

7 The variability of lexical diversity and its relationship to learning style

Paul Booth
Kingston University
p.booth@kingston.ac.uk

When L2 learners' texts are measured for lexical diversity we tend to find greater diversity with years of instruction. However, at the higher levels of proficiency, this linear relationship can disappear (Jarvis 2002; Read 2005). At higher proficiency levels we tend to find a wide variety of scores, which is difficult to explain. Therefore, instead of examining mean scores, which hide the wide variation, this study concentrates on the variance in learners' lexical performance. In this study learning style is used as way of understanding how variance in lexical diversity could be related to the individual differences in memory and analysis.

Lexical diversity was measured by using Meara and Miralpeix's *D-Tools* (Meara & Miralpeix 2007a). The learners' scores for lexical diversity (*D*) were taken from written texts produced under timed conditions. One group of learners ($n = 60$) wrote a descriptive text based on a cartoon story. The other group ($n = 62$) wrote a text elicited via a discursive type question on the topic of globalisation. Because low proficiency learners' texts tend to be low in lexical diversity, the learners were grouped into low proficiency (IELTS ≤ 5.5 , $n = 51$) and high proficiency (IELTS ≥ 6.0 , $n = 71$). Skehan's (1998) memory-analysis framework was used to categorise learners according to learning style because it has been shown to illuminate aspects of aptitude variability in L2 proficiency. The learners were tested for learning style using language aptitude tests (LAT) (Meara, Milton, and Lorenzo-Duz, 2001) of LAT B for memory and LAT C for analysis. Based on Meara et al's (2001) normative data, learners were categorised into bottom, middle, and top scores in memory and analysis. The lexical diversity scores were then analysed in relation to learners' performance on the memory and analysis tests. The coefficient of variation (CV) was used to measure the variability of the mean diversity (*D*) scores: the greater the variability, the greater the coefficient. The CV was used to compare the dispersion of scores across the different sub-groups.

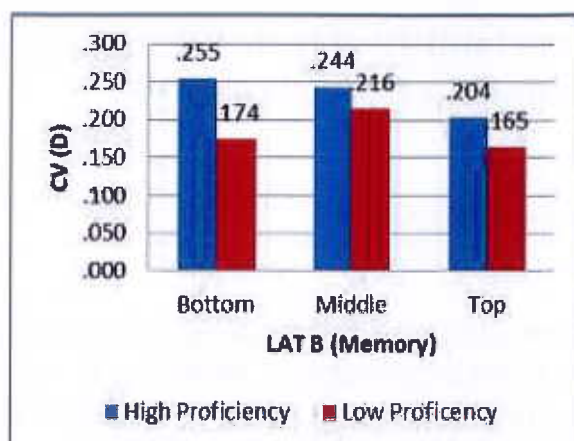


Figure 1: Variability (CV) in lexical diversity (D) and memory

In Figure 1, high proficiency learners' texts become more homogenous in diversity as memory scores increase. However, low proficiency learners do not mirror this pattern.

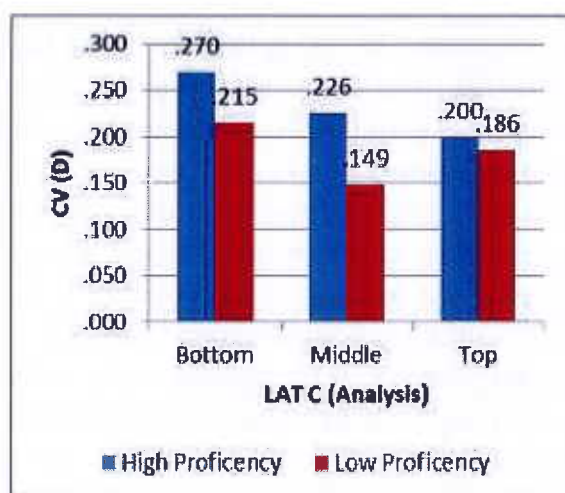


Figure 2: Variability (CV) in lexical diversity (D) and analysis

In Figure 2, a more marked pattern in relation to variability and analysis is found when high proficiency learners' diversity scores are calculated. Again, at low proficiency, the pattern is less clear. Levene's test was used to see whether differences in the variance of diversity scores in the analysis sub groups were significant. Levene's test for homogeneity of variances was significant $F(2, 68) = 4.804, p < 0.05$. This means that at high proficiency the higher analysis scores, the lesser the variability in lexical diversity. Learners who are particularly able to detect grammatical patterns show less variability than those who are less able in analysis.

What we see here is that lexical diversity scores from the highly analytic learners tend to cluster together. In other words, these learners tend to be more uniform in their recycling of words. It could mean that there is a ceiling effect in lexical diversity, above which greater diversity is not related to greater complexity. Meara and Miralpeix (2007b) argue that grammar words are the most recycled words which influence sentence structure. Sentence structure is related to sentence complexity. Learners who are less prone to grammaticise their language may be more telegraphic in their writing by avoiding certain function words and overusing content words which would increase diversity. On the other hand, they could overuse some function words with more generalised meaning (Ijaz, 1986). This may tell us why there is so much variability in lexical diversity from L2 learners. Over or under recycling of function words would produce more variability which could indicate less complexity in sentences. Kormos and Trebits (2012) hypothesise that greater grammatical sensitivity relates to clausal complexity. In order to create complexity in English, learners would be required to recycle a certain amount of grammar words. Higher grammatical sensitivity seems to encourage a more uniform recycling of vocabulary, including grammar words, which could be related to sentence complexity. The current study highlights a relationship between the variability of lexical diversity, learning style, and proficiency. Simply concentrating on the mean diversity scores glosses over the fact that there is wide variability in learners' lexical performance and that a shift to looking at individual differences is valuable in understanding the messy details in group performances.

References

- Helene Ijaz.** 1986. Linguistic and cognitive determinants of lexical acquisition in a second language. In *Language Learning* vol 36, 1986, pp401-51.
- Scott Jarvis.** 2002. Short texts, best-fitting curves and new measures of lexical diversity. In *Language Testing* vol 19 (1), 2002, pp57-84.
- Judit Kormos & Anna Trebits.** 2012. The role of task complexity, modality, and aptitude in narrative task performance. In *Language Learning* vol 62 (2), 2012, pp439-472.
- Paul Meara, James Milton, & Nuria Lorenzo-Duz.** 2001. *Language aptitude tests*. Newbury: Express.
- Paul Meara & Imma Miralpeix.** 2007a. *D_Tools*. v2.0. [Online]
<http://www.lognostics.co.uk/tools/index.htm>.
- Paul Meara & Imma Miralpeix.** 2007b. *D_Tools: The Manual* [Online]
<http://www.lognostics.co.uk/tools/index.htm> [Accessed 26 October 2007]
- Peter Skehan.** 1998. *A cognitive approach to language learning*. Oxford: Oxford University Press.

John Read. 2005. *Applying lexical statistics to the IELTS speaking test.* In Research Notes 20, 2005, pp12-16.