



Cultural Analysis of Spoken Linguistic Signalling: A Pipeline for the Alignment of Audio, Text, and Prosodic Features

Taylor Arnold, Nicolas Ballier

► To cite this version:

Taylor Arnold, Nicolas Ballier. Cultural Analysis of Spoken Linguistic Signalling: A Pipeline for the Alignment of Audio, Text, and Prosodic Features. DH2019, University of Utrecht, Jul 2019, Utrecht (Netherlands), Netherlands. <https://doi.org/10.34894/VDEFPR>. hal-04012540

HAL Id: hal-04012540

<https://hal.science/hal-04012540>

Submitted on 2 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

Cultural Analysis of Spoken Linguistic Signalling: A Pipeline for the Alignment of Audio, Text, and Prosodic Features

*Taylor Baillie Arnold (tarnold2@richmond.edu), University of Richmond, United States of America and
Nicolas Ballier (nicolas.ballier@univ-paris-diderot.fr), Université Paris Diderot*

Linguistic elements are known to be powerful signals for social categories such as class, race, education, political affiliation, and gender (Lakoff 1973; Lanehart 2015; Zappavigna 2012). Significant research has been conducted within the field of digital humanities to explore the ways in which language function to form communities across large corpora (Gavin 2018; Hoyt 2018; Orlikowski 2018). The vast majority of work on linguistic signalling in the digital humanities has focused on the analysis of print culture due to the availability of large textual datasets and readably available methods. Spoken language, however, is known to vary considerably within communities, even when they share a common written language and dialect (Cutler, 1997). Phonetic features such as tone, rhythm, and phoneme variation all serve to signal social identity. Methods for collecting and studying such variation offer, therefore, important insights into linguistic signaling that fail to be recognized by the study of text-only corpora.

In this poster, we present a general pipeline for the construction, alignment, and analysis of spoken linguistic data. Our pipeline uses a combination of open-source tools in the R programming language and will be made available as an open-source toolkit through GitHub. The goal in our alignment workflow is to produce a single table collectively representing each of the elements collected in the multiple annotations. As the smallest unit of analysis, we chosen to align the corpus at the phoneme level. Other larger linguistic units—such as syllables and words—and metadata are simply duplicated across the relative phonemes (see Figure 1). Unique identifiers for each unit are also included (these are not show in the figure only due to space), allowing for reconstruction of the original annotations. Once the data was collected as a single table, we were able to compute new lemmatised word forms, part of speech tags, dependency relationships, and named entities.

As a way of illustrating how this linguistic data pipeline is able to produce new scholarship, the poster focuses on an application to a corpus of spoken British English curated by the French-led Aix-MARSEC project (Auran, Bouzon, Hirst, 2004). The dataset provides features for analysing vowels, pitch, rhythm, phonosyntax, for prediction of phrase breaks for text-to-speech systems. It has even been used as a baseline for psycholinguistic experiments. In our analysis, we suggest that we can contribute to a finer-grained analysis of cultural and situational factors on the prosodic hierarchy by taking into account the original annotations of the corpus and adding new layers. We synthesize the earlier stages of the corpus, from the Spoken English Corpus to the Aix-MARSEC speech database.

Our poster lays out two experiments: the analysis of major and minor boundaries acknowledged in the corpus on the basis of a multidimensional analysis of the different subgenres of the corpus and of its prosodic and syntactic annotation when analysing the final nuclear contours of the prosodic units. Results of the distribution of the main intonation values (Major versus Minor) across the final tonal segment types in final positions according to discourse genres are shown in Figure 2. As, Brierly and Atwell explain, prosodic parsing can be based on the speaker's desire to highlight specific aspects of the syntax producing a break after the item she wishes to highlight as in '...The idea that it's important | for developing countries to become self-sufficient | in food | is widely | and uncritically accepted | not just in Brussels; | but from the orthodox economic standpoint | it's without foundation...' whereas the syntactic model would predict a break after idea and before to (2004). 'Highlighting' as a strategy means emphasizing the role of adjectives in final position of the intonation unit. The relative proportion of adjectives in final position of minor units should be monitored in relation to this 'highlighting' strategy. The dominance of Higher (H) pitch targets (see Left) for minor units confirms our previous observation as does the clustering of Same and Bottom for major units (consistent with finality). Our 3-gram analysis of final pitch targets in intonation units reveal phonosyntactic patterns. Considering the number of S-S-S sequences, a 4-gram analysis might be more relevant for the definition of the span of the pitch targets that characterize these units. Figure 3 illustrates a clear gender difference in tonetic stress marks and intonation unit. An overall patterning of major units with tonetic stress markings suggesting finality (falls) whereas rises co-occur with minor units, marking continuity.

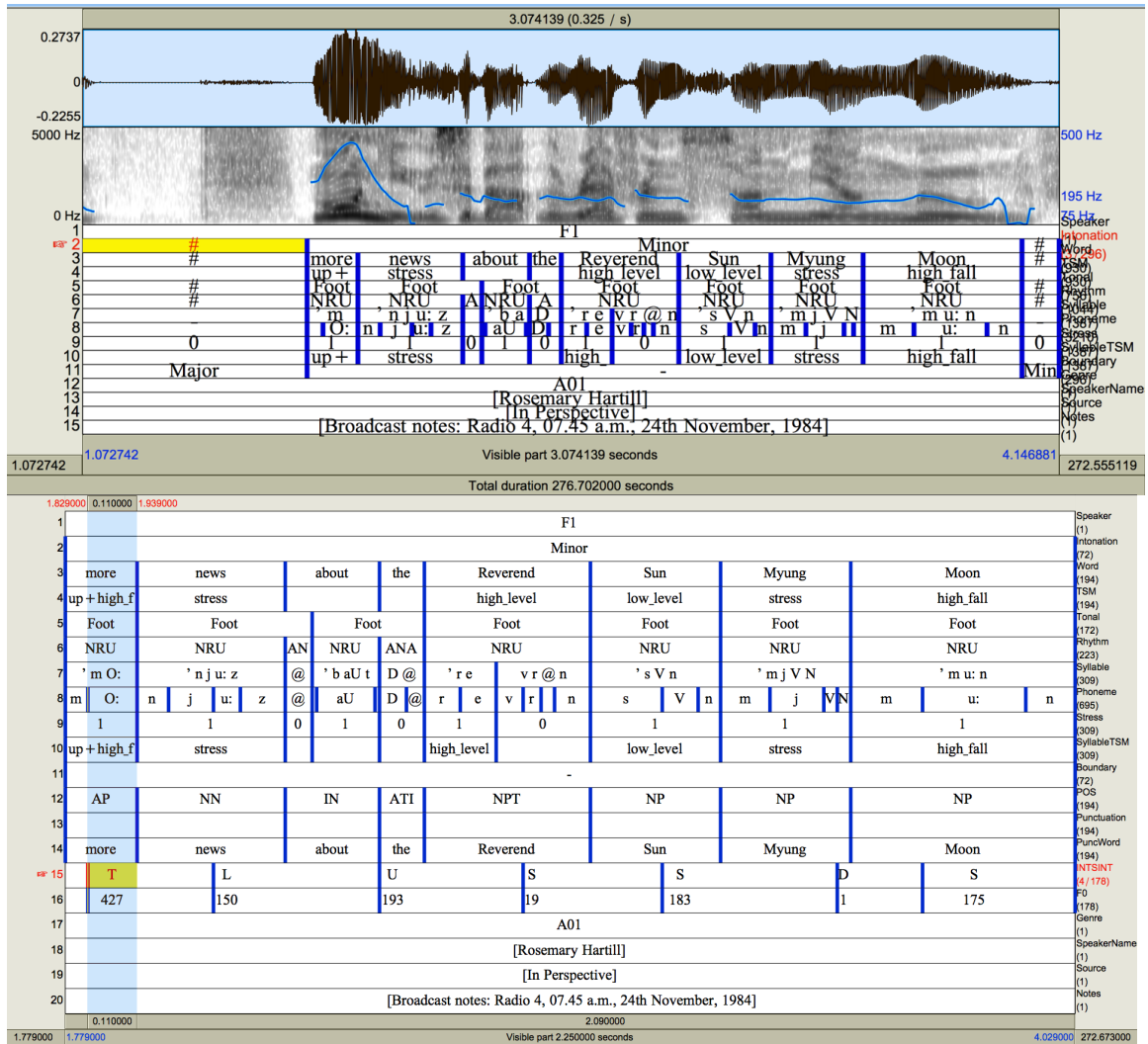


Figure 1. Example of the input data (top) and aligned corpus (bottom) using our alignment pipeline.

[Page]

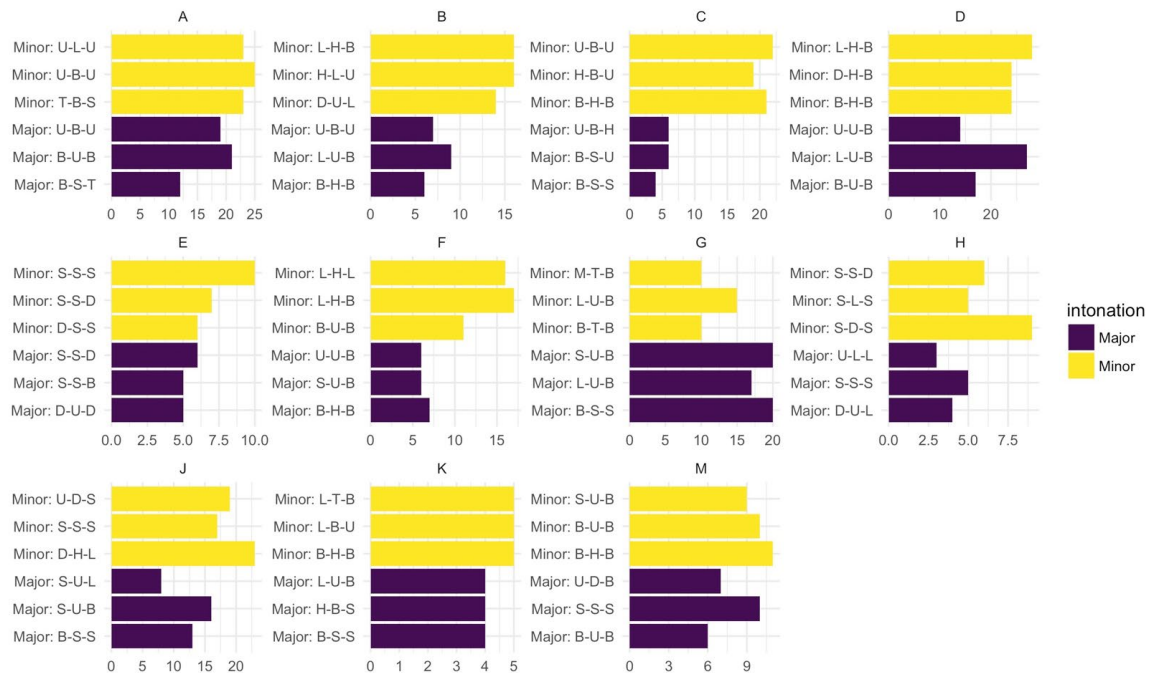


Figure 2. Distribution of the main intonation values in final positions according to discourse genres.

Figure 3. Distribution of tonetic stress marks by gender and intonation unit.

Appendix A

Bibliography

1. Auran, Cyril, Caroline Bouzon, and Daniel Hirst. "The AixMARSEC project: an evolutive database of spoken English." In In Bel, B. & Marlien, I.(eds) *Proceedings of the Second International Conference on Speech Prosody*. 2004.
 2. Beliao, Julie. "Characterizing speech genres through the relation between prosody and macrosyntax." In *Student Sessions at the European Summer School in Logic, Language and Information*, pp. 1-18. Springer, Berlin, Heidelberg, 2013.
 3. Brierley, Claire, and Eric Atwell. "Prosodic phrase break prediction: problems in the evaluation of models against a gold standard." *TAL Journal: Traitement Automatique des Langues* 48, no. 1 (2007): 187-206.
 4. Cutler, Anne, Delphine Dahan, and Wilma Van Donselaar. "Prosody in the comprehension of spoken language: A literature review." *Language and speech* 40, no. 2 (1997): 141-201.
 5. Degaetano-Ortlieb, Stefania and Elke Teich. *Using relative entropy for detection and analysis of periods of diachronic linguistic change*. Santa Fe, New Mexico: Association for Computational Linguistics, 2018, pp. 22-33.
 6. Gavin, Michael and Eric Gidal. "Scotland's Poetics of Space: An Experiment in Geospatial Semantics". In: *Cultural Analytics* (2018).
 7. Hoyt Long, Anatoly Detwyler, and Yuancheng Zhu. "Self-Repetition and East Asian Literary Modernity, 1900-1930". In: *Cultural Analytics* (2018).
 8. Lakoff, Robin. "Language and Woman's Place." *Language in society* 2, no. 1 (1973): 45-79.
 9. Lanehart, Sonja, ed. *The Oxford Handbook of African American Language*. Oxford University Press, 2015
 10. Orlikowski, Matthias, Matthias Hartung, and Philipp Cimiano. "Learning diachronic analogies to analyze concept change." In *Proceedings of the Second Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*, pp. 1-11. 2018.
 11. Zappavigna, Michele. *Discourse of Twitter and Social Media*. Bloomsbury, 2012.
-