INTERSPEECH 2023
20-24 August 2023, Dublin, Ireland

Speech reduction: position within French prosodic structure

Kübra Bodur[1], *Roxane Bertrand*[1], *James S. German*[1], *Stéphane Rauzy*[1], *Corinne Fredouille*[2], *Christine Meunier*[1]

[1]Aix-Marseille Université, LPL, CNRS, France
[2]Avignon Université, LIA, CNRS, France

kubra.bodur@univ-amu.fr

Abstract

Variation in the speech signal is a characteristic of spoken language, emerging partially as a result of interactions between various linguistic levels. One example of variation is phonetic reduction, where words are produced with missing or underspecified phonetic forms. Using a French conversational corpus, this paper focuses on the relationship between reduction and prosodic structure to see whether certain positions favor the occurrence of reduction. We annotated and observed the distribution of reduced sequences within specific prosodic domains (Intonational and Accentual Phrases). Preliminary analyses revealed that the detected reductions occur mostly mid-IP and very rarely at IP-final. However, this pattern may vary among speakers, as speakers have different patterns in terms of the number of reductions produced and their positions. It is also usually the case that the reduced sequences occurring mid-IP, coincide with the AP level boundaries, extending from one AP to another.

Index Terms: speech reduction, prosody, spontaneous speech

1. Introduction

Prosody can be defined as the supra-segmental organization of speech, encoding prominence, accent placement, phrasing, and speaking style. The way in which sounds are organized in speech, known as prosodic structure, follows a hierarchical order where larger units have more influence than smaller ones. The boundaries within prosodic structure are marked by changes in pitch, as well as variation in the acoustic/phonetic realizations of individual segments. At prosodic boundaries, the durations of segments might increase, gestures might be lengthened and this lengthening might increase as the prosodic boundary is hierarchically more important in the prosodic organization (e.g. [1, 2, 3, 4, 5]). The prosodic organization and its relationship with acoustic/phonetic realizations are crucial aspects to consider when exploring the role of prosodic structure in speech variation, even though these manifestations are specific to languages and their consonant categories. The influence of prosodic prominence might extend from individual segments to larger units such as syllables and words, contributing to a variation in speech signal.

Variation in the speech signal is an intrinsic feature of spoken language [6] as a result of the permanent adaptation of speech productions to the speaking context [7]. One way this variation might manifest itself in spontaneous conversations is through the hyper and hypo-articulation [7] of speech segments. In the case of the hypo-articulation, speech is not as clearly articulated as in formal or read speech [8, 9]. Segments (i.e. phonemes, syllables, words) may appear acoustically and phonetically weakened with shorter durations, reduced vowel spaces and/or they might be completely missing in extreme cases of reduction [10, 11, 12, 13] all the while allowing the interaction to flow smoothly between speakers and listeners. Even though we know little about all the driving forces behind the reduction processes, previous research has suggested that reduction in speech is influenced by various linguistic and conversational factors such as semantic and lexical predictability [14, 15, 16], lexical frequency [17, 18], previous mention in discourse [19, 20], speaking style [9] as well as the phonological properties of the words composing reduced sequences [21, 22, 23]. It has been shown that prosodic structure is also one of the factors contributing to the level of reduction in speech [3, 19, 24, 25]. Aylett & Turk [3] claimed that reduction in speech occurs in order to preserve the smoothness/fluency of speech, suggesting that contextually redundant elements should be reduced through prosodic and phonetic means. They found that speakers tend to reduce the duration and amplitude of unstressed syllables within a word, but maintain greater precision in stressed syllables. Bell et al.'s [26] findings support this claim, as they found that vowels in unstressed syllables were more likely to be reduced than those in stressed syllables, and that the degree of reduction was influenced by the position of the syllable within the prosodic structure of the word. The same was proven for consonants in English [27] and both for vowels and consonants in Korean [28].

A Model of French Prosodic Structure

This study is based on the phonological model of French prosodic structure developed by Jun and Fougeron [29, 30]. In French, there are two main prosodic boundaries, the *Accentual phrase* (AP) [29] and the *Intonational phrase* (IP), which is a larger hierarchical structure that includes the AP. The AP is characterized by an initial boundary tone (LHi), a nuclear pitch accent, and a final tone (LH), and its location within the larger IP can influence the realization of the pitch accent and the overall f0 contour. The IP can consist of one or more APs and is delimited by a boundary tone that marks the end of the IP (L% or H%) and the beginning of a new one. The boundary between two IPs is marked by a larger change in pitch and duration than between accentual phrases. Additionally, research suggests that there is a third boundary level in French prosodic structure, the *intermediate phrase* (ip) [31, 32] which is larger than the AP but smaller than the IP and is characterized by a downstep (L-) and a final tone (either L- or H-). The downstep concerns the height of the peak of the accent of a non-initial AP, only if it is also non-final.

Our study aims to investigate the distribution of speech reductions within the prosodic structure of French and determine whether reductions are more likely to occur in certain positions.

10.21437/Interspeech.2023-1651

Building on the work of Aylett and Turk [3], we make the assumption that reductions would be more prevalent in prosodically weaker positions, affecting unstressed syllables located in the middle of a prosodic domain for instance.

This study presents a novel approach to investigating reductions in spoken French by analyzing spontaneous conversations produced in a more natural context, in contrast to controlled and read laboratory speech. Additionally, this research focuses specifically on the French language, which has received comparatively less attention than English in terms of reduction studies. Moreover, we introduce a new method for detecting reductions in spontaneous conversations, enabling us to analyze reduced sequences involving multiple words (for more details, see Section 2). By combining these approaches, this study provides a descriptive examination of the relationship between prosodic structure and reductions in a set of conversations in French.

2. Methods

2.1. Corpus

For the study of reduction in spontaneous speech and its relation to the prosodic categories, we selected the *Corpus of Interactional Data* [33] as it includes phonetic and prosodic annotations performed both manually and automatically using tools that can process large amounts of speech data and provide relatively reliable outputs. The corpus consists of 8 audio-visual recordings (each lasting 1 hour) of conversational dyads between colleagues (10 females and 6 males, (M = 34 years old)) with relatively familiar relationships, which is thought to be a context enhancing the spontaneity of conversations.

2.2. Annotations

2.2.1. Detecting reductions

In order to study phonemic and temporal reduction in speech, we obtained reduced sequences by adopting a bottom-up approach. We used a script that detects sequences containing at least 6 phonemes in a 230ms window and continues to search for reductions through a sliding window. Taking individual alignment files containing phoneme sequences and their temporal boundaries as input, it counts from a phoneme's start time how many phonemes belong to the predefined window (230ms). If the threshold of the number of phonemes (6) is reached, the sequence is saved in a Textgrid and the script goes on to search for the following 230ms window. The script continues adding phonemes to the sequence as long as the threshold (6 phonemes) is exceeded. The 230ms window was selected after testing the script with other window sizes such as 180ms and 200ms, which pointed at more highly reduced (in terms of duration) areas in the corpus. The choice of 230ms is thus crucial, in that a larger window might contain more phonemes and words allowing us to obtain a sufficient number of reductions for a more reliable preliminary analysis.

2.2.2. Prosodic annotations

The prosodic annotation of the corpus was performed by semi-naive annotators following a perception based approach [34]. The annotation procedure follows Tones and Break Indices (ToBI) guidelines [35] in general by adapting the guidelines to the prosodic phrasing of French language following Jun and Fougeron's model [36]. The annotators labeled the ends of words from 0 to 3 according to the strength of their associations with the following word. In this annotation scheme, label 1 corresponds to the AP, 2 to the ip and 3 to the IP. However, this research is based on two main prosodic domains in French [30] since there is not enough evidence on the manifestation of the intermediate phrase [32] in conversational speech yet. Labels 2 and 3 were merged in the end, due to the lack of evidence for the presence of ip in spontaneous speech. Therefore, the coding scheme includes only AP and IP (see Figure 1). Additionally, the coding of numbers in the prosodic annotations ”IP11” and ”AP22” seen in the example are related to the inter-annotator agreement scores. For instance, when an accentual phrase is annotated as ”AP02”, 0 means no gold (no expert annotation) annotation while 2 means 2 naive annotators agreed upon the annotation. In the case of Figure 1, ”AP22” means that all four annotators (2 expert and 2 naive) agreed on the annotation.

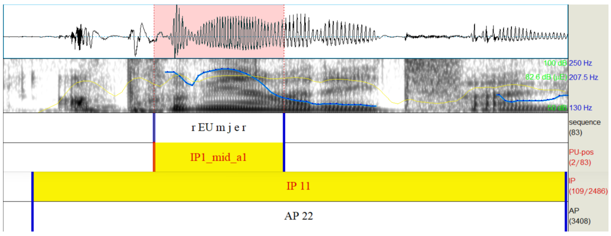

Figure 1: *An example of the annotation of reductions within prosodic domains. The first tier includes the reductions, the second shows how many APs form the IP (IP2), the position in the IP (mid), the AP containing the reduction (a2) and whether the reduction is on the frontier of a prosodic domain (Fa).*

2.2.3. Prosodic annotations and reductions

In order to analyze where in the prosodic domains the reductions occur, a second set of annotations based on the prosodic phrasing explained above was performed by the first author. The reduced sequences which had been detected automatically as well as the prosodic annotations of the corpus were brought together using Praat [37]. The annotation structure is as follows:

IP*n* (*n* = the number of APs making up the IP) - **position** in the IP (sta/mid/end/all)- **a*n***(n = the number of the AP within its IP).

Two complementary categories were also included in the annotation scheme: Fa and Fi. *Fa* is used when the reduced sequence crosses an AP boundary, meaning the sequence is stretching between two or more APs within the same IP, whereas *Fi* indicates that the reduced sequence crosses an IP boundary, extending to the previous or following IP (see Figure 1 for an illustration).

Disfluencies While annotating the positions of reductions within prosodic units, it was observed that there were many cases where the IPS contained disfluencies (repetitions, filled pauses, false starts) -which is a characteristic of spontaneous speech- as well as reductions. As disfluencies occurred quite frequently in the corpus we used (as frequent as 1 disfluency in every 15 words [38]), and as they interact with the prosodic structure [39, 40] thus making prosodic segmenting more difficult, we decided to exclude the annotations that coincide with disfluencies even if the distribution of reduction in disfluencies may be relevant. The disfluency annotations were already available for the corpus. This procedure resulted in the exclusion of 351 annotations (36,9 % of the reduced sequences). However, the excluded data will be useful for further investigation of reduction processes.

3. Results

Having extracted the annotations, we used a custom-made script to exclude reductions coinciding with disfluencies (351 cases out of 950 detected reductions) and a case for which there was no prosodic annotation available. We ended up with 598 annotations for prosodic units and counted the occurrences of reductions at various positions (initial, middle, end, all) with respect to the IPs.

3.1. Distribution of reduced sequences within IPs

We first observed the overall distribution of reduced sequences at various positions within the IP. The results are shown in Table 1. Based on our annotations, reduced speech was produced mostly at IP-mid position (56%). It was also quite frequent at IP-initial (32%) while it was rare at IP-final (9%) or over all of the IP (3%).

Position	Count	Ratio
start	193	32%
mid	336	56%
end	53	9%
all	16	3%
	598	

Table 1: *The various positions where reduction occurred and their distributions within the IP.*

3.2. Distribution of positions among speakers

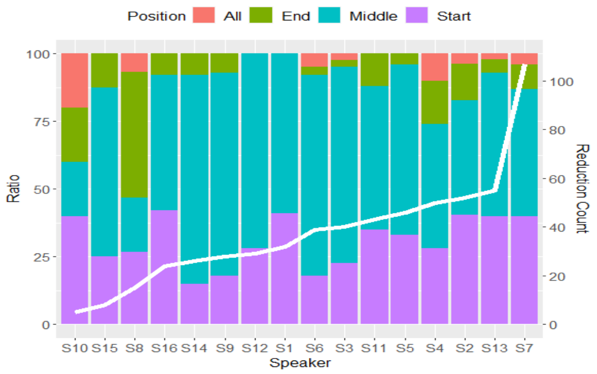

Figure 2: *The distribution of reductions in various positions within the IP per speaker. The bars show the ratios (%) of different positions where reductions occurred. The line shows the number of reductions produced by each speaker of the corpus.*

The distributions of above-mentioned positions for the annotated reduced sequences were observed separately for each speaker of the corpus. The general pattern is compatible with the overall results presented above. The reduced sequences occur mostly in the middle of an IP (see Figure 2 below). However, some speakers displayed different patterns. For example, the reductions produced by S10 tend to be found more frequently at IP-initial (40% of his reductions) than the other positions. Similarly, for S8, the reductions were more frequently at IP-final (46,68%) followed by IP-initial (26,66%) while they appear only 20% of the time in the middle. Figure 2 shows that there is no relation between the number of speakers' reductions and the specific distribution of their reductions.

3.3. Reductions on the boundaries

While annotating reduced sequences within prosodic domains, it has been observed that numerous sequences occurred on prosodic boundaries. It was usually the case that a sequence extended from the end of one IP to the start of another or from one AP to another in the middle of an IP (see Figure 3 for an example). Out of a total of 598 reductions, 116 (19,3%) appeared on AP boundaries, while 135 (22,5%) appeared on IP boundaries extending from one utterance to another.

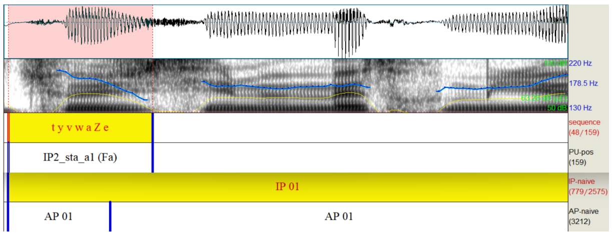

Figure 3: *An example of a reduced sequence occurring on two APs, the sequence is "**tu vois/j'ai** essayé de me mettre au milieu" ("you see/i tried to put myself in the middle").*

3.4. The length of the IPs containing reduced sequences

The total number of IPs found in the corpus being 31,119, only 2% of these were affected by reductions detected using the 230ms/6 phonemes parameter (after excluding the disfluent IPs). The annotation scheme allowed us to obtain the length of the IPs containing reductions, and the length of the IPs is calculated based on the number of APs found in them. The results suggest that, in the 598 detected sequences, reductions appeared more in shorter IPs (in terms of the APs they contain) than the IPs composed of 6 or 9 APs (see Table 2) which is rare but possible in the French prosodic structure.

IP length	Count	IP length	Count
IP1	309	IP5	6
IP2	195	IP6	2
IP3	59	IP9	1
IP4	26	**Total**	598

Table 2: *A summary of the IP length (based on the number of APs involved) containing reduced sequences.*

4. Discussion

The current paper investigates the relationship between reduction in spontaneous speech and prosodic structure, more specifically whether reductions occur in certain positions within prosodic domains. Based on a phonological model of French prosodic structure [30], it aims to investigate the distribution of reduced sequences within this structure.

We adopted a new and automatized method for the detection of reduced sequences to discover possible reductions in a corpus of conversational speech in French [33]. The selected window size (230ms/6 phonemes) focuses mostly on reductions involving multiple words -especially monosyllabic words-; suggesting that "smaller" reductions, affecting only one or two phonemes within a word, have not been taken into consideration in this study. We tried to analyze a reduction pattern in conversational speech using a bottom-up method. This is part of the originality of our study, compared to previous ones concentrating on reductions at the word level through a top-down approach for detection of reductions (based on well known lexicalized reduction forms).

For descriptive purposes, we observed where reductions occurred more frequently, whether there was a global tendency for speakers to produce reductions in similar positions and whether the length of a prosodic phrase played a role in the occurrence of reductions in these positions. As a characteristic of spontaneous speech, some reductions co-occurred with interruptions in the speech flow, which are called "disfluencies". These were not included in the analysis of the data. However, if the reduced sequence was found in the *Reparans* which is the repairing part after the disfluency, we included that sequence in the analysis. The goal of this elimination process was to have a coherent preliminary analysis, since disfluencies are known to be able to cause ambiguity in separating prosodic units [41], as well as the segmentation of words making up the reduced sequences. This exclusion allowed us to have a first impression of how reductions behave in speech when they are not disrupted by conversational factors. However, we believe that the study of the reductions containing disfluencies will also prove very useful for understanding the reduction mechanisms in the future, as there might also be an interaction between disfluency and reduction in spontaneous speech.

Looking at the distributions of reduced sequences at various positions within the Intonational phrase (IP), the results suggest that reductions occurred mostly in the middle of the IPs (56% of the detected reductions) in the corpus. Compatible with our hypothesis that reductions would be more prevalent in mid-domain positions as these areas are susceptible to being prosodically weaker, thus less informative, this finding could be also considered compatible with previous research [3, 26] considering the French prosodic structure. It is important to note that, although reduced sequences occurred preferably in the middle of an IP, no position was excluded for them to occur. Another finding is that, even if the distribution pattern of reductions in prosodic structure is quite consistent among speakers, it still displays variability to a certain degree. Some speakers tend to produce more reductions at the beginning or end of the IP. Seeing that the variation is also very high among the speakers in terms of the numbers of reductions produced in conversations (e.g., S7 produced 107 reduced sequences while S10 had only 5 -after the exclusion of disfluent IPs), these differences might be explained by individual characteristics of each speaker. Despite the fact that speaker related variability has rarely been addressed in previous research (except for Labov's work on sociological factors [42, 43]), it contributes significantly to the heterogeneity in reduction rates; speakers behave differently regarding the number of reduced productions (see Figure 2 for an illustration). In a previous analysis, we found that articulation rate is one of the factors contributing to inter-speaker variation: a positive correlation was found between the articulation and reduction rates ($r(14) = 0.78$, $p = 0.0002996$). However, articulation rate itself is not enough for accounting for the differences between speakers in terms of reduction behavior, further investigation is required to see what might be causing these differences.

We also observed that some reductions can be realized at the junction between two APs. The AP ends with a pitch accent realized by a movement of f0 (most often rising) and by a lengthening of the final full syllable of the lexical word. The observed reduced sequences concern cases where function words are often found at the beginning of the following APs (e.g., je crois/que...; chaque fois /qu'il y avait; tu vois/que... ;elle avait/un...). This finding is in line with the results obtained in previous studies, (not only on lexical predictability but also on unstressed versus stressed syllables) since these reductions mainly concern function words found at the beginning of the following APs which can even be considered informationally redundant in many cases. For further analysis, it would be interesting to measure the durations of stressed syllables (from AP 1) which, although they are perceived as stressed since they are associated with an AP boundary, could nevertheless be reduced compared to other syllables in this position which are not followed by reduction.

Furthermore, when we looked at the lengths of the IPs containing reduced sequences, we saw that shorter IPs were more common than the longer ones among the IPs containing reduction (IP1 and IP2 make up 84,3% of the relevant IPs). On the other hand, the relationship between reduction and the length of an utterance can be very complex and may depend on a number of factors, including the context in which the utterance occurs, the speaker's fluency, and the degree of stress or emphasis placed on certain words or phrases. Considering that disfluencies might cause an IP to be longer due to pauses, repetitions, and restarts, it can be thought that this dominance of the shorter IPs might be the result of the excluded IPs containing disfluencies. However, when we compared the disfluency including and disfluency excluding datasets, the overall distribution did not change for this analysis. In this case, it is difficult to generalize and to predict what type of IPs are more susceptible to be affected by reductions, since we do not have the information regarding the lengths of all the IPs comprising the corpus.

Our research being limited in terms of the number of reductions we were able to detect using the selected parameters (230ms/6 phonemes), we believe more spontaneous speech data would provide more opportunities to understand the relationship between reduction and its place in the prosodic structure. The fact that the evidence for the manifestation of the intermediate phrase [32] is not adequate for conversational speech may have limited our interpretation of the results, however, with more evidence in the future, this level could also prove to be useful in further accounting for the interaction between reductions in French prosodic structure.

In conclusion, the preliminary results show that the distribution of reductions in various prosodic positions displays variability as regards to the speakers and the length of the utterances. Even though we aim at exploring the relationship between reduction and the prosodic structure, the prosody is not enough for accounting for the occurrence of reductions on its own. When enriched with more detailed acoustic and phonetic analyses, this preliminary research would serve as a valuable basis for further research to fully comprehend the mechanisms causing reduction during speech.

5. Acknowledgements

This project was partially supported by grants ANR-16-CONV-0002 (ILCB) and the Excellence Initiative of Aix-Marseille University (A*MIDEX).

6. References

[1] C. W. Wightman, S. Shattuck-Hufnagel, M. Ostendorf, and P. J. Price, "Segmental durations in the vicinity of prosodic phrase boundaries," *The Journal of the Acoustical Society of America*, vol. 91, no. 3, pp. 1707–1717, 1992.

[2] D. Byrd, J. Krivokapić, and S. Lee, "How far, how long: On the temporal scope of prosodic boundary effects," *The Journal of the Acoustical Society of America*, vol. 120, no. 3, pp. 1589–1599, 2006.

[3] M. Aylett and A. Turk, "The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech," *Language and speech*, vol. 47, no. 1, pp. 31–56, 2004.

[4] D. Byrd and S. Choi, "At the juncture of prosody, phonology, and phonetics—the interaction of phrasal and syllable structure in shaping the timing of consonant gestures," *Laboratory phonology*, vol. 10, pp. 31–59, 2010.

[5] J. Cole, "Prosody in context: A review," *Language, Cognition and Neuroscience*, vol. 30, no. 1-2, pp. 1–31, 2015.

[6] M. T. C. Ernestus, *Voice assimilation and segment reduction in casual Dutch, a corpus-based study of the phonology-phonetics interface*. Utrecht: LOT, 2000.

[7] B. Lindblom, "Explaining phonetic variation: A sketch of the h&h theory," in *Speech production and speech modelling*. Springer, 1990, pp. 403–439.

[8] M. Ernestus and N. Warner, "An introduction to reduced pronunciation variants," *Journal of Phonetics*, vol. 39, no. SI, pp. 253–260, 2011.

[9] C. Gendrot, M. Adda-Decker, and C. Schmid, "Comparaison de parole journalistique et de parole spontanée: analyses de séquences entre pauses," in *Proceedings of the Joint Conference JEP-TALN-RECITAL 2012*, 2012, pp. 649–656.

[10] B. Schuppler, M. Ernestus, O. Scharenborg, and L. Boves, "Acoustic reduction in conversational dutch: A quantitative analysis based on automatically generated segmental transcriptions," *Journal of Phonetics*, vol. 39, no. 1, pp. 96–109, 2011.

[11] O. Niebuhr and K. J. Kohler, "Perception of phonetic detail in the identification of highly reduced words," *Journal of Phonetics*, vol. 39, no. 3, pp. 319–329, 2011.

[12] M. Ernestus, "Acoustic reduction and the roles of abstractions and exemplars in speech processing," *Lingua*, vol. 142, pp. 27–41, 2014.

[13] C. G. Clopper, R. Turnbull, F. Cangemi, M. Clayards, O. Niebuhr, B. Schuppler, and M. Zellers, "Exploring variation in phonetic reduction: Linguistic, social, and cognitive factors," *Rethinking reduction*, pp. 25–72, 2018.

[14] M. Aylett and A. Turk, "Language redundancy predicts syllabic duration and the spectral characteristics of vocalic syllable nuclei," *The Journal of the Acoustical Society of America*, vol. 119, no. 5, pp. 3048–3058, 2006.

[15] C. G. Clopper and J. B. Pierrehumbert, "Effects of semantic predictability and regional dialect on vowel space reduction," *The Journal of the Acoustical Society of America*, vol. 124, no. 3, pp. 1682–1688, 2008.

[16] R. Turnbull, "The role of predictability in intonational variability," *Language and speech*, vol. 60, no. 1, pp. 123–153, 2017.

[17] M. Pluymaekers, M. Ernestus, and R. H. Baayen, "Lexical frequency and acoustic reduction in spoken dutch," *The Journal of the Acoustical Society of America*, vol. 118, no. 4, pp. 2561–2569, 2005.

[18] J. Bybee, R. J. File-Muriel, and R. N. De Souza, "Special reduction: A usage-based approach," *Language and Cognition*, vol. 8, no. 3, pp. 421–446, 2016.

[19] R. E. Baker and A. R. Bradlow, "Variability in word duration as a function of probability, speech style, and prosody," *Language and speech*, vol. 52, no. 4, pp. 391–413, 2009.

[20] J. M. Kahn and J. E. Arnold, "Articulatory and lexical repetition effects on durational reduction: Speaker experience vs. common ground," *Language, Cognition and Neuroscience*, vol. 30, no. 1-2, pp. 103–119, 2015.

[21] N. Pharao, "Consonant reduction in copenhagen danish," *Unpublished Ph. D. dissertation, University of Copenhagen, Denmark*, 2010.

[22] C. Meunier and B. Bigi, "Répartition des phonèmes réduits en parole conversationnelle. approche quantitative par extraction automatique," in *Actes de la conférence conjointe JEP-TALN-RECITAL 2016*, 2016, pp. 615–623.

[23] ——, "Variations temporelles des phonèmes en parole conversationnelle: propriétés phonétiques et facteurs lexicaux," *Studii de lingvistică*, 2021.

[24] C. é Fougeron and S.-A. Jun, "Rate effects on french intonation: Prosodic organization and phonetic realization," *Journal of Phonetics*, vol. 26, no. 1, pp. 45–69, 1998.

[25] R. S. Burdin and C. G. Clopper, "Phonetic reduction, vowel duration, and prosodic structure." in *ICPhS*, 2015.

[26] A. Bell, J. M. Brenier, M. Gregory, C. Girand, and D. Jurafsky, "Predictability effects on durations of content and function words in conversational english," *Journal of Memory and Language*, vol. 60, no. 1, pp. 92–111, 2009.

[27] L. C. Dilley and M. A. Pitt, "A study of regressive place assimilation in spontaneous speech and its implications for spoken word recognition," *The Journal of the Acoustical Society of America*, vol. 122, no. 4, pp. 2340–2353, 2007.

[28] T. Cho and P. A. Keating, "Articulatory and acoustic studies on domain-initial strengthening in korean," *Journal of phonetics*, vol. 29, no. 2, pp. 155–190, 2001.

[29] S.-A. Jun and C. Fougeron, "The accentual phrase and the prosodic structure of french," in *Proceedings of the 13th International Congress of Phonetic Sciences*, vol. 2, 1995, pp. 722–725.

[30] ——, "A phonological model of french intonation," *Intonation: Analysis, modelling and technology*, pp. 209–242, 2000.

[31] A. Michelas and M. d'Imperio, "Durational cues and prosodic phrasing in french: evidence for the intermediate phrase," in *Speech Prosody*, 2010, p. 4.

[32] A. Michelas, "Caractérisation phonétique et phonologique du syntagme intermédiaire en français: de la production à la perception." Ph.D. dissertation, Université de Provence-Aix-Marseille I, 2011.

[33] R. Bertrand, P. Blache, R. Espesser, G. Ferré, C. Meunier, B. Priego-Valverde, and S. Rauzy, "Le cid-corpus of interactional data-: protocoles, conventions, annotations," *Travaux interdisciplinaires du Laboratoire parole et langage d'Aix-en-Provence (TIPA)*, vol. 25, pp. 25–55, 2006.

[34] L. Prevot, R. Bertrand, K. Peshkov, S. Rauzy, and P. Blache, "Prosody, Discourse and Syntax in French Conversations: Resource creation and Evaluation," Feb. 2023, working paper or preprint. [Online]. Available: https://hal.science/hal-01231884

[35] M. E. Beckman and G. A. Elam, "Guidelines for tobi labelling, version 3," *Ohio State University*, 1997.

[36] S.-A. Jun and C. Fougeron, "Realizations of accentual phrase in french intonation," 2002.

[37] P. Boersma and D. Weenink, "Praat: Doing phonetics by computer (version 6.0. 14)," *Retrieved from (last access: 29.04. 2018)*, 2016.

[38] B. Pallaud, R. Bertrand, L. Prevot, P. Blache, and S. Rauzy, "Suspensive and disfluent self interruptions in french language interactions," 2019.

[39] E. Shriberg, R. Bates, and A. Stolcke, "A prosody only decision-tree model for disfluency detection," in *Fifth European Conference on Speech Communication and Technology*, 1997.

[40] J. Beliao and A. Lacheret, "Disfluency and discursive markers: when prosody and syntax plan discourse," in *DiSS 2013: The 6th Workshop on Disfluency in Spontaneous Speech*, vol. 54, no. 1, 2013, pp. 5–9.

[41] L. Prevot, R. Bertrand, and S. Rauzy, "Investigating disfluencies contribution to discourse-prosody mismatches in french conversations," in *The 10th Workshop on Disfluency in Spontaneous Speech*, 2021.

[42] W. Labov, "Sociolinguistic patterns (university of pennsylvania, philadelphia)," 1972.

[43] ——, "Principies oflinguistic change: Social factors," *Malden, Massachusetts: Blackwell*, 2001.