Published in: LabPhon 14 The 14th Conference on Laboratory Phonology: Association for Laboratory Phonology, July 25-27, NINJAL, Tokyo. 2014.

Verbal feedback: positioning and acoustics of French "ouais" and "oui"

Laurent Prévot, Jan Gorisch

Aix Marseille Université & CNRS, Laboratoire Parole et Langage, France

laurent.prevot@lpl-aix.fr jan.gorisch@lpl-aix.fr

Embedded in a larger study that examines the form-function relationship for French positive feedback markers this study examines the form of *'oui'* and *'ouais'* with respect to their relative placement to the own speaker's and the interlocutor's speech production, and with respect to acoustic features including F0, aperiodicity and intensity.

Previous corpus-based research on these items has shown the importance of positional features for determining the choice of a specific item (Gravano et al., 2012). Neiberg et al., (2013) has shown the importance of phonological operations in Swedish (similar to our *'oui'* vs. *'ouais'* opposition) for communicating different attitudes. Other studies took a qualitative approach towards the phonetics of response tokens in naturally occurring talk in interaction. Gardner (2001) shows that the response token *'mm'* in English has specific acoustic features (fall-rising, falling or rise-falling pitch) when it is used in specific sequences of talk (continuation, acknowledgement, assessment). Ogden (2010) found that specific phonetic-prosodic properties (loudness, F0 height, F0 span and voice quality) of complaints have specific interactional consequences. In French, *'oui'* vs. *'ouais'* is assumed to be opposed in terms of level of language ('*ouais'* being more colloquial). However, their simultaneous presence in most interactive situations as well as some recent descriptive work (Péroz, 2009) tend to suggest different communicative functions for these two items. Here, we apply basic machine-learning techniques on a large dataset of these feedback items. The driving research question is how the various positional and acoustic properties interact and produce distinct properties of feedback items. The hypothesis is that independent from the lexical content (*'ouais'* vs. *'oui'*), the selected features can be used to train a classifier that distinguishes between different speech-exchange situations, e.g. conversational vs. task-oriented dialogue.

The material used for this study contains two corpora. The CID corpus (Bertrand et al., 2008) are face-to-face conversations that have the mere instruction to talk about "particular" events. The Aix Maptask Corpus (Gorisch et al., 2014), which is the French version of the famous Edinburgh corpus. Two conditions are recorded: face-to-face vs. remote interaction.

Since manual coding for such short items has been found to be difficult (D'Imperio et al., 2013) for the time being, we used a completely automatic and data-driven approach using only the sound files and the manual transcription that has been force aligned using SPPAS software (Bigi, 2012). Our final dataset was built from the 12 hours (CID: 8 hours, Maptask: 4 hours) of a transcribed and force-aligned corpus. We extracted from this data sets all the isolated *'oui'* (N: 352) and *'ouais'* (N: 1359) (isolated = item surrounded by over 200 ms pauses). We then extracted a set of features concerning their environment (including normalized duration of overlap, normalized position in the dialogue, position with regard to the other speaker's speech) and more acoustic features (duration, intensity median, F0 slope, F0 standard deviation, aperiodicity, …). We were then in the position of classification exploratory experiments. We used Weka (Hall et al., 2009) to try to predict meta-data information (*'oui'* vs. *'ouais',* corpus, speaker role) based on our feature set. About *'oui'* vs. *'ouais'* classification results, Naïve Bayes and Simple Logistic classifiers achieve about 79% accuracy without having speaker. This is about the score of majority-class baseline. When speaker information is added, this score rises to 88%. However, inspection of the decision-tree (J48) version of the classification shows that the increase in accuracy is partly due to a majority-class choice per speaker. Other features used by the classifiers include overlap information, duration and F0 standard deviation. The results on speaker role prediction (Co-narrator for the CID, Giver and Follower for the MapTasks) are stronger with an accuracy of 79.5% (without using meta-data) using first F0 slope, then duration, overlap features and intensity. Finally, corpus prediction, gave 71.5 % accuracy using mostly duration and then overlap and slope features. The current work consists of refining and improving the features. We are also finishing to transcribe a third corpus of 2 hours argumentation and negociation.

References

Bigi, B. (2012). SPPAS: a tool for the phonetic segmentations of speech, *8th LREC*, Istanbul. / D'Imperio, M., Petrone, C., Prévot, L. (2013). Basic prosodic transcription of short French feedback utterances. In *Workshop Advancing Prosodic Transcription for Spoken Language Science and Technology II*. / Bertrand, R., Blache, P., Espesser, R., Ferré, G., Meunier, C., Priego-Valverde, B., Rauzy, S. (2008). CID — Corpus of Interactional Data — Annotation et exploitation multimodale de parole conversationnelle. *Traitement Automatique des Langues 49(3)*, 1-30. / Gardner, R. (2001). *When listeners talk: Response tokens and listener stance*. John Benjamins. / Gorisch, J., Astésano, C., Gurman Bard, E., Bigi, B., Prévot L. (2014). Aix Map Task corpus : The French multimodal corpus of task-oriented dialogue, *9th LREC 2014*, Reykjavik. / Gravano, A., Hirschberg, J., Benus, S. (2012). Affirmative cue words in task-oriented dialogue. *Computational Linguistics 38(1)*, 1-39. / Hall, M., Eibe, F., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I. (2009). The WEKA Data Mining Software: An Update; *SIGKDD Explorations, 11(1).* / Neiberg, D., Salvi, G., Gustafson, J. (2013). Semi-supervised methods for exploring the acoustics of simple productive feedback. *Speech Comm. 55*, 451-469. / Ogden, R. (2010). Prosodic constructions in making complaints. In D. Barth-Weingarten, E. Reber, & M. Selting, *Prosody in Interaction* (pp. 81-103). John Benjamins. / Péroz, P. (2009). « « On ne dit pas ouais ! » », *Langue française* 1/2009 (n° 161), 115-134.