

Correction Notice

In Vol. 11.1, 2014 of *Taiwan Journal of TESOL*, the title of “Constructing the Taiwanese Component of the Louvain International Database of Spoken English Interlanguage (LINDSEI)” on page 31 written by Lan-fen Huang was misprinted. *Taiwan Journal of TESOL* hereby re-prints the correct page in this volume.

Taiwan Journal of TESOL
Vol. 11.1, 31-74, 2014

CONSTRUCTING THE TAIWANESE COMPONENT OF THE LOUVAIN INTERNATIONAL DATABASE OF SPOKEN ENGLISH INTERLANGUAGE (LINDSEI)

Lan-fen Huang

ABSTRACT

This paper reports the compilation of a corpus of Taiwanese students' spoken English, which is one of the sub-corpora of the Louvain International Database of Spoken English Interlanguage (LINDSEI) (Gilquin, De Cock, & Granger, 2010). LINDSEI is one of the largest corpora of learner speech. The compilation process follows the design criteria of LINDSEI so as to ensure comparability across the sub-corpora. The participants, procedures for data collection and process of transcription are all recorded. Fifty third- or fourth-year English majors in Taiwan were given recorded interviews in English. Each interview was accompanied by a profile containing information about such learner variables as age, gender, mother tongue, country, English learning context, knowledge of other foreign languages, and amount of time spent in English-speaking countries and such interviewer variables as gender, mother tongue, knowledge of foreign languages and degree of familiarity with the interviewees. Data on another variable, the learners' English proficiency level based on the results of international standardised tests, was collected; this is not available in other sub-corpora of LINDSEI. The participants' proficiency was similarly distributed across B1 to C1 levels in the Common European Framework of Reference. The structure of the Taiwanese sub-corpus is discussed in comparison with eleven other published sub-corpora. The preliminary investigation, using corpus-linguistic approaches, reveals overall statistical information about the Taiwanese component and Version 1 of LINDSEI. The lexical analyses of the top 50 words and chunks show the characteristics of spoken English in the Taiwanese sub-corpus. The contributions and research potential of this newly-developed learner corpus are discussed, followed by an example of Contrastive Interlanguage Analysis of the most common chunk, *I think*, in the Taiwanese learners' speech. The release of this learner corpus is merely the first step. It is hoped that more corpus research will be done on Taiwanese learners, that corpora of other speech genres will be compiled and that research results will contribute to relevant areas in Applied Linguistics.

Key Words: LINDSEI, interlanguage, learner corpus, Taiwanese learners of English, *I think*