



HAL
open science

Tonal density characterises the scope of the overbid use of connector "mais" in French conversation

Cristel Portes, Marie Kolenberg, Stephane Rauzy, Roxane Bertrand

► To cite this version:

Cristel Portes, Marie Kolenberg, Stephane Rauzy, Roxane Bertrand. Tonal density characterises the scope of the overbid use of connector "mais" in French conversation. *Speech Prosody 2024*, Leiden University, Jul 2024, Leiden (NL), Netherlands. hal-04617715

HAL Id: hal-04617715

<https://hal.science/hal-04617715v1>

Submitted on 19 Jun 2024

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.

Tonal density characterises the scope of the overbid use of connector “mais” in French conversation

Cristel Portes¹, Marie Kolenberg², Stéphane Rauzy¹, Roxane Bertrand¹

¹Aix Marseille Univ, CNRS, LPL, Aix-en-Provence, France ; ²KU Leuven, Leuven, Belgium
christelle.portes@univ-amu.fr, marie.kolenberg@gmail.com, stephane.rauzy@univ-amu.fr, roxane.bertrand@univ-amu.fr

Abstract

Most uses of *mais* ('but') as a connector in French (i.e. correction, opposition, concession) imply a contrast between their antecedent p and their scope q but differ in terms of argumentative orientation. Conversely, p and q in overbid *mais* ('des grands mecs *mais* très grands!' 'tall guys *mais* very tall!') present a co-oriented argumentation and the contrast is expressed by a semantic strengthening from p to q. We hypothesized that this should be reflected by both a lexical and a prosodic strengthening of the scope constituent q. To test this hypothesis, we compared the overbid use with the concessive use in conversational data. We found more adverbial intensifiers in the scope of the overbid than the concessive use. We also investigated both the phonetic and phonological realisation of *mais* itself and of its scope. We found that both *mais* are phrased with their scope rather than alone. Crucially, we found a significantly greater H tone rate (both per second and per syllable) on the scope of the overbid. We also found a lower articulation rate but only in the absence of adverbial intensifiers. This demonstrates the role of tonal density in prosodic strengthening.

Index Terms: prosody, tonal density, discourse markers, French, "mais"/"but"

1. Introduction

Although the contrastive functions of French connector *mais* ('but') have been described extensively ([1], [2], [3], [4], [5], [6], a.o.), the overbid function is rarely studied (but see [7], [8] and [9]) and very few works mention the role of prosody (but see [10] and [11]). In the present pilot study, we aim at showing that the semantic strengthening between the antecedent p and the scope q that characterises the overbid use of p *mais* q is conveyed by both lexical and prosodic means.

In most of the contrastive uses of connector *mais*, the antecedent p and the scope q are in contrast because they diverge argumentatively: p is explicitly refuted and replaced by q in *correction*, p and q give rise to opposite implications in *opposition*, and q is true despite p in *concession*. Conversely, in *overbid*, p and q have the same argumentative orientation and the contrast between them is between two wordings of the same predicate where p is presented as the ordinary wording and q as the more appropriate wording, although more extreme [12]. Example (1), (2) and (3) below, extracted from the CID [13], are prototypical explicit examples of the overbid use:

(1) on a vu arriver au milieu de la cour des espèces de **grands** mecs ... euh *mais* très **grands**

'We saw some kind of **big** guys ... uh MAIS **very big** arrive in the middle of the courtyard'

(2) et puis dans les deux trois maisons qui avaient été construites comme ça y en avait une euh **tout mais absolument tout** était de travers

'And then in the two three houses that were built like that there was one uh **everything** MAIS **absolutely everything** was crooked'

(3) y avait [...] une pauvre personne qui était tellement fatiguée qu'elle **ronflait ... mais** qui **ronflait ... bien**

'There was a poor person who was so tired that she **snored ...** MAIS who **snored ... well**'

In these examples, a predicate proposed by the speaker in p - an adjective in (1), a pronoun in (2), a verb in (3) - is resumed in q with an additional adverbial intensifier which indicates the preferred strengthened meaning. Example (4) below illustrates a case where the predicate in p is not resumed and completed but entirely reworded in q:

(4) et dès qu'ils ont servi à manger on s'est tous **présentés mais jetés comme des euh ... comme des miséreux** sur la bouffe

'And as soon as they served the food we all **came forward** MAIS **threw ourselves into the food** like uh ... like wretches'

Interestingly, examples like (5) below illustrate cases where the ordinary predicate in p is neither resumed nor reworded but elided in q (elided text is crossed out):

(5) enfin des trucs *mais* ~~des trucs~~ **fous** tu vois
'well **things** MAIS **crazy things** you see'

The ordinary predicate may even be elided in p like in example (6):

(6) et elle était ~~envahissante~~ *mais* su- **genre super envahissante** tu vois

'And she was ~~annoying~~ MAIS su- **like super annoying** you see'

The overbid use of connector *mais* just described has been said to originate in conversational speech ([9]) but analyses of this use all involve invented examples or citations from dialogues in literary works ([7], [8]) which make it difficult to study the role played by prosody. To our knowledge, no study has been conducted on the prosodic treatment of the overbid use of *mais*. Only [7] and [9] mention the presence of particularly salient accents on the scope q of *mais*, and of a silent pause after *mais*. A lengthened realisation of the connector *mais* itself is also evoked by [7].

The present pilot study therefore aims at proposing a first lexical and prosodic characterisation of the overbid *mais* in conversational speech by means of a systematic comparison with another contrastive use of connector *mais*, namely the concessive use, in a corpus of conversational data.

2. Material and method

2.1. Corpus

2.1.1. The Corpus of Interactional Data (CID)

The Corpus of Interactional Data [13] is an audio-video recording of French spontaneous face-to-face conversations (16 speakers, 8 hours, about 115.000 words). The manual transcription is automatically aligned at the phoneme level on the speech signal and various annotations (phonemes, syllables, part-of-speech, disfluencies, syntax, discourse units, prosodic accentual and intonational phrases, etc.) are available.

2.1.2. Material selection

As can be seen from the examples (1) to (6) above, *mais* in its overbid use coordinates mainly constituents smaller than the clause in elliptic structures. We therefore selected the 114 occurrences (out of 1617) of *mais* in the CID which correspond to this type of constructions, using the available syntactic annotation [14]. This does not rule out the possibility of this use in the case of clause coordination, but we leave this question for further research.

2.2. Discursive annotation

One of the authors manually annotated the contrastive functions of the selected occurrences of *mais*. We used the set of functions described in the second paragraph of section 1 above, namely: correction, opposition, concession and overbid. We also found a few cases of ‘non seulement... *mais* aussi’ (not only... but also) which form a separate category in [3]. Among the 114 selected *mais*, only 74 were classifiable without ambiguity. Only those occurrences are included in the present pilot study. Table 1 below shows the distribution of these occurrences in the different functional categories.

Table 1: Number of occurrences of each unambiguous discourse function of connector *mais*.

Discourse function	Nb of occurrences
Correction	10
Opposition	14
Concession	21
Overbid	27
Non seulement... <i>mais</i> aussi	2
TOTAL	74

In order to perform the lexical and prosodic characterisation of the overbid *mais* in our data, we compared the 27 occurrences of overbid *mais* with the 21 occurrences of concessive *mais* whose numbers are comparable.

2.3. Lexical comparison

As can be observed in the examples (1) to (6) of section 1 above, the semantic strengthening between the antecedent *p* and the scope *q* that is conveyed by the overbid use of *mais* is mostly achieved using adverbial intensifiers like ‘très’ (very) in (1) and ‘absolument’ (absolutely) in (2) within *q*. Moreover, [7] signals that the scope of the overbid use is often associated with tag questions like ‘hein?’ in French. We therefore annotated the presence of adverbial intensifiers and tags within the scope *q* for both selected overbid and concessive occurrences of *mais*.

2.4. Prosodic comparison

Our prosodic analysis aims at comparing the prosodic design of overbid and concessive uses of *mais* in our data. Based on informal observations by [7] and [9] mentioned in section 1 and our own observations, our main hypothesis was that prosody may contribute to the semantic strengthening characterising the overbid use. Since a few previous works like [15] showed that phonology (pitch accents, phrasing, etc.) may be relevant to contrast connector uses, we perform both a phonetic and a phonological analysis of the prosody of *mais* itself and of its scope *p*.

2.4.1. Phonological annotation

Following French ToBI [16], French intonation has three levels of phrasing: the accentual phrase (AP), the intermediate phrase (ip) and the intonational phrase (IP). The AP is defined by its underlying tonal pattern LHiLH* with a mandatory final pitch accent H* occurring on the last syllable of the last content word, and an optional initial hybrid accent Hi which may occur on the first syllable of the first content word. Optional L tones may occur on the function words or between the two H tones. Phrasal tones (L- or H-) occur at ip edges, and boundary tones (L% or H%) at IP boundaries. In focal position, other pitch accents may replace H*, like L* or a bitonal H+IH*.

Two authors of the present study performed the phonological annotation of *mais*, its antecedent *p* and its scope *q* using French ToBI, as illustrated by Figure 1 below.

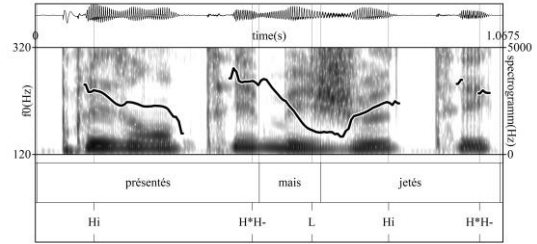


Figure 1: Tonal annotation of the antecedent *p*, the connector *mais* and the scope *q*, using French ToBI.

2.4.2. Prosodic hypotheses and analyses

We planned the prosodic analysis of *mais* itself on the one hand and of its scope on the other hand.

- Prosodic analysis of *mais* itself

Following the informal observations by [7] and [9], we hypothesised that overbid *mais* could be phrased alone more often, notably because following pauses have been observed, and that it may be longer compared to concession *mais*.

Consequently, we first observed the phrasing of *mais* relying on the tonal analysis described above, and assuming that *mais* is phrased alone in its own accentual phrase if it contains at least one pitch accent T* and if it is immediately preceded by a T* or an edge tone T- or T%.

We also manually counted the presence/absence of silent and filled pause before and after *mais* in both conditions.

Finally, we compared the duration of *mais* in milliseconds (ms) in both conditions.

Note that we didn't perform any acoustic measure of f0 on *mais* because of the nature of our data: an informal inspection showed that the f0 on *mais* was often highly influenced by the f0 on the word preceding *mais*, which introduced a strong bias in the acoustic analysis. Figure 1 above illustrates this point: the f0 curve on *mais* shows a falling movement which corresponds to the transition between the contingent high f0 values corresponding to the H*H- tonal targets on the preceding syllable and the L tonal target realised on *mais*. If the preceding word had ended in a low tone, the f0 on *mais* would have been flat.

- Prosodic analysis of the scope of *mais*

Based on our own informal inspection, and on the observation by [7] and [9] that the overbid use is often accompanied by some particularly salient accents, we hypothesised that its scope q may show a higher tonal density compared to concession *mais*. We also relied on a previous study by [17], showing that tonal density was a reliable prosodic mark of focus.

Consequently, we also hypothesised that raw f0 measures of f0 mean and f0 span [18] could reflect the semantic strengthening characterising overbid *mais* compared to concession *mais*. As for duration, we formulated no hypotheses concerning the length of the scope since it varied tremendously in number of words in our corpus data, a variability that we had no means to control for.

We calculated tonal density as the number of H tones per syllable, independent of the type of the tone: Hi, H*, H-, H%. Whenever two H tones were superposed in the frequential dimension on the same syntagmatic position, we counted only one H tone. This is the case when an H* and an edge H- or H% tones are realised on the same syllable (see for instance the last tonal position H*H- on the last syllable of the word 'jetés' in figure 1 above). In doing so, we aimed at being closer to the definition of the notion of tonal density (in the temporal dimension) and thus more conservative.

We also used Praat (Boersma & Weenink, 1992–2022) to compute the following phonetic measures : the f0 mean over the scope of *mais* normalised by the ratio between the mean of f0 means and the speaker's f0 mean (normalised $f0_{mean} = f0_{mean_scope} * \text{mean}f0_{means} / f0_{mean_speaker}$), the difference between the f0 maximum and the f0 minimum over the scope of *mais* normalised by the ratio between the mean of f0 differences and the speaker's f0 mean (normalised $f0_{span} = f0_{max} - f0_{min} * \text{mean}f0_{differences} / f0_{mean_speaker}$).

3. Results

3.1. Results of the lexical comparison

Recall that we compared the number of adverbial intensifiers and tag questions within the scope q for the 27 occurrences of the overbid use versus the 21 occurrences of the concessive use of *mais* in our selected data.

Indeed, in only 3 out of the 21 occurrences of concessive *mais*, an adverbial intensifier was present in the scope, whereas this was the case for 12 out of 27 occurrences of the overbid *mais*. According to Fisher's exact test, this difference is statistically significant ($p = 0.0319$).

Conversely, although the proportion of tags in the scope of overbid *mais* (12/27) was greater than in that of concessive *mais* (6/21), this difference is not significant according to Fisher's exact test ($p = 0.3693$).

We can conclude that, in our data, the presence of an adverbial intensifier is a robust distinctive characteristic of overbid *mais* compared to concessive *mais*, but this is not true for the presence of a tag question.

3.2. Results of the prosodic comparison

3.2.1. Results concerning the prosody of *mais* itself

- H1: overbid *mais* is phrased alone (in its own AP)

Among the 48 occurrences of *mais* that we examined, and following our tonal analysis, only one was phrased apart in its own accentual phrase (see figure 2 below), and one was phrased with the word 'enfin' with which it formed a compound. In the remaining 46 occurrences, *mais* was phrased with what followed, most of the time its scope (43 occurrences, i.e. 93%) and 3 times with some verbal material preceding its scope (like disfluent pre-formulation of the scope). Figure 3 below illustrates the most common phrasing of *mais* in our data. Indeed, no reliable difference could be observed between the concession and the overbid use of *mais*.

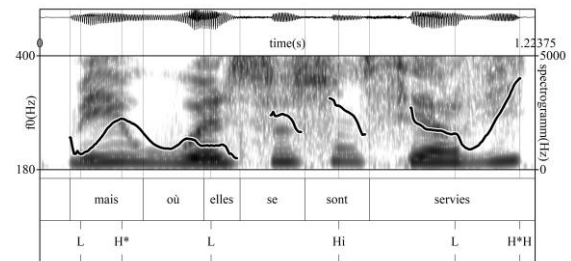


Figure 2: Phrasing of *mais* alone in its own accentual phrase LH*.

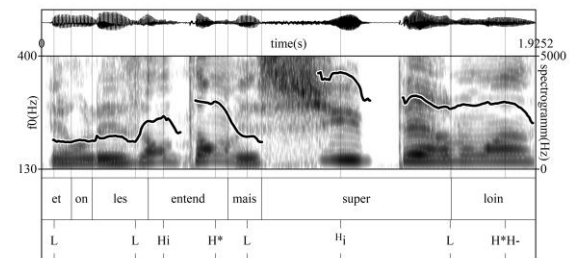


Figure 3: Phrasing of *mais* with its scope.

In our data, the realisation of pauses differed slightly for concession compared to overbid. For concession, we observed no pauses in 55% of the cases, but 35% of pauses before *mais* and 10% of pauses after *mais*. For overbid uses, we observed no pauses in 67% of the cases, but 15% of pauses before *mais* and 18% of pauses after *mais*. Hence, the preferred realisation of *mais* in our data is without pauses both in concession and overbid uses. In conclusion, H1 is not confirmed. Rather, both overbid and concession *mais* are phrased with their scope.

- H2: overbid *mais* is lengthened compared to concession *mais*

The mean duration of overbid *mais* in our data is 162.10 (± 27.448) milliseconds (ms) while the mean duration of concession *mais* is 146.14 (± 30.9) ms where the error bars are given with 95% confidence interval. Although there is a

tendency for overbid *mais* to be longer on average than concession *mais*, this difference is not statistically significant (one sided T-test p-value of 0.21), and H2 is not confirmed.

3.2.2. Results concerning the prosody of the scope of *mais*

- H3: the scope of overbid *mais* has a higher density of H tones compared to the scope of concession *mais*

The first two statistical analyses compare the distribution of the H tone rate per second and one of its variants, the distribution of the H tone rate per syllable, for the two conditions. For each selected sentence, the number of H tones in the scope area is computed by counting the number of target points labelled with at least one H tone (see section 2.4.2). The ‘H tone rate per second’ (HRD) is afterwards obtained by dividing this number of H tones with the duration (in seconds) of the scope area, excluding the duration of the pauses which might be present in the scope. The ‘H tone rate per syllable’ (HRS) is the number of H tones divided by the number of syllables in the scope area. The two distributions are plotted in figure 4. The two indicators show a higher rate of H tones for the overbid condition. For the variable ‘H tone rate per second’, the median value is $\text{med}(\text{HRD}, \text{‘concessive’}) = 1.84$ for concessive and $\text{med}(\text{HRD}, \text{‘overbid’}) = 2.24$ for overbid, and respectively $\text{med}(\text{HRS}, \text{‘concessive’}) = 0.33$ and $\text{med}(\text{HRS}, \text{‘overbid’}) = 0.56$ for the median value of the variable ‘H tones rate per syllable’.

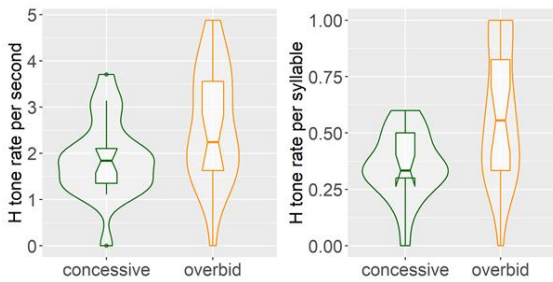


Figure 4: The distribution of the ‘H tones rate per second’ (left panel) and ‘H tones rate per syllable’ (right panel) for the concessive (left diagram) and overbid (right diagram) conditions.

These two effects are statistically significant. The results of a Wilcoxon rank test give a p-value of 0.024 for the H tones rate per second and a p-value of 0.0059 for the H tones rate per syllable. Compared to the concessive use of *mais*, the prosody of the scope in the overbid condition is amplified by an increase of the H tones rate. H3 is therefore confirmed by both measures.

- H4: the scope of overbid *mais* has a lower articulation rate compared to the scope of concession *mais*

We computed the ‘articulation rate’ (AR) in the scope area. This rate is herein defined as the number of syllables divided by the duration in seconds, excluding the duration of pause areas belonging to the scope. We found a lower articulation rate for the overbid condition than for the concessive condition. The median articulation rate is $\text{med}(\text{AR}, \text{overbid}) = 4.76$ for overbid to be compared with $\text{med}(\text{AR}, \text{concessive}) = 5.55$ for the concessive sentences. However, this syllable lengthening effect is not statistically significant for our dataset (i.e. a Welch t-test on the mean gives a p-value of 0.595).

We identified the presence/absence of an intensifier in the scope (see section 4.3) among the parameters allowing to control the variability of the articulation rate measurements. The scope intensifier parameter can potentially reduce the variability of the articulation rate measurements. There are 33 sentences over 48 in our data that are produced without an intensifier in the scope. Among this subsample of 33 sentences, the counts of concession versus overbid are well-balanced (18 concession for 15 overbid). We found again the lengthening effect on this subsample. The median articulation rate is $\text{med}(\text{AR}, \text{‘concessive’}) = 5.64$ for concession to be compared with $\text{med}(\text{AR}, \text{‘overbid’}) = 4.57$ for the overbid sentences. But this time the Wilcoxon test assesses the statistical significance of the discrepancy (p-value = 0.0162). This result can be interpreted as follows. The scopes without intensifier show a larger lengthening effect on average than the scopes with an intensifier since the latter already fulfills the role of semantic strengthening expected for the overbid. The production of an additional prosodic marker of semantic strengthening is therefore optional and is less present on average than for the scopes without intensifier. Only this latter case confirms H4.

- H5: the scope of overbid *mais* has a higher f0 mean compared to the scope of concession *mais*

The normalized f0 mean of overbid *mais* in our data is 194.50 (± 11.39) Hertz (Hz) while the normalized f0 mean of concession *mais* is 195.21 (± 10.21) Hz, where the error bars are given with 95% confidence interval. These values are not different (one sided T-test p-value of 0.45), so H5 is not confirmed in our data set.

- H6: the scope of overbid *mais* has a higher f0 span compared to the scope of concession *mais*

The normalized f0 span of overbid *mais* is 59.83 (± 12.07) Hertz (Hz) while the normalized f0 span of concession *mais* is 62.93 (± 12.9) Hz where the error bars are given with 95% confidence interval. These values are not different (one sided T-test p-value of 0.36), so H6 is not confirmed on our data set.

4. Conclusions

To sum up, the present pilot study showed that the overbid use of French connector *mais* in conversation conveys a contrast between an ordinary degree of a predicate and its extreme expression by three linguistic means implemented on its scope: the presence of adverbial intensifiers, a lower articulation rate in the absence of these lexical markers and a greater tonal density. Those are preliminary results obtained on a small data set. A larger scale study may confirm the tendencies observed for the lengthening of the connector *mais* itself and the lower articulation rate even in presence of lexical intensifiers.

Interestingly, the modest size of the data set allowed its manual annotation for the contrastive functions of *mais* and for the phonological tonal configurations. The success of the latter analysis also showed that adding a phonological ,to the usual phonetic ones gave clear, fine grained results, which were able to capture the prosodic specificity of the overbid use of connector *mais* in French conversation.

5. Acknowledgements

This work was carried out within the Labex BLRI (ANR-11-LABX-0036) and the Institut Convergence ILCB (ANR-16-CONV-0002).

6. References

- [1] J. C. Anscombre and O. Ducrot, “Deux mais en français?,” *Lingua*, vol. 43, no. 1, pp.23-40, 1977.
- [2] S. Bruxelles, G. Dos Reis Nunes, J. Gouazé, E. Fouquier, É., O. Ducrot and A. Remis, “Mais occupe-toi d’Amélie,” in D. Bourcier, O. Ducrot, & L. Maury, *Les mots du discours*, pp. 93–130, Ed. de Minuit, 1980.
- [3] J. M. Adam, “Deux ou cinq fonctionnements de MAIS en français?,” in J.-M. Adam, *Éléments de linguistique textuelle: Théorie et pratique de l’analyse textuelle*, pp. 192–211, Mardaga, 1990.
- [4] V. Hancock, *Quelques connecteurs et modalisateurs dans le français parlé d’apprenants avancés. Étude comparative entre suédophones et locuteurs natifs*, Doctoral dissertation, Stockholm University, 2000.
- [5] A. M. Silletti, “Mais et la rectification atténuative,” *Revista Electrónica de Lingüística Aplicada*, pp. 12–20, 2010.
- [6] G. Winterstein, *La dimension probabiliste des marqueurs de discours : nouvelles perspectives sur l’argumentation dans la langue*, Doctoral dissertation, Paris 7, 2010.
- [7] M. M. Garcia Negroni, *Réinterprétation et Scalarité: les instructions de relecture dans la langue*, Doctoral dissertation, Paris, EHESS, 1995.
- [8] F. Torterat, “Quel Apport des fonctions communicatives en lexicographie? Le cas de la conjonction mais,” *Revue romane*, vol. 38, no. 2, pp. 215-238, 2003.
- [9] A. R. Somolinos, L’évolution de mais surréalisant en français : De Fui t’en, mais vitelement! À Je kiffe, mais grave! *Revue de sémantique et de pragmatique*, vol. 33, pp. 231–249, 2013.
- [10] A. C. Simon and A. Grobet, “Intégration ou autonomisation prosodique des connecteurs”, in *Proc. Speech Prosody 2002*, pp. 647-650.
- [11] N. Dajko and K. Carmichael, “But qui c’est la différence?,” *Discourse markers in Louisiana French: The case of but vs. mais* *Language in Society*, vol. 43, no. 2, pp. 159–183, 2014.
- [12] Authors, in preparation.
- [13] P. Blache, R. Bertrand, G. Ferré, B. Pallaud, L. Prévot and S. Rauzy, “The Corpus of Interactional Data: a Large Multimodal Annotated Resource,” in Ide N. & Pustejovsky J. (eds), *Handbook of Linguistic Annotation, Springer bookseries Text, Speech, and Language Technology*, pp. 1323-1356, 2017.
- [14] S. Rauzy, G. Montcheuil and P. Blache, “MarsaTag, a tagger for French written texts and speech transcriptions,” *Second Asian Pacific Corpus linguistics Conference*, Mar 2014, Hong Kong, China. pp. 220-220, 2014.
- [15] J. Hirschberg and D. Litman, “Empirical studies on the disambiguation of cue phrases,” *Computational linguistics*, vol. 19, no. 3, pp. 501-530, 1993.
- [16] E. Delais-Roussarie *et al.*, “Chapter 3 - Intonational phonology of French: Developing a ToBI system for French,” in: Sonia Frota/ Pilar Prieto (eds), *Intonation in Romance*, Oxford, Oxford University Press, pp. 63–100, 2015.
- [17] C. Portes and U. Reyle, “Combining syntax and prosody to signal information structure: the case of French,” in *Proc. Speech Prosody 2022*, pp. 87-91, 2022.
- [18] D. Patterson and D. R. Ladd, Pitch range modelling: linguistic dimensions of variation. In *Proc. of ICPhS*, Vol. 99, pp. 1169-1172, 1999.