Corpus-Linguistic and Cognitive Approaches to Determiner Usage in Chinese Student Writing Testing the Fluctuation Hypothesis

Sven Albrecht (TU Chemnitz, Germany)

This study investigates determiner usage by Chinese learners of English, taking account of the Fluctuation Hypothesis by Ionin, Ko, and Wexler (2004), which states that learners' article choice fluctuates for mixed settings of the semantic parameters definiteness and specificity. In an attempt to combine corpus-linguistic and cognitive approaches, the study employs a dual methodology: corpus data is complemented by data collected through a forced choice elicitation test. The forced choice elicitation test data was obtained using an online survey with 39 participants, most of whom are students at Sun Yat-sen University. The corpus data, from which three 100 sentence random samples were analyzed, was taken from the Sun Yat-sen University Corpus (SYSU-C), compiled by Küchler (2015).

The semantic concept of definiteness can be found in most languages. However, the way this concept is grammaticalized in different languages varies strongly and the realization can take various lexical, morphological or positional forms. Some languages, such as English or German, have a fully functional grammatical category of articles that encode definiteness, while in contrast, other languages, such as Samoan, encode specificity. Other languages, such as Korean or Russian, do not have an article category. Definiteness in these languages is expressed by different means, for example demonstratives.

In the area of first language acquisition, there has been strong evidence suggesting the existence of unconscious knowledge of language that lead Chomsky to postulating the notion of Universal Grammar (Chomsky, 1965, 1986). In addition, there has been research that suggest an effect of UG and its parameter setting model in second language acquisition as well (see Schwartz and Sprouse, 1994, 1996). Investigations of this effect revealed that L2 learners of English whose L1 does not have a fully functional grammatical category of articles have difficulties acquiring the correct usage of the English article system. Ultimately, that lead to the postulation of the Fluctuation Hypothesis (Ionin, 2003; Ionin et al., 2004), which states that until learners have received sufficient input to set the article choice parameters correctly, they will fluctuate in certain environments of these parameters.

Based on the assumption of the learners' full access to universal grammar, Ionin et al. propose the Article Choice Parameter, presented in (1), that determines how articles are selected in languages. While English (neglecting the colloquial use of referential this) has the first setting, there are other languages, like Samoan, that have the second setting (Ionin et al., 2004, p. 12).

(1) The Article Choice Parameter (for two-article languages) A language that has two articles distinguishes them as follows: The Definiteness Setting: Articles are distinguished on the basis of definiteness. The Specificity Setting: Articles are distinguished on the basis of specificity.

(Ionin et al., 2004, p. 12)

Resulting from the two possible settings of the Article Choice Parameter, Ionin et al. predict that learners go through a period of fluctuation before sufficient language input leads them to setting the parameter correctly. The fluctuation is likely only to occur for mixed parameter values ([+definite] [-specific] and [-definite] [+specific]) and not for either positive or negative parameter values ([+definite] [+specific] and [-definite] [-specific]). These predictions, as seen in table 1, are consistent with Lyons (1999), who argued that languages with articles mark either definiteness or specificity. According to the definition of the Fluctuation Hypothesis given below, the learners' interlanguage grammars are constraint by UG, however a single learner might have two rivaling interlanguage grammars in the process of language acquisition.

(2) The Fluctuation Hypothesis (FH) for L2 article choice:
1.   L2-learners have full access to the features that can underlie article choice cross-linguistically: the features [+definite] and [+specific].
2.   L2-learners fluctuate between dividing English articles on the basis of definiteness vs. specificity, until the input leads them to choose the definiteness option.
(from Ionin et al. (2009), based on Ionin et al. (2004, p. 8))

As Ionin et al. note, the explanation that the FH offers for non-random article errors is, however, limited to contexts where no other features, such as partitivity, interact. This is mirrored in their means of data collection, as the forced choice elicitation tasks carefully controlled for these circumstances.

	[+definite] (target: the)	[-definite] (target: a)
[+specific]	correct use or *the*	overuse of *the*
[-specific]	overuse of *a*	correct use of *a*

Table 1: Predictions for Article Choice in Chinese L2 English (from Ionin et al. (2004))

Related to article misuse, as predicted by the Fluctuation Hypothesis, article omission, which can also be described as overuse of the zero article, is frequently reported. Different studies in SLA have found that even advanced learners whose L1 lacks articles exhibit an overuse of the zero article in their interlanguage grammars (Bergeron-Matoba, 2007, p. 7). To account for this phenomenon, there are two opposing views. On the one hand, there are researchers, utilizing a no UG access approach, arguing that the variability in the interlanguage grammars is due to some form of either temporary or permanent impairment (Prévost and White, 2000, p. 108). On the other hand, there are researchers, following an approach permitting UG access, claiming that the interlanguage grammars contain abstract functional categories and features. The variability found in empirical studies results from difficulties mapping that abstract knowledge to the surface morphological manifestation (Prévost and White (2000, p. 108), following Haznedar and Schwartz (1997)). They state that:

(3) [t]he Missing Surface Inflection Hypothesis (MSIH) proposes that L2 learners have unconscious knowledge of the functional projections and features underlying tense and agreement. However, learners sometimes have a problem with realization of surface morphology, such that they resort to non-finite forms [...]. (Prévost and White, 2000, p. 103)

Combining Full Transfer/Full Access (Schwartz and Sprouse, 1996) with missing surface inflection (Haznedar, 2001), the MSIH states that the initial state of the L2 grammar is the final state of the L1 grammar, which then needs to be restructured as it fails to account for input data. In contrast to first language acquisition, the activation of functional categories is dependent on the acquisition of morphology (Haznedar, 2001, p. 279). Hence, issues in L2 production stem from morpholexical aspects rather than systematic syntactic deficits (Haznedar, 2001, p. 280). This hypothesis offers a possible explanation for article omission errors a posteriori and has therefore been debated controversially (Granfeldt, 2000).

The learner groups that have been researched extensively in previous studies include Korean (Crosthwaite, 2014; Ionin et al., 2009), Russian (Chrabaszcz and Jiang, 2014; Ionin, 2003; Ionin, Zubizarreta, et al., 2008) and, to a lesser extent, Turkish (Dikilitaş and Altay, 2011), Japanese and Chinese native speakers (Snape et al., 2006). However, especially L1 Chinese learners are of particular interest, since there is research suggesting that the grammatical category of articles is in the process of grammaticalization in modern Chinese (P. Chen, 2003, 2004; C.-T. J. Huang, 1987; Li and Bisang, 2012). Consequently, it would be interesting to see, how Chinese learners of English perform in their article choice in comparison to other learners with an article-less L1 as well as to learners whose L1 does have articles (c.f. Snape et al., 2006; White, 2008).

While most of these studies have used forced choice elicitation tasks as means of data collection, it might be interesting to study the usage in academic writing as well, because it is rather natural language that was not collected in a focused test. Furthermore, the overall frequency of articles is highest in academic English across all genres (Biber et al., 1999, p. 267). Moreover, depending on the corpus used in the analysis, it can be investigated in how far some variation in the usage patterns can possibly be attributed to different levels of proficiency.

	[+definite] (target: *the*)		[-definite] (target: *a*)	
L1 Chinese (n=37)	the	a	the	a
[+specific]	84 %	14 %	10 %	90 %
[-specific]	89 %	9 %	4 %	93 %

Table 2: Total Article Choice

The results of the analysis of this study do not indicate fluctuation in the learners' article choice as predicted by the Fluctuation Hypothesis, with low overall error frequencies (c.f. Table 2). In more detail, a regression analysis revealed that the most significant influence on article choice is exerted by an interaction group of specificity, scope and speaker

knowledge. Moreover, noticeable usage differences are found between participants from the fields humanities and engineering. The results furthermore indicate increased article substitution error rates with partitive indefinites and simple (in)definites, which is unexpected (c.f. Table 3). Reasons explaining these results could either be found in the ongoing grammaticalization process of articles in Mandarin Chinese or the proficiency level of the participants, since the Fluctuation Hypothesis is an explicit learner phenomenon. Similar to previous studies on determiner usage, omission errors are generally low in the data.

	simple definitiv	partitivity	difficult
correct	81 %	78 %	48 %
substitution	18 %	22 %	40 %
omission	1 %	1 %	13 %

Table 3: Results for Simple Definites, Partitivity and Lexically Complex Contexts

Tentatively interpreting the results, this study finds that specificity, in combination with scope and speaker knowledge, has an influence on learners' article choice, even though article choice in English is solely governed by definiteness. Additionally, different proficiency levels seem to favor different error patterns. However, these findings should not be overgeneralized, as the apply only to the relatively small datasets that control for numerous sociolinguistic variables, such as age and education. In the future, it would therefore be interesting to carry out a larger scale study with regression analysis in mind from the very beginning, as this would allow a more definite answer to the possible influence of the different variables.

References

Bergeron-Matoba, J. (2007). *Acquisition of the English article system in SLA and the Missing Surface Inflection Hypothesis*. The University of Queensland Working Papers in Linguistics, 1.

Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E., & Quirk, R. (1999). *Longman grammar of spoken and written English*. MIT Press.

Chen, P. (2003). *Indefinite determiner introducing definite referent: a special use of 'yi 'one'+classifier' in Chinese*. Lingua, 113(12), 1169–1184.

Chen, P. (2004). *Identifiability and definiteness in Chinese*. Linguistics, 42(6), 1129–1184.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, Mass: M.I.T. Press.

Chomsky, N. (1986). *Knowledge of language: its nature, origin, and use*. New York: Praeger.

Chrabaszcz, A. & Jiang, N. (2014). *The role of the native language in the use of the English nongeneric definite article by L2 learners: A cross-linguistic comparison*. Second Language Research, 30(3), 351–379.

Crosthwaite, P. R. (2014). *Definite Discourse-New Reference in L1 and L2: A Study of Bridging in Mandarin, Korean, and English*. Language Learning, 64(3), 456–492.

Dikilitaş, K. & Altay, M. (2011). *Acquisition sequence of four categories of non-generic use*

of the English definite article THE by Turkish speakers. Novitas-ROYAL (research on Youth and Language), 5(2), 183–198.

Granfeldt, J. (2000). *The acquisition of the determiner phrase in bilingual and second language French.* Bilingualism: Language and cognition, 3(03), 263–280.

Haznedar, B. (2001). *The acquisition of the IP system in child L2 English*. Studies in second language acquisition, 23(01), 1–39.

Haznedar, B. & Schwartz, B. D. (1997). *Are there optional infinitives in child L2 acquisition*. In Proceedings of the 21st annual boston university conference on language development (pp. 257–268). Somerville, MA: Cascadilla Press.

Huang, C.-T. J. (1987). *Existential sentences in Chinese and (in) definiteness*. The representation of (in) definiteness, 226–253.

Ionin, T. (2003). *Article semantics in second language acquisition* (Doctoral dissertation, Massachusetts Institute of Technology).

Ionin, T., Ko, H., & Wexler, K. (2004). *Article Semantics in L2 Acquisition: The Role of Specificity*. Language Acquisition, 12(1), 3–69.

Ionin, T., Ko, H., & Wexler, K. (2009). *L2-acquisition of English articles by Korean speakers*. The handbook of East Asian psycholinguistics, 3, 286-304.

Ionin, T., Zubizarreta, M. L., & Maldonado, S. B. (2008). *Sources of linguistic knowledge in the second language acquisition of English articles*. Lingua, 118(4), 554– 576.

Küchler, J. (2015). *Usages of may and will in Chinese and German Student Writings.* In J. Schmied (Ed.), Academic writing for South Eastern Europe (pp. 99–118). Göttingen: Cuvillier.

Li, X. & Bisang, W. (2012). *Classifiers in Sinitic languages: From individuation to definiteness-marking*. Lingua, 122(4), 335–355.

Lyons, C. (1999). *Definiteness*. (1. publ.). Cambridge: Cambridge Univ. Press.

Prévost, P. & White, L. (2000). *Missing Surface Inflection or Impairment in second language acquisition? Evidence from tense and agreement*. Second Language Research, 16 (2), 103–133.

Schwartz, B. D. & Sprouse, R. (1994). *Word order and nominative case in nonnative language acquisition: a longitudinal study of (L1 Turkish) German interlanguage*. Language acquisition studies in generative grammar, 31(4), 71–89.

Schwartz, B. D. & Sprouse, R. (1996). *L2 cognitive states and the Full Transfer/Full Access model*. Second Language Research, 12(1), 40–72.

Snape, N., Ting, H.-C., & Leung, Y.-k. I. (2006). *Comparing Chinese, Japanese and Spanish Speakers in L2 English Article Acquisition: Evidence against the Fluctuation Hypothesis?*

White, L. (2008). *Different? Yes. Fundamentally? No. Definiteness effects in the L2 English of Mandarin speakers*. In Proceedings of the 9th generative approaches to second language acquisition conference (gasla 2007) (pp. 251–261).