THE DEVELOPMENT OF A HUNGARIAN–ENGLISH LEARNER SPEECH DATABASE AND A RELATED ANALYSIS OF FILLED PAUSES

Mária Gósy – Dorottya Gyarmathy – András Beke

Research Institute for Linguistics, Hungarian Academy of Sciences, Budapest
gosy.maria@nytud.mta.hu – gyarmathy.dorottya@gmail.com – beke.andras@gmail.com

Keywords: English–Hungarian database, three proficiency levels, five speech styles, annotations.

1. INTRODUCTION

The third revolution in the history of speech science has been identified as the development of large databases. By now, the importance and usefulness of such databases are unquestionable. They provide a unique possibility to do research on various speech phenomena from fine phonetic events to diverse characteristics of long narratives. Beside native speakers' databases language learner corpora have also started being developed. The size, contents, annotation level, and other characteristics of existing spoken corpora are very different. They focus on English as L2 and mainly contain written materials. The largest corpus of informal interviews of learners of English (from higher intermediate to advanced students) has been coordinated by the University of Louvain [1]. This corpus contains 554 interviews produced by learners with eleven first languages (Hungarian is not represented among them).

The present study has two principal purposes. (i) It intends to present the design and development of our learner database (HunEng-D). (ii) We present our research concerning filled pause patterns of Hungarian and English speech based on the database.

The two languages differ in a number of linguistic facts. Hungarian is an agglutinative language belonging to the Finno-Ugric language family. It has an extremely rich morphology with extensive affixation and, as a consequence, syntactic and semantic functions of noun phrases are primarily expressed via suffixes and postpositions. Case endings are used extensively with nouns, but pronouns, adjectives, and numerals also take case and number endings. Verbs also have a considerable number of prefixes and suffixes. It has an inventory of 39 phonemes, including also phonologically short and long ones (but there are no diphthongs, no neutral vowels, and no aspirated consonants).

2. THE LEARNER DATABASE (HUNENG-D)

This is the first spoken language learner database that (i) contains speech materials of Hungarian-speaking participants whose second language is English, (ii) is recorded under the same conditions and following the same protocol, and (iii) contains annotated speech materials. The aim of the project is to develop a large database of 150 monolingual speakers whose native language is Hungarian and whose command of English represents various L2 proficiency levels.

2.1. Participants

At the time of writing, the total recorded material involves 60 speakers (half of them females) and amounts to about 30 hours. Hungarian is the mother tongue of all participants; they learn English as L2 at school. 20 of them are basic learners (BL) with two to three years of L2 instruction at school (14–15-year-olds), another 20 of them (18–19-year-olds) are intermediate learners (IL) and the third group consists of 20 advanced learners (AL) (ages between 22 and 28 years). The L2 language proficiency of the youngest group is B1, of the second group is B2 while that of the third group is C1 according to the categories of the Common European Framework.

2.2. Recording protocol

The database contains various types of spontaneous speech materials including also a word list to be read. The protocol consists of 5 modules. 1. Narratives about the participant's life, family, job, and hobbies. 2. Narrative-like opinions about a topic of current interest, provided by the interviewer. The topics are adjusted to the participants' age. 3. Précis (summary of content), that is, directed spontaneous speech. The participant reads a story and then s/he has to summarize its content in his/her own words. 4. Two types of methods (a map task and a speech game) to elicit quasi-natural conversations between the participants. 5. A word list consisting of forty-eight Hungarian words and another forty-eight English words (with voiced and unvoiced stop consonants as the initial segments of the words).

2.3. Recording conditions

Recordings are invariably made in the same sound-proof booth (at the Research Institute for Linguistics), under identical technical conditions, digitally, direct to the computer using the same professional microphones. In all recordings (both in English and in Hungarian) the interviewer was the second author.

2.4. Annotation

The speech material has been manually annotated by two trained transcribers, phoneticians with high-level proficiency in English) while two authors have continuously double-checked the annotations. The transcription was done in Praat at two levels (phrase level and word level) according to the criteria and rules that had been developed.

3. FILLED PAUSE ANALYSIS

The term ‘filled pause’ refers in this paper to the phenomenon of diverse vocalic, nasal, mixed or other sound events inserted into the spontaneous utterances. Filled pauses do occur in spontaneous narratives in any language. They may be used for several reasons and in various functions, and occur in diverse verbal forms of the languages (e.g., [2]). English speakers use *uh* and *um*, Portuguese ones use *uum*, [ɐ], and [ə], Japanese speakers use *ano*, *e*, *eto*, *ma*, Basque speakers use *e*, *m*, *zera*, while Hungarians prefer neutral vowels [4, 5, 6], etc. Data of analysis of filled pauses in L1 and L2 speech show that they are more frequent and longer in L2 than in L1, and occur characteristically within the clause in L2 narratives (e.g., [3]).

The main theoretical interest that guided our research was to find out whether filled pauses show different occurrences and phonetic patterns in the two languages, whether they are flanked by lexical items or by silent periods, and whether they show different patterns suggesting different proficiency levels. Three hypotheses were formulated. (i) Filled pauses occur more frequently and with longer durations in L2 than in L1 narratives. (ii) Phonetic patterns and flanking character of the filled pauses reflect the influence of L1 on pausing in L2. (iii) Filled pauses occur more frequently at clause boundaries in L1 than in L2 utterances.

Thirty recordings were randomly selected for analysis from HunEng-D, ten from each proficiency level. Occurrences of the filled pauses in terms of their position in clauses, their phonetic forms, durations, and the first two formant frequencies were defined. To test statistical significance, cross-tabulation and GLMM tests were used (SPSS 19.0).

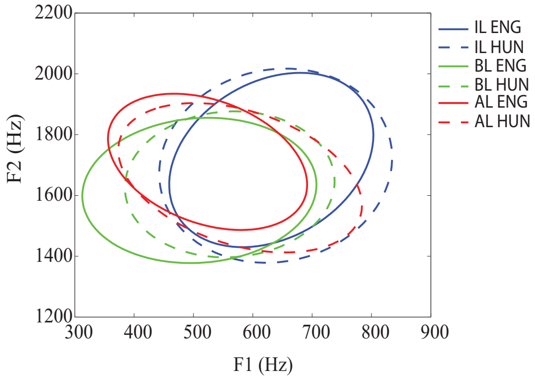
Results show that, as expected, the great majority of the filled pauses (72%) were neutral vowels in both L1 and L2 (out of 4612 filled pauses). Analysis was focused on these pauses. Language proficiency had a significant effect on the occurrence of filled pauses (Chi-square = 89.76; p<0.05). More filled pauses were found in L2 (2411 instances, 2.5/min) than in L1 (928 instances, 1.2/min) irrespective of level of proficiency. In L1, 29.4%, in L2 21% of all filled pauses were flanked by two silent pauses. They occurred most frequently with IL learners.

Durations of the filled pauses were analysed in terms of four fixed factors: L1 vs. L2, level of proficiency, position of the filled pause in the clause, and whether it was flanked by silent pauses. Statistical analysis revealed significant differences across positions (F(1, 3113) = 228,626; p=0.001) and languages (F(1, 3113) = 41,312; p=0.001). However, no significant differences were found in filled pause durations depending on language proficiency with the exception of IL vs. AL. The durational differences among speakers depending on level of proficiency were 10 to 26 ms, on average. The mean duration of the filled pauses was longer in L1 than in L2 (329 ms and 366 ms, respectively). The shortest durations occurred with AL and BL, while the longest ones were found with IL both in L1 and L2.

Pairwise contrasts confirmed significant durational differences depending on the positions of the filled pauses. Filled pauses were shorter within clauses than at clause boundaries both in L1 and in L2, particularly in the case of the BL and IL groups. In addition, filled pauses flanked by two silent pauses were significantly longer (397 ms, on average) than those surrounded lexically (304 ms, on average).

Formant frequencies of the filled pauses showed significant differences depending on level of proficiency (for the first formant: F(2, 3113) = 104.330; p=0.001 and for the second formant: F(2, 3113) = 53.161; p=0.001). F2s significantly differed also depending on the language (F(1, 3113) = 4.048; p=0.044), see Fig. 1. The formant differences suggest that speakers produce the same neutral vowels as filled pauses but with different vowel qualities.

Figure 1: Formant frequencies of filled pauses across languages and groups.



Conclusions will be drawn on the usage of filled pauses in terms of L1 vs. L2, language proficiency and phonetic patterns of the filled pause articulation.

4. REFERENCES

[1] Clark, H. H., Fox Tree, J. E. 2002. Using *uh* and *um* in spontaneous speaking. *Cognition* 84, 73–111.

[2] Gilquin, G., De Cock, S. & Granger, S. (eds.) 2010. *LINDSEI Louvain International Database of Spoken English Interlanguage*. Presses Universitaires de Louvain.

[3] Trouvain, J., Fauth, C. 2014. Zu phonetischen Details von Pausen – Untersuchungen von Lesesprache in L1 vs. L2. *Phonetik und Phonologie 10*. Konstanz.

[4] Urizar, X., Samuel, A. 2014. A corpus-based study of fillers among native Basque Speakers and the role of *zera*. *Language and Speech* 57, 338–366.

[5] Veiga, A., Candeias, S., Lopes, C., Perdigão, F. 2011. Characterization of hesitations using acoustic models. *Proc. of ICPhS*. Hong Kong, 2054–2057.

[6] Watanabe, M., Hirose, K., Den, Y., Minematsu, N. 2008. Filled pauses as cues to the complexity of upcoming phrases for native and non-native listeners. *Speech Communication* 50, 81–94.