Stance in the BAWE Corpus: New Revelations from Multidimensional Analysis

Hilary Nesi and Sheena Gardner (Coventry University, UK)

The British Academic Written English (BAWE) corpus contains about 6.5 million words of proficient university student writing, categorised in terms of 'genre families' and distributed fairly equally across levels of study and disciplinary groupings. The corpus has been examined from various perspectives, including multidimensional analysis (MDA) (Nesi & Gardner 2012) using Biber's 1988 dimensions. This analysis revealed clear distinctions between texts produced by first, second and final year undergraduates and Masters students, and between texts belonging to different disciplinary groupings. In terms of stance, epistemic modality was the most discriminating feature, being considered a tool for persuasion on Dimension 4. Published academic prose from the Lancaster-Oslo-Bergen (LOB) Corpus had a neutral score on Dimension 4 in Biber's own 1988 study, but scores for BAWE assignments were negative on Dimension 4, and became increasingly so across levels of study (Table 1). Scores for Arts and Humanities assignments were particularly low (Table 2).

Level	Score
1st year undergraduate	-1.4
2nd year undergraduate	-1.4
3rd year undergraduate	-1.5
Masters	-2.0

Table 1: BAWE Dimension 4 scores by level

Disciplinary Grouping	Score
Physical Sciences	-1.2
Social Sciences	-1.3
Life Sciences	-1.5
Arts and Humanities	-2.3

Table 2: BAWE Dimension 4 scores by disciplinary grouping

Biber's 1988 dimensions still have validity, because the results from studies which employ them can easily be compared. For example, the 1988 dimension 4 scores for BAWE and Biber's LOB subcorpus can be taken as an interesting indicator of differences between student and expert writing. Student writing below PhD level appears to have fewer overtly persuasive features than expert writing, perhaps because undergraduate and Masters students tend to discuss the work of others rather than presenting their own original research.

Stance theory has developed considerably since 1988, and features have now been added to Biber's tagset, enabling clearer distinctions to be made between registers which differ predominantly in terms of stance. Biber (2006) uses this tagset as the basis for a detailed linguistic description of registers encountered by university students, such as classroom teaching, coursepacks and institutional writing, but does not include any analysis of texts (such as assignments) produced by students themselves. A new multidimensional analysis of BAWE (BAWE2016) has built on this work, using the extended tagset to create new dimensions specific to university student writing. These bespoke dimensions distinguish between BAWE texts grouped by discipline and genre family, without reference to registers in other corpora.

Following MDA methodology, 39 linguistic features were retained and four new factors were identified (Gardner, Nesi & Biber, under review), to be treated as dimensions along which the BAWE registers could be situated. Two of the four new dimensions depend strongly on stance features. The variation along each dimension was then

interpreted by qualitative means, manually examining high- and low-scoring texts to ascertain the communicative effect created by the clustering of features.

Stance adverbials and stance nouns controlling *that*-clauses clustered with 3rd person pronouns, proper nouns and communication verbs at the negative pole of Dimension 1. Together these have been interpreted as indicating 'stance toward the work of others'. Essays tend to have the lowest scores on this dimension, particularly in Arts and Humanities disciplines (Table 3).

Discipline	Score
Comparative American Studies (n=74)	-6.78
History (n=95)	-6.89
English (n=106)	-7.38
Classics (n=82)	-10.53
Philosophy (n=106)	-10.67

Table 3: Disciplines with the lowest scores on Dimension 1

Stance nouns controlling *that*-clauses were relatively rare in Biber's university language corpus, with the exception of *fact* (Biber 2006:112). In texts with low scores on Dimension 1 most of these nouns have an epistemic function, but some fall into the category of 'Communication' or 'Attitude' nouns (Biber 2006:93). Stance nouns in the 20 lowest-scoring texts are listed in Table 4.

Likelihood		**Certainty**		**Communication**		**Attitude**	
Belief	9	Certainty	2	Argument	1	Fear	4
Assumption	2	Conclusion	3	Confession	1	Feeling	1
Claim	1	Fact	30	Explanation	1	Hope	1
Hint	1	Precept	1	Interpretation	1	Illusion	1
Hypothesis	1	Principle	1	Intimation	1	Indication	2
Idea	11	Proof	1	Objection	1	Understanding	2
Interpretation	1	Realization	1	Proposal	1	View	9
Intimation	1	Revelation	2	Proposition	3		
Notion	3	Statement	1				
Possibility	1						
Premise	3						
Presumption	1						
Sign	2						
Suggestion	4						
Total	41	**Total**	42	**Total**	10	**Total**	20

Table 4: Stance nouns in the 20 lowest-scoring texts on Dimension 1.

Likewise stance adverbials mainly serve as epistemic devices, particularly to indicate certainty. Table 5 lists these items under the categories identified by Biber (2006:92).

Likelihood		**Certainty**		**Style**		**Attitude**	
Apparently	3	Actually	33	According to	9	Importantly	5
Arguably	3	Always	26	Clearly	13	Ironically	2
Evidently	3	Certainly	15	Frankly	1	Paradoxically	1
Kind of	13	Decidedly	2	Fundamentally	3	Surprisingly	1
Maybe	1	Doubtless	1	Generally	4	Tragically	2
Perhaps	10	Evidently	3	Mainly	4	Unfortunately	1

Possibly	3	Indeed	21	Primarily	4		
Presumably	3	In fact	17	Technically	1		
Probably	12	Inevitably	5	Typically	2		
Sort of	3	Never	33	Usually	3		
		No doubt	1				
		Obviously	3				
		Of course	4				
		Really	20				
		Undeniably	1				
		Undoubtedly	3				
		Unquestionably	1				
Total	54	**Total**	189	**Total**	44	**Total**	12

Table 5: Stance adverbials in the 20 lowest-scoring texts on Dimension 1.

At the positive pole of Dimension 2, stance verbs controlling *to*- and *that*-clauses occur alongside mental verbs, *that*-deletions, 1st person pronouns and past tense verbs. Together, these have been interpreted as indicating 'Personal stance'. Genres that are not typically considered 'academic' - reflective writing, legal Problem Questions discussing the implications of everyday situations, and Empathy Writing engaging with non-experts – tend to contain more of these features. Scores are highest in the soft and applied disciplines (Table 6).

Discipline	**Score**
Philosophy (n=106)	8.24
Health (n=81)	7.60
Psychology (n=95)	5.51
Linguistics (n=115)	5.29
Classics (n=82)	4.55

Table 6: Disciplines with the highest scores on Dimension 2

Stance verbs in the 20 highest-scoring texts on Dimension 2 are shown in Tables 7 and 8, using the categories identified by Biber (2006:92).

Likelihood		**Certainty**		**Communication**		**Attitude**	
Assume	1	Conclude	3	Argue	1	Agree	2
Believe	15	Discover	3	Claim	1	Ensure	2
Imagine	1	Find	6	Explain	4	Expect	1
Think	3	Know	3	Insist	1	Feel	38
		Learn	8	Mention	1	Forget	1
		Mean	9	Propose	1	Hope	5
		Notice	4	Say	16	Wish	4
		Observe	2	State	2	Worry	3
		Realise	10	Suggest	1		
		Recognise	1	Tell	1		
		See	4				
		Show	7				
		Understand	1				
Total	20	**Total**	53	**Total**	20	**Total**	56

Table 7: Stance verbs controlling *that*-clauses in the 20 highest-scoring texts on Dimension 2.

Probability/ Cognition/ Perception		**Desire/ Intention/ Decision**		**Speech Act/ Communication**		**Causation/ Modality/ Effort**	
Appear	3	Agree	2	Ask	3	Attempt	5
Believe	1	Choose	9	Be said	1	Enable	22
Expect	4	Decide	7	Teach	7	Encourage	9
Find	1	Hope	3			Fail	2
Forget	1	Intend	2			Help	31
Know	1	Like	6			Manage	5
Learn	8	Mean	3			Persuade	1
Seem	12	Need	37			Require	4
Tend	2	Plan	2			Try	29
Suppose	3	Prefer	3				
		Prepare	1				
		Want	78				
		Wish	3				
Total	28	**Total**	156	**Total**	11	**Total**	108

Table 8: Stance verbs controlling *to*-clauses in the 20 highest-scoring texts on Dimension 2.

At the positive pole of Dimension 2 stance features indicate a different register, more expressive of personal attitudes, desires and efforts. Stance verbs suit texts that are less like conventional university assignments because they appeal to readers at a more emotional level, and are essentially designed to persuade.

These findings indicate some distinct differences in the expression of stance across genres and disciplines, and provide evidence to support the more general claims made in prior studies. For example Hyland has argued that the soft disciplines are "more inclined to explicitly recognise the role of human agency in constructing knowledge" (2002:124) and to take "more involved and personal positions than those in the science and engineering fields" (2015:34).

Epistemic nouns and adverbials are particularly prevalent in soft discipline Essays, because argumentative claims in this register tend to be developed on the basis of certainty and likelihood (see, for example McEnery & Kifle 2002). Stance nouns which control *that* -clauses also enable arguments to unfold, by serving as "shell nouns' (Hunston & Francis 1999) or 'signalling nouns' (Flowerdew & Forest 2014) which label and encapsulate information presented elsewhere in the text. Epistemic and attitude adverbials were strongly associated with spoken registers in Biber's university language corpus (2006:103-4), but in texts at the negative pole of Dimension 1 they are used to boost and hedge, signalling the writer's direct involvement in the text. Again, this is typical of the soft disciplines where, in contrast to the hard sciences, there is "less control of variables, more diversity of research outcomes, and fewer clear bases for accepting claims" (Hyland 2015:35).

Our study has also shown that there is a distinctive register found in student writing that employs first person pronouns and 'stance' *to*- and *that*- clauses. Future studies could investigate the extent to which the personal stance registers of student writing are also

found in expert writing, perhaps extending beyond research writing to include texts produced in the professions.

References

Biber, D. (2006). *University Language: A Corpus-based Study of Spoken and Written Registers*. Amsterdam: John Benjamins.

Flowerdew, J. & Forest, R. (2014). *Signalling Nouns in English: A Corpus-Based Discourse Approach*. Cambridge: Cambridge University Press.

Gardner, S. Nesi, H. & Biber, D. (under review) Discipline, level, genre: Integrating multiple situational perspectives in a new MD analysis of university student writing. Hyland, K. (2015). Genre, discipline and identity. *Journal of English for Academic Purposes* 19 32-43.

Hyland, K. (2002). Activity and evaluation: Reporting practices in academic writing. In J. Flowerdew (Ed.), *Academic Discourse* (pp. 115-130) Harlow, Essex: Pearson Education.

Hunston, S. & Francis, G. (1999). *Pattern Grammar*. Amsterdam: John Benjamins.

McEnery, T., & Kifle, N.A. (2002). Epistemic modality in argumentative essays of second language writers. In J. Flowerdew (Ed.), *Academic Discourse* (pp. 182-195). Harlow, Essex: Pearson Education.

Nesi, H. & Gardner, S. (2012). *Genres across the Disciplines: Student writing in higher education.* Cambridge: Cambridge University Press.