

The COREIL corpus : a learner corpus designed for studying phrasal phonology and intonation

Delais-Roussarie, Elisabeth and Hiyon Yoo

Université Paris-Diderot

elisabeth.roussarie@wanadoo.fr, yoo@linguist.jussieu.fr

Any study on L2 intonation and phrasal phonology is interesting not only to get more knowledge on L2 language acquisition, but also to get better insights on the phonology of the target language itself. For many intonational and phrasal phenomena, clear descriptions are still missing. Studying learner speech may give interesting perspectives to the analysis of phenomena that have remained unnoticed up to now (e.g. grammatical and prosodic constraints that appear in case of self-repairs, phonological status of some events, etc.). The examination of well-built data is needed to work in such direction. The aim of the contribution is to present the COREIL corpus, an electronic oral learner corpus that has been designed to study the acquisition of phrasal phonology and intonation in French and English as a foreign language. Focus will be mostly given on the data collection protocol and the annotation scheme, but few research questions that can be worked on from this corpus will be thoroughly described.

The data have been gathered and annotated in such a way as to carry research on the acquisition of phrasal and tonal phenomena in second/ foreign language. Up to now, research on the L2 acquisition of phonology have mostly focused on segmental phenomena, but very few studies have been done on the acquisition of suprasegmental phenomena (see Archibald 1997, Trouvain and Gut 2007, Rasier and Hiligsmann 2007 among others). In that domain, many questions need to be investigated:

- is there a huge difference in the way phrasal and intonational phonology are acquired in L2, in comparison to what happened in L1 ?
- does transfer play a crucial role in the acquisition of suprasegmental phenomena ?
- are the answers given to the previous questions valid for any domain of intonation and phrasal phonology (stress patterns, phrasing, tonal phenomena, tonal inventory, etc.) ?
- is there a difference in the way segmental and suprasegmental phenomena are acquired?
- for prosodic phenomena such as rhythm, which encompass phonological and phonetic knowledge, is there a fixed order in the way the acquisition occurs (phonology before phonetics, status of phonetic implementation, etc.) ?

Moreover, the data collection protocol has been developed in order to allow several comparisons along different dimensions: i) influence of the learner L1, ii) proficiency/accuracy level in the second/ foreign language, iii) differences in the acquisition of prosody in L1 and L2, etc. The protocol has been thought in such a way as to be adaptative in order to take into account the age, the L1 and the level of the speaker, the experimental procedure (cohort study or not), etc.: for each language skill chosen for data collection (text reading, narrative or story telling, role plays, picture description, interview, etc.), a large panel of tasks have been elaborated in order to be picked up according to the recording situation at hand (age and level of the speaker, procedure, etc.).

References:

- Archibald, J. (1997). The acquisition of second language phrasal stress: a pilot study. In S.J. Hannahs and M. Young-Scholten (eds) *Focus on phonological acquisition*. Amsterdam/ Philadelphia : John Benjamins, pp. 263-289
- Trouvain, J and Gut, U. (2007). *Non-native prosody*. Berlin : Mouton de Gruyter
- Rasier, L. and Hiligsmann, P. (2007). *Prosodic transfer from L1 to L2. Theoretical and methodological issues*. Nouveaux cahiers de linguistique française 28/2007. Université de Geneve