

# **The Spoken Proficiency of Future English Teachers in Brazil: a Corpus-based Analysis**

Paula Tavares PINTO-PAIVA  
Universidade Estadual Paulista - UNESP  
Department of Modern Languages  
São José do Rio Preto, Brazil  
paula@ibilce.unesp.br / paulapaivaibilce@gmail.com

The teaching of a foreign language to secondary school students in Brazil has mostly been based on reading comprehension texts as required by the National Curriculum Parameters guidelines (1998). Although this practice has encouraged the reading of several genres in English and, more recently, in Spanish, it has also discouraged students from practicing the spoken foreign languages at this stage. One of the results of this approach is that undergraduate students who are admitted to language and linguistics courses at the university level lack the oral communication skills in the foreign language they have chosen to study. In addition, students do not generally appreciate the communicative approach during their time at university. Therefore, when they graduate, they are not employed by schools which emphasize the development of oral skills (Consolo, 2005; Consolo, Guerreiro & Hatugai, 2001).

Based on the findings of a research group which analyzed the linguistic competence of foreign language teachers, the proposal entitled ENAPLE-CCC (Education and Foreign Language Learning: Beliefs, Competencies and Constructs) was put forward involving the creation and implementation of a “Proficiency Test for Teachers of English as a Foreign Language” (EPPLE). One section of the EPPLE is an oral interview in which, among other items, the speaker’s lexical competence is evaluated. Rodrigues (2009) defines lexical competence as “the ability to recognize and use words in a language the way native speakers do”. Lexical competence includes the understanding of the relationships between different families of words and the placement of the most common words.

A pilot study was carried out which involved the compilation of a spoken corpus. The corpus consisted of the oral transcripts of the interviews conducted with the last year’s undergraduate students in a public university in Sao Paulo state. This corpus was analyzed using the theoretical and methodological approach of Corpus Linguistics (Sinclair, 1991; Berber Sardinha, 2000) in order to test whether students’ final marks corresponded to the high or low lexical variation of their oral productions. According to Baker (1996) and McCarthy (1990), a low type/token ratio would suggest a minor lexical variation whereas a higher rate would suggest the use of diversified vocabulary. We hypothesized that the interviewees who had used a higher lexical variety would be the ones who had had the highest marks in the oral assessment or vice-versa. Besides the analysis of lexical variation, we also retrieved the keywords for each interview in order to analyze, in a more detailed way, the use of lexical words as opposed to function words.

For the purpose of this pilot study, the interviews that had been transcribed were compiled as a corpus of spoken language according to the criteria listed by Berber Sardinha (2000, 2004) as fundamental to the compilation of a computerized corpus: origin and authenticity (transcribed interviews in a real context of a proficiency exam), purpose (lexical analysis), format (texts in a plain text format), representation (oral language) and extension (9,080 items/a small corpus in that year). Using the software WordSmith Tools (Scott, 1999), the statistics for each interview, were generated. The results were compared to the students' final grades and they showed a correlation between the type/token ratio numbers, the use of lexical words and student's final mark in 75% of the group.

We must emphasize that the analysis proposed in this pilot study did not take into account other aspects of oral language such as pronunciation, pauses and syntactic structure used in the speech of each respondent. This was a lexical analysis to assess the lexical component of oral interviews. For this reason, we cannot state that students' final grades were only connected to the use of vocabulary they used in their tests.

In order to collect more data for type/token analyses and to provide more empirical data to check validity and reliability of EPPL, since 2009 we have been compiling a corpus of spoken interviews of undergraduate students in Languages and Teaching program, in their last course year. This corpus presently contains 19,462 items and is still being enlarged as the interviews are transcribed. Up until now we have transcribed the interviews from the years 2009 and 2010. We have also carried out small research projects with undergraduate students who have investigated the frequency of words and the keywords in this corpus. We believe that by the end of this year we will have finished the compilation of this corpus and have more tangible results to show.

## References

- Baker, M. (1996). 'Corpus-based translation studies: the challenges that lie ahead.' In: Somers, H. (ed). *Terminology, LSP and translation studies in language engineering. In honour of Juan C. Sager*. Amsterdam: John Benjamins. 177-186.
- Consolo, D. A., Guerreiro, G. M. S., & Hatugai, M. R. (2001). *Mapeamento da competência lingüístico-comunicativa em línguas estrangeiras de alunos de Licenciatura em Letras*. Relatório final de projeto de iniciação científica. UNESP/São José do Rio Preto.
- Consolo, D. A. (2004). 'Construção de um instrumento de avaliação da proficiência oral do professor de língua estrangeira.' *Trabalhos em Lingüística Aplicada* 43.
- Berber Sardinha, A. P. (2000). 'Lingüística de corpus: histórico e problemática.' *DELTA: documentação de estudos em lingüística teórica e aplicada*, São Paulo, 16: 323-367.
- Berber Sardinha, T. (2004). *Lingüística de Corpus*. Barueri, São Paulo: Manole.
- McCarthy, M. *Vocabulary*. (1990). Oxford: Oxford University Press.
- Rodrigues, D. F. A. (2009). 'Competência lexical e o ensino de oralidade em aulas de língua inglesa.' In: Consolo, D. A. & Silva, V. L. T. (eds). *Olhares sobre competências do professor de língua estrangeira: da formação ao desempenho profissional*.
- Sinclair, J. (1991). *Corpus, Concordance and Collocation*. Oxford: Oxford University Press.