form, it is known that schwa is mainly not realised by German L1 speakers, especially in the final position of the verb (Kohler & Rodgers, 2001). This study tested if learners of German realise more schwa in inflectional endings than native speakers, as the schwa is canonically presented in the written form. The analysis was based on free dialogues in the CoNNAR corpus which is fully transcribed and aligned with the audio (including 40 dialogues between L1-L1 speakers and L1-L2 speakers of German). The whole dataset consists of 26597 tokens whereas L2 German speakers with L1 English (n = 4) were compared to L1 German speakers (n = 4). The analysis included two types of inflectional endings with schwa: <e> as a marker for the first person singular present indicative and <en> for the first, second, and third person plural present indicative. In the annotation, the realisation of inflectional endings was analysed acoustically and divided into two categories: whether the schwa was realised in inflectional endings or not. In order to avoid arbitrariness in the annotation process, the following criteria were established: 1) if periodic movements were observed in the waveform and 2) if there were visible formant movements and/or glottalisation in the sonagram when the schwa was preceded or followed by a vowel. Where this was the case, the schwa was annotated as 'present' and in other cases as 'absent'. A total of 1252 cases were analysed using generalised linear mixed effects models. The results suggest that L2 speakers produce significantly more schwa in inflectional endings than L1 speakers, with a stronger main effect of the inflectional ending <en>. The results of this study suggest that orthography may also indirectly influence spontaneous L2 speech production.

Hayes-Harb, R., & Barrios, S. (2021). The influence of orthography in second language phonological acquisition. *Language Teaching*, 54(3), 297–326. <https://doi.org/10.1017/S0261444820000658> Kohler, K. J., & Rodgers, J. (2001). Schwa deletion in German read and spontaneous speech. *Spontaneous German Speech: Symbolic Structures and Gestural Dynamics*, 97–123. https://www.ipds.uni-kiel.de/kjk/pub_exx/aipuk35/kkjr.pdf (retrieved: 27.01.2023) Young-Scholten, M., & Langer, M. (2015). The role of orthographic input in second language German: Evidence from naturalistic adult learners' production. *Applied Psycholinguistics*, 36(1), 93–114. <https://doi.org/10.1017/S0142716414000447>

Linguistic and interactional development of interrogatives in French L2: proficiency or exposure?

Anita Thomas (University of Fribourg) and France Rousset (Institute of Multilingualism, Fribourg).

Research on the development of interrogatives in L2 spoken French shows that classroom learners at beginner levels tend to overuse formal variants (i.e. subject-verb inversion) as well as the question marker “est-ce que” ‘is it that’ (Dewaele, 1999; Myles et al., 1999). These structures are present but not frequent in native spoken French (Gadet, 1997). A recent study by Donaldson (2016) showed that the production of interrogatives by advanced learners of French was very similar to what is found in native speakers, suggesting that the level of proficiency has an effect on the type of interrogatives produced by L2 learners. The aim of the present study is to examine the production of interrogatives from a linguistic and interactional point of view in learners at (low) intermediate level of French who are mainly exposed to French spoken language. It is based on the longitudinal data from seven adult migrants (L1 Tigrinya) who are training in manual professions. They were audio-recorded for ten minutes during free peer-to-peer interactions, four times over one year. The results are based on the manual coding of 320 questions. They show no use of subject-verb inversion and only eight tokens of “est-ce que”. The learners all produced different variants of interrogatives, verbless questions (30%), subject-verb questions (30%) and WH-questions (30%). Moreover, from an interactional point of view, the question-answer sequences included, from the first recording, elements considered central to active listenership (Salaberry & Kunitz, 2019). The results suggest that the linguistic variants and interactional aspects of interrogatives in L2 French is less a factor of language proficiency than of the type of language input. The results are discussed with respect to the role of written and oral language norms in L2 learning.

Dewaele, J.-M. (1999). Word order variation in French interrogative structures. *ITL - International Journal of Applied Linguistics*, 125(1), 161–180. <https://doi.org/10.1075/itl.125-126.01dew>

Donaldson, B. (2016). Aspects of interrogative use in near-native French: Form, function, and register. *Linguistic Approaches to Bilingualism*, 6(4), 467–503. <https://doi.org/10.1075/lab.14024.don>

Gadet, F. (1997). *Le français ordinaire*. (2nd ed.). Armand Colin.

Myles, F., Mitchell, R., & Hooper, J. (1999). Interrogative Chunks in French L2: A Basis for Creative Construction? *Studies in Second Language Acquisition*, 21(1), 49–80. <https://doi.org/10.1017/S0272263199001023>

Salaberry, M. R., & Kunitz, S. (Eds.). (2019). *Teaching and Testing L2 Interactional Competence: Bridging Theory and Practice*. Routledge.

Explaining academic achievement among international students in HE in the UK: the role of creative coping strategies.

Jeanine Treffers-Daller (Emeritus Professor) and Anne Vicary (Honorary Fellow).

Many international students in Higher Education in the UK struggle to understand the compulsory texts for their course and obtain lower scores for their modules than their monolingual peers. A key reason for this achievement gap is the fact that international students have lower language and literacy skills (Hu and Trenkic, 2021; Trenkic and Warmington, 2018). Here we explore the relationship between vocabulary knowledge, reading behaviour and academic achievement in terms of 30 international students on a postgraduate Law module. They took the 20k version of the Vocabulary Size test (Coxhead, Nation and Slim 2015) and provided summative assignment grades for their Law module and their IELTS scores (overall scores and separate scores for reading). A subgroup of 14 students also took part in qualitative interviews to explore their reading behaviour. Students' mean IELTS reading sub-scores were relatively low (6.63), and there were 15 students with reading scores of 6.0 or below. The VST scores appeared to indicate that students had a vocabulary of 10,000 words on average. However, students' actual vocabulary sizes were probably lower than those indicated by the VST; the poorer sampling of items for the 20k version by comparison with the 14k version may have led to inflated vocabulary size values (Gyllstad, Vilkaite, and Schmitt, 2015). The fact that our students' mean vocabulary scores were higher than those in Trenkic and Warmington (2018), who used the 14k version, while our students' mean reading scores were lower also suggest that the VST scores in the current sample were inflated. Nevertheless, we found mid strength correlations between the IELTS reading sub-score and the VST ($r_s = 0.594$, $p = .002$). Contrary to expectations, the module score did not correlate with either the VST or the IELTS reading subscore. Put differently, neither students' vocabulary sizes, nor their reading levels could explain variance in academic achievement. In-depth interviews revealed that the absence of correlations was likely due to students with IELTS scores of 6.0 or lower using 'creative' reading strategies to pass assignments, which included the use of Google translate or similar online tools for reading compulsory texts and writing assignments. We conclude by formulating implications for admissions tutors and teachers in HE.

Coxhead, A., P. Nation, and D. Sim (2015). "Creating and Trialling Six Versions of the Vocabulary Size Test." *The Tesolanz Journal* 22: 13-27. Gyllstad, H., L. Vilkaite, and N. Schmitt. 2015. "Assessing Vocabulary Size through Multiple-Choice Formats: Issues with Guessing and Sampling Rates." *ITL-International Journal of Applied Linguistics*, 166 (2): 278-306. Hu, R., and D. Trenkic, D. 2021. "The Effects of Coaching and Repeated Test-Taking on Chinese Candidates' IELTS Scores, their English Proficiency, and Subsequent Academic Achievement." *International Journal of Bilingual Education and Bilingualism* 24 (10): 1486-1501. Trenkic, D., and M. Warmington. 2018. "Language and Literacy Skills of Home and International University Students: How Different are they, and Does it Matter?" *Bilingualism* 22 (2): 349-365.

Assessing syntactic complexity in L2 academic writing: teacher judgments, student ratings and complexity indices

Ineke Vedder (University of Amsterdam).

The contribution focuses on the assessment of syntactic complexity as a predictor of L2-development in academic writing, by holistic ratings of language teachers and L2 learners, compared with the outcomes on overall complexity measures. For assessing syntactic complexity various indices have been proposed (Bulté & Housen, 2012; Kyle, Crossley & Verspoor, 2021). Overall syntactic complexity has typically been assessed with measures focusing on the length of syntactic units (mean length of T-unit, sentence, clause; number of Words/T-unit) and clausal subordination measures (Clauses/T-unit, Dependent Clauses/Clause). To a lesser extent, indices for coordination and phrasal complexity have been employed, such as Bardovi Harlig's Coordination Index (1991) or number of complex noun phrases per T-unit (Biber et al., 2011). The question arises how far the results obtained with these complexity indices may also be relevant for language pedagogy, and to what extent they correlate with perspectives on syntactic complexity of stakeholders in the L2 classroom. Few studies, so far, have taken a bottom-up approach by analysing judgments of teachers (Kuiken & Vedder, 2019) and L2 learners in the process of acquiring the target language, or by comparing holistic human ratings with the results yielded by automated complexity tools (Granfeldt & Ågren, 2014). Another under-studied area of research is the potential influence of raters' native or non-native speaker status, which might lead to L1-based preferences for particular syntactic L2 features and a major leniency/severity when evaluating syntactic complexity in L2 (Bogorevich, 2018, 2019; Duym et al., 2018). The proposed study aims to fill this gap. The following research questions have been formulated: 1) How do teacher judgments of syntactic complexity converge with ratings by L2 learners? 2) How far do judgments of syntactic complexity by native and non-native raters correlate? 3) To what extent are teacher and student ratings of syntactic complexity correlated with scores obtained by overall indices of syntactic complexity? A group of 16 expert raters, language teachers with Italian as their L1, were asked to rate on a six-point Likert scale the syntactic complexity of six argumentative texts written by intermediate Dutch L2 learners of Italian (level A2-B1). Teacher judgments, score justifications and suggestions for feedback were confronted with the ratings and feedback of 28 non-expert L2 raters (university students of Italian, level B2-C1), in two different linguistic contexts (Finland, Hungary), and with the judgments of 20 native Italian university students. All holistic human ratings were then compared with the outcomes on three indices of overall complexity (Clauses/T-unit, Dependent Clauses/Clause, Words/T-unit). The results showed that scores of teachers and students highly agreed, although teacher ratings appeared to be motivated by accuracy rather than by reasons concerning syntactic complexity. Students' feedback turned out to be more encouraging compared to feedback provided by teachers. Teacher and student ratings corresponded most with Words/T-unit. No impact of (non) nativeness and linguistic background was found. In the paper the background, methodology and the implications of the study for complexity research are discussed, together with the implications for classroom practice and teacher training.

Relationships between bilingual exposure at daycare and vocabulary growth in a linguistically diverse group of two- to four-year-olds

Josje Verhagen (University of Amsterdam), Jan Boom (Utrecht University), Suzanne Aalberse (University of Amsterdam), Darlene Keydeniers (Royal Auris Group), Folkert Kuiken (University of Amsterdam), Anne-Mieke Thieme (University of Amsterdam) and Sible Andringa (University of Amsterdam).

Earlier work has shown that the amount of exposure bilingual children receive in a language is associated with their development in this language. However, previous research has typically focused on the home context, looking at exposure by caregivers, and to a lesser extent, siblings (Unsworth, 2013). In many countries, children attend daycare from a young age onward, and an increasing number of children attend bilingual programs where an additional language is spoken next to the majority language (e.g., English in Spain). Few studies to date have examined the effects of bilingual exposure at daycare on children's language development, but the available studies indicate that bilingual daycare has positive effects on children's development of the minority language and does not negatively impact their development of the majority language (Thieme et al., 2021). However, it is currently unknown if these effects remain if home language exposure is taken into analysis: earlier studies collapsed children irrespective of home language background or looked at children from majority language only families. The aim of this study was to investigate how bilingual exposure at daycare relates to children's dual language development when differences in home language exposure are taken into account.

Participants were 584 two- to four-year-old children who attended Dutch-only or Dutch-English daycare in the Netherlands. In the Dutch-English daycare centers, English as a foreign language was used between 11% and 50% of the time. Children formed a diverse group, as they were exposed to Dutch, English, other languages, or combinations thereof at home. Dutch and English versions of the Peabody Picture Vocabulary Test were administered in four waves that were approximately nine months apart. Estimates of language exposure to Dutch and English at home and at daycare were derived from a parental questionnaire, and weighted for the actual time children spent at home and daycare in a typical week. Latent Growth Modeling analyses showed that the amount of English exposure at daycare was significantly and positively related to growth of English receptive and expressive vocabulary. Amount of Dutch exposure at daycare did not show significant relationships with growth of Dutch receptive and expressive vocabulary. The strengths of the relationships between Dutch and English exposure at daycare and vocabulary growth did not differ between children depending on whether they were exposed to Dutch or English at home.

These results corroborate the effects of bilingual daycare found in earlier research and indicate that these effects remain if home language exposure is taken into account. Thus, our findings indicate that, at least with the current sample and in the current context, bilingual daycare supports dual language development from a young age onward.

Thieme, A. M. M., Hanekamp, K., Andringa, S., Verhagen, J., & Kuiken, F. (2022). The effects of foreign language programmes in early childhood education and care: a systematic review. *Language, Culture and Curriculum*, 35, 334-351.

Unsworth, S. (2013). Current issues in multilingual first language acquisition. *Annual Review of Applied Linguistics*, 33, 21-50.

Lexical overlap in foreign language speech segmentation in primary-level students

Katie Von Holzen (English Linguistics, TU Braunschweig), Marie Schnieders (English Linguistics, TU Braunschweig), Sophia Wulfert (English Linguistics, TU Braunschweig) and Holger Hopp (English Linguistics, TU Braunschweig).

Policy makers base decisions about the ideal starting date for foreign language (FL) teaching on FL learning outcome studies at secondary level which often support a later start (Jaekel et al., 2017). However, their results are influenced by the quality of instruction rather than the learning abilities by the students (Baumert et al., 2020).

To study students' initial learning abilities, we focus on a key skill: speech segmentation. Adults initially struggle to extract or segment word forms from continuous speech in an FL but succeed when target words partially overlap in form and meaning with their L1 equivalents (i.e. cognate: English: /kraʊn/; German: /kro:nə/; noncognate: English: /skɪn/; German: /haʊt/; Shoemaker & Rast, 2013) or when target words are adjacent to previously familiarized words (Cunillera et al., 2010). To identify the optimal age for starting FL learning at primary-school level, we study the role of lexical overlap with German for the segmentation of English speech among 1st graders early in the school year and among 2nd graders towards the end of the school year before they have received instruction in English.

In a word recognition study, English-German cognate and noncognate word pairs ($n = 160$) were embedded into an otherwise identical English sentence frame (She reduced her crown/skin mursk to poverty). Within each sentence frame cognate and noncognate words were followed by a noncognate pseudoword (i.e. mursk). Students listened to the sentence followed by an isolated probe-word and indicated via button press whether they heard the probe word in the sentence.

In the talk, we will report results from 48 1st graders and 48 2nd graders to be collected by July. So far, we have analyzed results from 27 first-graders. A general-linear mixed model revealed no difference in target accuracy for cognate and non-cognate words ($p = 0.83$), nor was there a difference in d' prime between them ($p = .97$; Figure 1). There was also no difference in target accuracy for cognate and non-cognate following words (i.e. mursk, $p = 0.53$) or in d' prime between them, ($p = .97$; Figure 2). These preliminary results provide suggestive evidence that 1st grade students do not yet transfer knowledge from their L1 to segment words in an unknown FL. In our final, fully powered sample, we will also examine individual differences in L1 knowledge on the segmentation of FL speech. Results will show whether age (grade level) or L1 knowledge predict initial FL learning skills.

References

- Baumert, J., et. al. (2020). The Long-Term Proficiency of Early, Middle, and Late Starters Learning English as a Foreign Language at School: A Narrative Review and Empirical Study. *LangLearn*, 70(4), 1091–1135. <https://doi.org/10.1111/lang.12414>
- Cunillera, T., et. al. (2010). Words as anchors: Known words facilitate statistical learning. *ExpPsych*, 57(2), 134–141. <https://doi.org/10.1027/1618-3169/a000017>
- Jaekel, N., et. al. (2017). From Early Starters to Late Finishers? A Longitudinal Study of Early Foreign Language Learning in School. *LangLearn*, 67(3), 631–664. <https://doi.org/10.1111/lang.12242>
- Shoemaker, E., & Rast, R. (2013). Extracting words from the speech stream at first exposure. *SecLangRes*, 29(2), 165–183. <https://doi.org/10.1177/0267658313479360>

French Immersion vs. Core French L2 Accentedness: Proficiency Scores and Native Speaker Ratings

Hilary Walton (University of Toronto).

This study evaluates whether there are differences in the second language (L2) speech of students enrolled in French immersion and core French programs by comparing learners' oral proficiency as evaluated by native speaker accentedness ratings and their overall French language ability as evaluated by a validated French proficiency test. French immersion and core French programs are the primary French-as-a-second language (FSL) programs in Canada. While students in both programs acquire French in formal classroom contexts in English-speaking communities, the amount of French instruction and the subjects that are taught in French differ between programs (Canadian Parents for French, 2017). Notably, French immersion speakers have been shown to have a particular non-native French accent, distinct from that of core French speakers (Poljak, 2015); however, the structures that distinguish the L2 speech of these groups remain virtually unexplored. Thus, we seek to distinguish whether there are, indeed, differences in the overall accentedness of these learner groups and if so, to determine whether such differences can be attributed to L2 proficiency. This study has two primary research questions (1) Are there differences in the accentedness of the L2 French speech of French immersion and core French speakers? and (2) If so, are such differences due to program- or proficiency-related factors? To assess these questions, this study examined the French oral production of 58 Grade 12 FSL students enrolled in either a French immersion (n = 28) or a core French program (n = 30). Participants first read the French passage "Mes parents m'énervent" aloud. Next, participants completed the University of Toronto Test of French as an overall measure of language ability. Lastly, recordings of the participants' passage reading underwent an accentedness rating task which was evaluated by six native speaker judges along a 5-point Likert scale from 1 (Definitely non-native; very strong foreign accent) to 5 (Definitely native; no foreign accent). Preliminary results indicate that native speakers do, indeed, score the accentedness of French immersion and core French speakers differently. Overall, French immersion speakers were given a lower accentedness rating, indicating that native speakers perceived this group as sounding more native-like. This group also had higher levels of overall French proficiency than the core French group. However, when evaluating the recordings of proficiency-matched participants, native speakers seem to associate core French learners with a lower accentedness rating, indicating that they have a more native-like accent than their French immersion counterparts. Further analyses will confirm whether between-group accentedness differences between the French immersion and core French program-level and proficiency-matched speaker groups exist. This study evaluates whether differences in the L2 speech of French classroom learners are program- or proficiency-based. The significance of the results will contribute to the field of L2 speech learning and could highlight the influence of FSL programs on the L2 acquisition of French speech.

Canadian Parents for French. (2017). Category placement/hours of instruction/French language experience. <https://on.cpf.ca/wp-content/blogs.dir/1/files/Category-Placement-Hours-of-Instruction-French-Language-Experience-3.pdf> Poljak, L. (2015). A search for "Immersionese": Identifying French immersion accents in BC. Masters Thesis. Simon Fraser University, British Columbia.

Multilingualism negatively predicts prosocial behaviour in young adults, mediated by empathic concern

Yifan Wang (University of Birmingham) and Andrea Krott (University of Birmingham).

Multilingual speakers demonstrate increased empathy [1] compared to monolinguals (but see [2]). Also, empathy is closely related to prosocial behaviour [3]. The present study tested whether multilingualism, especially frequency of language use and language proficiency, promotes prosocial behaviour and, if so, whether this is mediated by empathy. 126 monolinguals, 126 bilinguals, and 50 participants knowing three to five languages filled in the Language and Social Background Questionnaire [4], the Prosocialness Scale [5] and the Interpersonal Reactivity Index [3] (measuring empathy), plus the Culture Orientation Scale [6] and a Socioeconomic Status (SES) questionnaire to control for confounds. We measured multilingualism as self-reported language usage frequency and proficiency across languages. We first established whether multilingualism promoted empathy in our sample. We conducted mixed-effect model analyses for empathy and its four subscales (Perspective taking, Fantasy Score, Empathic Concern, Personal Distress), with the two multilingualism measures as predictors in separate models; and age, SES and cultural orientation as covariates. Both multilingualism measures were significant negative predictors for Empathic Concern (next to collectivism and individualism) and Personal Distress (next to SES). We then tested whether multilingualism promoted prosocial behaviour and, if so, whether it is mediated by Empathic Concern or Personal Distress. Both multilingualism measures were negatively related to prosocial behaviour, and Empathic Concern (but not Personal Distress) was positively related to prosocial behaviour. The mediation analysis showed only one and only an indirect relationship of multilingualism with prosocial behaviour: Language usage frequency predicted prosocial behaviour indirectly via Empathic Concern. These results suggest that being multilingual, or more precisely, high language usage frequency, does not lead to higher but lower prosocial behaviour than being monolingual, and that this is due to lower empathic concern. The negative relationship with empathy stands in contrast to previous findings [1] and might be due to multilinguals being more emotional stable than monolinguals and being less impacted by emotional distraction [7]. Furthermore, multilinguals scoring lower in personal distress confirmed that bilinguals are less anxious than monolinguals [8].

References: [1] Dewaele, J.-M., & Wei, L. (2012). Multilingualism, Empathy and Multicompetence. *International Journal of Multilingualism*, 9(4), 352-366. [2] Ghoreyshi Rad, F., Ahmadi, E., & Gorbani, F. (2020). Comparison of Social Cognition and Executive Functions of Motivation, Inhibitory Control, and Empathy in Bilingual and Monolingual Individuals. *International Journal of Psychology*, 14(1), 59-82. [3] Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, 44(1), 113-126. [4] Anderson, J. A. E., Mak, L., Keyvani Chahi, A., & Bialystok, E. (2018). The language and social background questionnaire: Assessing degree of bilingualism in a diverse population. *Behavior Research Methods*, 50(1), 250-263. [5] Caprara, G. V., Steca, P., Zelli, A., & Capanna, C. (2005). A New Scale for Measuring Adults' Prosocialness. *European Journal of Psychological Assessment*, 21(2), 77-89. [6] Triandis, H. C., & Gelfand, M. J. (1998). Converging measurement of horizontal and vertical individualism and collectivism. *Journal of Personality and Social Psychology*, 74, 118-128. [7] Barker, R. M., & Bialystok, E. (2019). Processing differences between monolingual and bilingual young adults on an emotion n-back task. *Brain and cognition*, 134, 29-43. [8] Dewaele, J.-M., & Wei, L. (2013). Is multilingualism linked to a higher tolerance of ambiguity?. *Bilingualism: Language and Cognition*, 16(1), 231-240.

Processing cognates in idiomatic expressions: a real 'nut to crack'

Kristina Weissbecker (University of Kassel).

Cognates make up a significant amount of a bilingual's vocabulary, and many psycholinguistic studies have attested, they provide processing advantages as opposed to non-cognates (e.g., Costa et al. 2000, Dijkstra, 2005, Sánchez-Casas and García-Albea, 2005). Cognates are also part of numerous idiomatic expressions. Idioms, in turn, present an intricate case for study concerning their (non-)compositionality and the question if they require their own entry in the mental lexicon (see, e.g., Gibbs 1993, Jackendoff 2007). Also, it is yet an open question whether the literal meanings of the corresponding words in an idiom are accessed during processing.

This study deals with the processing of cognates within idiomatic expressions. The main research question is: Can bilingual speakers access a cognate's literal meaning and translation equivalent when it is inside an idiom? In other words, can we also find cognate facilitation in the processing of idioms? To investigate this issue, we are conducting a self-paced reading experiment (SPR) paired with a subsequent lexical decision task (LD). The experimental items consist of English sentences including idiomatic VPs with direct objects as target words. The dependent variable is the cognate status of the target word (cognate vs. non-cognate). The stimuli in the LD are German translations of the targets from the SPR. Fillers contain non-idiomatic phrases in the SPR and pseudo-words in the LD. The measured variables are reaction times (RTs) and accuracy. If speakers can access the literal meaning of the words in an idiom, then faster RTs and a higher accuracy should occur in the LD for experimental items than for fillers. If speakers can access a cognate's literal meaning in an idiom, a processing advantage should occur in the LD as opposed to items with non-cognates.

We expect a processing effect for cognate items which would indicate strong connections of cognate equivalents in the lexicon, possibly even a joint storage within one entry. Alternatively, absence of cognate facilitation would imply a separation of cognate translations (similar to non-cognates). Absence of significant differences between experimental items and fillers would mean that accessing the literal meaning of the words within an idiom is not likely. If this is the case, then idiomatic expressions presumably require separate storage in the lexicon.

References:

Costa, A., Caramazza, A., and Sebastian-Galles, N. (2000). The Cognate Facilitation Effect: Implications for Models of Lexical Access. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26.5 (pp. 1283-96). Gibbs, R.W. (1993). Why idioms are not dead metaphors. In C. Cacciari & P. Tabossi (Eds.), *Idioms. Processing, structure and interpretation* (pp. 57-78). Psychology Press. Dijkstra, T. (2005). Bilingual visual word recognition and lexical access. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism. Psycholinguistic approaches* (pp. 179-201). Oxford UP. Jackendoff, R. (2007). *Language, Consciousness, Culture*. MIT Press. Sánchez-Casas, R. & García-Albea, J. E. (2005). The representation of cognate and noncognate words in bilingual memory. In: J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of Bilingualism* (pp. 226-250). Oxford UP.

Local and long-distance classifier-noun agreement in L2 Chinese sentence processing

Fei Yuan (University of Cambridge) and Boping Yuan (Shanghai Jiao Tong University/University of Cambridge).

Background Second Language (L2) studies of agreement have yielded rich insights into gender and number agreement in L2 grammars. However, relatively few L2 studies have examined classifier-noun (dis)agreement. In the current study, we delve into the processing of classifier-noun (dis)agreement in L2 Chinese among learners from different first language (L1) backgrounds (English: a non-classifier language; Korean: a classifier language). Specifically, we compare the participants' processing of local and long-distance agreement to explore the distance effect. We also examine the cue-based retrieval model in explaining our findings (Lewis & Vasishth, 2013; Cunnings, 2017). Our goals are to examine L1/L2 differences in processing classifier-noun (dis)agreement and to assess the impact of L1 transfer on processing.

Methodology A self-paced reading (SPR) task and an acceptability judgment task (AJT) were administered in this study. Chinese native speakers (NSs) ($n=28$) were recruited as a control group, with L1-English learners of L2 Chinese (EN group) ($n=54$) and proficiency-matched L1-Korean learners of L2 Chinese (KR group) ($n=51$) serving as two experimental groups. The same items were used in both tasks (see example stimuli on the next page). Each type was represented by six tokens, with 77% of the test items being filler items. The data were analysed using R software, and linear mixed-effects models were constructed to compare the types.

Results In the AJT, all three groups rated Types 1 and 2 as acceptable (z -scores >0), but judged Type 3, the disagreement condition, as unacceptable (z -scores <0). This suggests that both native speakers and L2 learners could detect classifier-noun disagreement in an off-line setting. In the SPR (see Figures 1-3), the NS group demonstrated sensitivity to the local disagreement, reflected in significantly and marginally shorter RTs of Types 1 and 3 in Region 6 (Types 1 vs 2: $p=0.05$; Types 2 vs 3: $p=0.089$). The RT of Region 9 in Type 3 was significantly longer than those in Types 1 and 2 (Types 1 vs 3: $p=0.005$; Types 2 vs 3: $p=0.002$). However, the EN group did not exhibit any effect in their SPR results. In the KR group, the RT of Region 6 in Type 2 was marginally shorter than those in Types 1 and 3 (Types 1 vs 2: $p=0.087$; Types 2 vs 3: $p=0.091$), similar to the sensitivity observed in the NS group to local disagreement. However, no effect was observed for the remainder of the regions in the KR group, indicating that the KR group was not sensitive to long-distance disagreement during real-time processing.

Discussion The L1/L2 differences were found in our study, indexed by the NS group's sensitivity and the L2 groups' insensitivity to the long-distance agreement. This phenomenon is likely a result of L2 learners' inability to retrieve cues from working memory in non-local disagreement. L1 transfer was also observed, as L1-Korean learners, but not L1-English learners, behaved native-like and were sensitive to local disagreement. This is attributed to a facilitative influence of their L1 Korean, which is a classifier language as well.

Relativized Minimality in L2 revisited: (non-)effects of L1 and tense on processing of object relative clauses

Vera Yunxiao Xia (University of Alberta), Natália Brambatti Guzzo (St. Mary's University) and Lydia White (McGill University).

Object relative clauses (ORCs) are known to cause greater problems than subject relatives (SRCs) in acquisition and processing (Adani et al. 2010). Friedmann et al. (2009) advance a Relativized Minimality (RM) account: the subject of the ORC intervenes between the head noun and its source position, potentially making interpretation hard. The intervention effect is reduced if the relative head and the intervenor mismatch in features such as number. See (1).

(1) I know the musicians (P) who the waiter (S) likes _.

For L2 acquisition, Authors (2022) report that, contrary to RM predictions, L2ers were faster in the case of matched items (SS vs. SP) in English. Authors speculated that this might be due to the L1 (Mandarin) not having number marking. A second issue is that, in the previous literature, there are always two cues to number: plural inflection on the noun and verbal agreement. In Authors, all verbs were past tense, so the only mismatch involved nominal inflection; compare (1) and (2).

(2) I know the musicians who the waiter liked _.

In this paper, we explore effects of L1 and tense, comparing speakers of Mandarin with Spanish, a language with number marking, and including stimuli with present and past tense verbs. Research questions are:

i. Do Spanish speakers (unlike Mandarin) respond faster to mismatched items than matched? ii. Do L2ers respond faster to present tense items than past?

An experiment was conducted with intermediate level learners of English (L1s Mandarin, n=18, Spanish, n=22), as well as native speaker controls (n=20), using a self-paced reading task. Sentences were presented in segments, followed by a comprehension question. The task included 32 sentences manipulating RC type (subject/object), number on relative head and intervenor (\pm singular), and tense (\pm past), plus distractors. Results show high accuracy overall for the comprehension questions: Spanish speakers were more accurate on SRCs than ORCs ($p = 0.01$) and the Mandarin speakers trended in the same direction, consistent with RM. Looking at overall RTs, none of the groups shows a difference between matched and mismatched items, or an effect for tense. The same is true for RTs at the critical region where match or mismatch becomes evident. However, number did make a difference: Mandarin speakers performed significantly faster on PP versus SS clauses ($p = 0.04$), suggesting that they are in fact sensitive to number. These results are consistent with Contemori and Marinis (2014), who also found no mismatch effect for child and adult English-speakers in a self-paced listening task, and an advantage for plural over singular. Contemori and Marinis propose that mismatch effects are late effects which show up only in offline tasks. Our results are consistent with this proposal.

References Adani, F. et al. 2010. Grammatical feature dissimilarities make relative clauses easier. *Lingua* 120: 2148-2166. Contemori, C., & Marinis, T. 2014. The impact of number mismatch and passives on real-time processing of relative clauses. *JCL* 41: 658-689. Friedmann, M. et al. 2009. Relativized relatives: types of intervention in the acquisition of A-bar dependencies. *Lingua* 119: 67-88.

What looks native-like may not necessarily be native-like: Evidence from L2 Chinese covert objects

Lilong Xu (Shanghai Jiao Tong University) and Boping Yuan (University of Cambridge; Shanghai Jiao Tong University).

Chinese is well-known for allowing ‘null’ objects, but this is subject to constraints. For example, an object in a following-up sentence in a Chinese parallel sentence can be missing only if the verb is identical to that in the preceding sentence (see the minimal pair in (1)). This is called the verbal identity condition. Liu (2014) argues that the missing object (‘covert object’ from here on) should be viewed not as a ‘null object’ but as a result of verb raising and VP ellipsis, as in (2). In English, objects are obligatorily overt, as in (3). This contrast between Chinese and English provides an opportunity to investigate whether covert objects are allowed and constrained by the verbal identity condition in English speakers’ L2 Chinese grammars.

173 English-speaking learners and 30 native Chinese speakers completed a cross-modal picture-description task (CPT) and an acceptability judgement task (AJT). In the CPT, participants first listened to a Chinese parallel sentence like (1a) where the two verbs were identical and the second object was covert. This served as an audio prime. Then participants were asked to describe a pair of pictures with key verbs written above each picture. These two verbs were either identical or non-identical. In the AJT, participants were asked to rate the naturalness of (c)overt objects in Chinese parallel sentences where the two verbs were (non-)identical.

The results of the AJT show that all L2 groups accepted missing objects in parallel sentences with identical verbs but failed to make a distinction between *the missing and overt objects in Chinese parallel sentences with non-identical verbs. In the CPT, beginners only produced complete sentences despite the fact that they were primed with objectless parallel sentences. Although intermediate and advanced L2ers could be primed to produce more objectless sentences, their production of complete and objectless sentences was not affected by whether the verbs were identical or not.

The finding that beginners tend to produce complete sentences supports the findings of Yuan and Zhang (2020), who argue that elliptical structures, which involve additional computational complexity (i.e. movement and ellipsis), are difficult for beginners to produce. Nevertheless, beginners can accept objectless sentences in the AJT, where beginners can probably rely on explicit knowledge. With improved L2 Chinese proficiency, L2ers can accept and produce objectless sentences. However, L1-L2 non-convergence is still observed in the CPT and AJT regarding the verbal identity condition, suggesting that L2ers’ apparently native-like performance is actually driven by mechanisms different from that for L1ers’ performance (i.e. the overwhelming evidence of null objects in the Chinese input may mislead L2ers into overgeneralizing null objects to missing objects in Chinese parallel sentences). The findings advance our understanding of L1 vs. L2 different mechanisms for phonetically unrealised objects in Chinese, suggesting that what looks native-like may not necessarily be native-like.

(1)a. John kan-le baozhi, Ben ye kan-le Φ . John read-ASP newspaper Ben also read-ASP b. *John kan-le baozhi, Ben ye du-le Φ . John read-ASP newspaper Ben also read (a synonym for kan)-ASP
(2) [TP Ben [vP kan-le [VP (kan-le baozhi).]]] Ben read-ASP read-ASP newspaper (3) *John read the newspaper, and Ben read Φ , too.

References Liu C. (2014) A modular theory of radical pro drop. [dissertation] Harvard University Yuan B. & Zhang L. (2020) An incremental model of L2 speech production mechanisms. *International Journal of Bilingualism*. 24(4)

The value of sign and print: Language proficiency predicts deaf signers' occupational prestige and income

Ye qiu Zheng (Erasmus University Rotterdam), Hao Lin (Shanghai international Studies University) and Yan Gu (University of Essex & University College London).

Language impacts immigrants' earnings and financial wealth concurrently and longitudinally (Zheng, Gu, Backus, & van Soest, 2022). However, no research has investigated deaf signers' language proficiency and their occupational prestige and income, though deaf signers are disadvantageous in labour market and their language use is of policy interest.

109 Shanghai Chinese Sign Language (CSL)-Mandarin print deaf signers were paid to fill in an online Mandarin questionnaire about their language history (Li et al., 2020), employment and background information. As Zheng et al. (2022) found that results of language and earnings were generally robust when using a self-assessed language proficiency compared to a language placement test, we followed most past research to use a self-assessment "Please assess your Mandarin/CSL proficiency on a scale of 1-7 (1=completely do not speak, 7= native proficiency)".

Deaf signers in our sample had a high employment rate (96.3%). To analyse whether CSL sign and Mandarin print proficiency related to deaf signers' occupational prestige and income, we coded two binary dependent variables "prestigious_job" based on categories in Li (2005), and "high_income" (=1 if > 6000 yuan, average net income, Statistics Shanghai, 2020; otherwise=0). As deaf signers who had a higher CSL proficiency level also had a higher Mandarin print proficiency, $r=0.53$, $p<.001$, we used Mandarin print and CSL proficiency separately as a predictor for dependent variables rather than putting both in the same model to avoid multicollinearity. Given that age of acquisition (AoA) is a strong predictor for sign proficiency (Mayberry, 2011), for the model of Mandarin print, we controlled for the AoA of CSL as a proxy of CSL language proficiency (same for Mandarin in the model of CSL). We also controlled for demographic factors, type of schools attended, severity of hearing loss, etc. We used the GLM generalizes linear regression in R starting with the full model and then the backwards step function based on AIC values to choose the best fit model. Results showed that a higher education ($\beta=1.22$, $p<.001$) and CSL proficiency level ($\beta=.44$, $p=.024$) predicted a more prestigious job. In addition, CSL proficiency ($\beta=.75$, $p=.034$) and education ($\beta=1.10$, $p=.045$) significantly predicted deaf signers' probability of having a high income, whereas Mandarin print proficiency was marginally significant ($\beta=.76$, $p=.049$).

Furthermore, as a sensitivity analysis, we coded the proficiency level (1-7) of CSL or Mandarin print as 'high' if the self-assessed score was 6 or 7, otherwise 'low' for the rest. This time we were able to put both CSL and Mandarin proficiency (high/low) in the same model together with other factors. The final model showed that only CSL proficiency level ($\beta=1.11$, $p=.022$) and education ($\beta=1.19$, $p<.001$) predicted whether having a more prestigious job, and a higher CSL proficiency ($\beta=1.88$, $p=.023$) and education level ($\beta=1.06$, $p=.043$) predicted a higher income.

Our study for the first time showed the relationship between deaf bilinguals' language proficiency and their occupational prestige and income, with strong implications for policymakers, such as the emphasis on the importance of learning sign language in deaf education.

Award Winners

Distinguished Scholar Award

Paul Meara (Swansea University)

LAB Award

Yolanda Xavier (NOVA University Lisbon)

Student stipend

- Hatice Akgün (Boğaziçi University)
- Maria Jose Centelles Cuart (University of Barcelona)
- Bridget Murphy (University of Barcelona)
- Louise Shepperd (University of York)
- Jing Zhou (Auckland University of Technology)

Sponsors

Benjamins

<https://benjamins.com/conferences/eurosla.2023>

Multilingual Matters

<https://www.multilingual-matters.com/>