# SPEAKER-SPECIFIC ANTICIPATORY LABIAL COARTICULATION IN FRENCH

Fanny Guitard-Ivent[1,2], Louise Wohmann-Bruzzo[1], Nicolas Audibert[1], Cécile Fougeron[1]

[1]Laboratoire de Phonétique et Phonologie (UMR7018, CNRS/Sorbonne Nouvelle)
[2]Laboratoire de NeuroPsychoLinguistique, Université de Toulouse, Toulouse, France
fanny.guitard-ivent@univ-tlse2.fr; {louise.wohmann, nicolas.audibert, cecile.fougeron}@sorbonne-nouvelle.fr

## ABSTRACT

This study examines inter-speaker variability in anticipatory labial coarticulation in French. Production of /sy/ sequences by fourteen French speakers originating from two databases, are compared to unrounded counterparts (/si/-/se/), in terms of spectral lowering measured dynamically over the /s/ duration. Results show speaker-specific coarticulatory patterns: (a) in the degree of labial anticipation measured by the lowering of the center of gravity of /s/ in the rounded context; (b) in the span of anticipation with speakers who anticipate the /y/ much earlier than others; and (c) in the variability of the coarticulatory patterns across recording sessions, with some speakers being more stable in their labial anticipation than others. Results are discussed in relation to individual preferences in the coordination of speech units, which implementation in the speech signal may index information about the speaker.

**Keywords**: labial anticipation, coarticulation, speaker variability, fricatives, spectral center of gravity.

## 1. INTRODUCTION

Although we are used to segmenting speech into well-defined units, such as phonemes, syllables, etc., and representing them in a regular and more or less identical manner, we know that these units vary to a greater or lesser extent not only from one individual to another, but also from one production to another of the same individual. Inter-individual variability is usually related to speaker-specific physiology, but also to individual production strategies [1, 2, 3, 4].

In the context of a larger project looking at individual properties in the speech signal, the present paper focuses on speaker-specific patterns of coarticulation. Inter-speaker variability in coarticulation has been reported in various languages and for different types of contextual effects. For instance, individual differences have been observed for vowel nasalization in CVN and NVN sequences in American English [5], for lingual Consonant-to-Vowel coarticulation in Dutch [6] and in French [7], for labial V-to-C coarticulation in French [8], and for lingual anticipatory V-to-V coarticulation [9]. In their study with 10 French speakers producing non-sense iCy sequences (with non labial Cs) [8] showed that some speakers anticipate the protrusion gesture of the /y/ much earlier in the consonant than others.

The present study aims to investigate further this speaker-specific pattern of anticipatory labial coarticulation in French, but in continuous and meaningful speech, and by looking at its acoustic consequences. Indeed, if coarticulation can be used to perceptually differentiate speakers as suggested by [10], it needs to have acoustic consequences. For this reason, we investigate /sy/ sequences, where anticipation of the labial gesture of the vowel can be tracked dynamically on the spectral characteristic of the fricative noise. Several studies have shown that fricatives exhibit high inter-speaker variability [11, 12, 13]. Based on the fact that there is speaker-specific information both in /s/ and, presumably, in labial anticipation, this sequence is particularly interesting.

Speaker-specific patterns will be defined by looking at different aspects: the amount of contextual acoustic difference between a [s] produced in a rounded context compared to an unrounded context; the span of anticipation as measured by dynamic acoustic variation over the /s/; but also by comparing speakers in terms of variability of their coarticulatory patterns across recording sessions.

## 2. METHOD

In order to study coarticulation on the production of 14 speakers in total, we have merged here productions extracted from two different corpora of read speech. The first corpus (corpus A) is extracted from the French PATAFreq database [14] that included 6 female speakers (F01 to F07) and 3 male speakers (H01-03) recorded in 8 to 10 sessions, over a two-month period. Each session includes the reading of three short texts, which contain 6 words

presenting a [sy] syllable and 6 words containing a [se] syllable (unfortunately no /si/ were present in the texts). In total, this corpus includes 24 to 60 /sy/ per speakers and 24 to 62 /se/ per speaker. The second corpus (corpus B) includes the production of five French female speakers (AE, AN, AP, EDF, ES) recorded in 5 sessions over two weeks [15]. In each session, speakers read a set of sentences, among which two contained [si] and [sy] sequences, and these sentences were repeated 9 times in a random order. In total, this corpus contains 135 /sy/ per speakers and 108 to 134 /si/ per speaker.

In each sound file, the fricative /s/ and the following vowel were segmented manually. The spectral Centre of Gravity (COG) was computed over the /s/ on 10 equally spaced points with a 10 millisecond Hanning window centered on the target point, and filtered to retain frequencies ranging from 350Hz to 18kHz.

COG trajectories were analyzed using GAMMs on R with the package mgcv [16]. Our analyses include autoregressive error models (AR1 models) that take into account the dependence between neighboring points within the same COG trajectory. Estimates of the difference between the rounded and the unrounded contexts were computed using the function get_difference of package itsadug [17], at each time point. Using the same method, pairwise differences between speakers (all sessions pooled) give a measure of inter-speaker variability, and pairwise differences between sessions of a given speaker give a measure of intra-speaker variability.

## 3. RESULTS

### 3.1 Is coarticulation speaker dependent?

COG trajectories in rounded and unrounded V2 contexts were modeled for each speaker separately, taking into account the duration of /s/ and including a random smooth per repetition.

All 14 per-speaker models account for 93 to 95% of the variance in our data and indicate an effect of V2 on the shape of the COG trajectories of /s/ as well as on the average COG frequency of these trajectories, except for one male speaker H03. For this speaker, the COG of /s/ was not significantly lowered in the rounded vs. unrounded context. For all 13 other speakers, as expected, the COG of /s/ is significantly lowered and its trajectory is affected by the rounding of the following vowel.

Figure 1 highlights the differences found between speakers in (a) the degree of anticipatory labial coarticulation, displayed as the magnitude of the difference in COG between the two V2 contexts, with darker shades showing more lowering of COG in the rounded V2 context, but also (b) in terms of span of anticipation, illustrated by the proportion of /s/ which is significantly different in the two contexts as indicated by the dark red line overlaid on the x-axis. The difference is considered significant when the 95% credible interval of the difference does not include value 0.

Speakers can be separated into subgroups based on differences in COG trajectories between the unrounded and rounded context. Speakers F02, F05, F07, H01, H02, AN, EDF, ES, F01 show a striking difference between the two contexts both in magnitude of the acoustic difference and in the span of anticipation. AE and F03 show an early start of anticipation but, if the acoustic difference is on the largest part of the /s/, it is not that strong in magnitude, like F04 for which in addition the effects of coarticulation appear much later. AP and H03 present little acoustic differences, which appear only at the end of the /s/.

### 3.2. Are speakers who strongly coarticulate stable in their coarticulatory patterns across sessions?

To measure whether the variability in coarticulatory pattern is speaker dependent, we selected the 9 speakers who coarticulate the most (see Figure 2) and modeled the trajectories of their /s/ in /y/ context as a function of the speaker and the recording session (8 sessions for speakers of corpus 1 and 5 for those of corpus 2). Our models explain 92% of the variance for corpus 1 and 95% for corpus 2.

In order to focus on intra- versus inter-speaker variability, we report the results of the pairwise comparisons only (see section 2). Regarding individual variability, some speakers are more variable than others across sessions. While speakers like EDF and F07 are very consistent in their coarticulatory patterns, other speakers like F02 and H02 are much more variable. Moreover, variability patterns do not depend on the overall degree of coarticulation: for instance speaker F07 is both a strong and stable coarticulator, while speaker F02 is a strong and variable coarticulator.

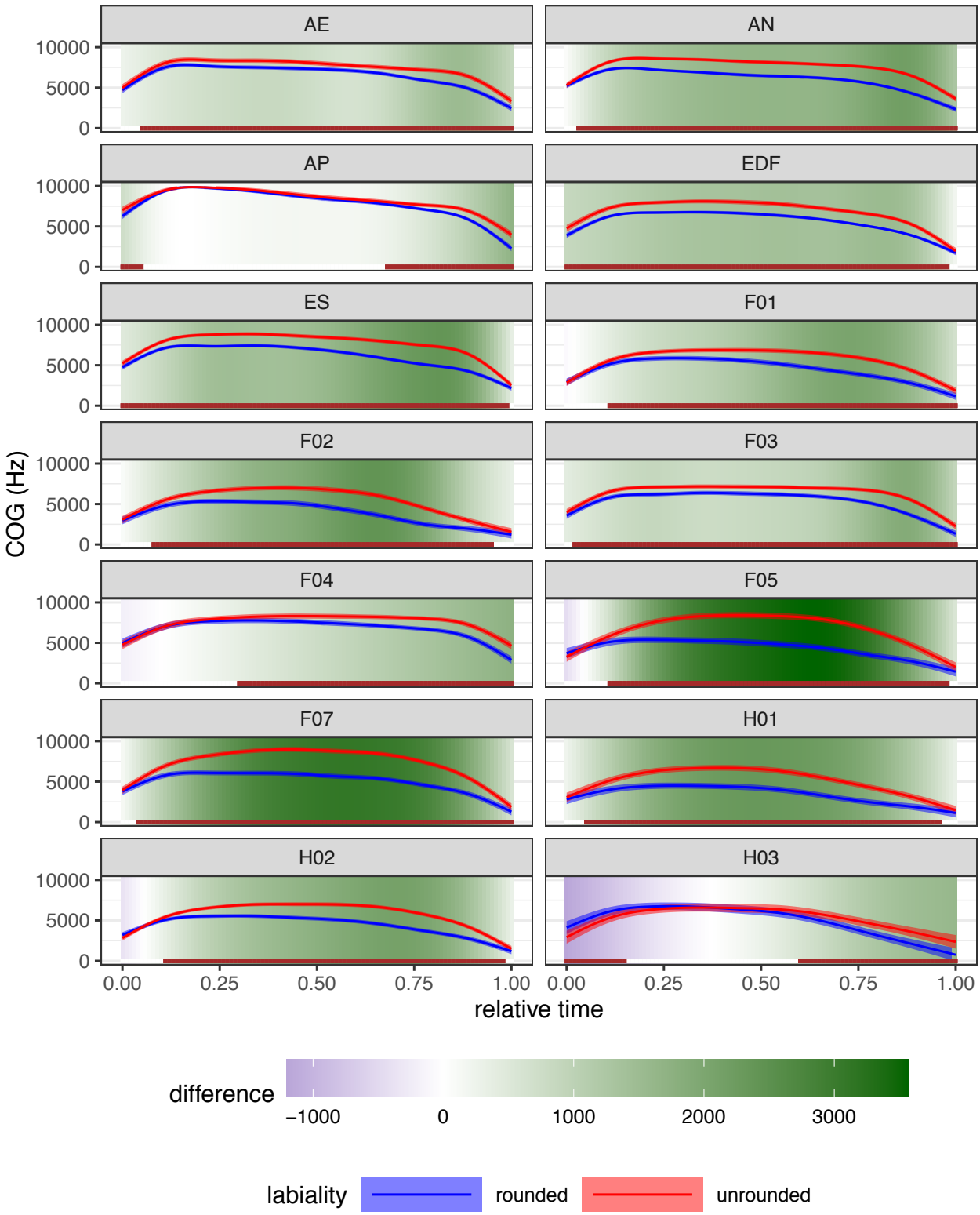


**Figure 1:** COG trajectory of /s/ estimated by GAM models as a function of V2 (rounded in blue and unrounded in red) for each speaker. The background color quantifies the difference between the two trajectories. Highlighting of the x-axis in dark red indicates that the difference is significant.

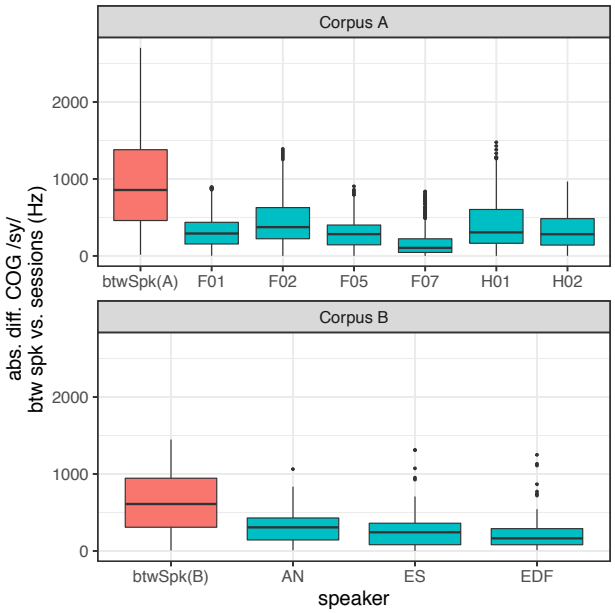


**Figure 2:** Distribution of absolute pairwise differences between speakers (all sessions pooled, in red) as a measure of inter-speaker variability, and between sessions of a given speaker (in blue) as a measure of intra-speaker variability, for both corpora.

## 4. DISCUSSION AND CONCLUSION

In this study we confirm with 14 speakers, numerous tokens per speaker, and in continuous (read) speech production, that the anticipation of the labial gesture of /y/ in a preceding /s/ is speaker-specific. Thanks to the dynamical tracking of the spectral characteristics of the fricative noise, we can better characterize these individual coarticulatory patterns. Indeed, speakers differ both in terms of the amount of spectral lowering of the fricative noise induced by the anticipation of the labial gesture of /y/, and in terms of the shape of the COG lowering trajectory indicating how much of the /s/ duration is affected by the anticipation of this labial gesture. While most of the speakers show a large amount of spectral lowering and an early anticipation, quite a few speakers show very little acoustic cues of coarticulation (H03 & AP), and one speaker appears to anticipate the labial gesture later than the others (F04).

Moreover, our comparison of the same speakers over multiple recording sessions separated by several days, showed that for most of the speakers showing labial anticipation, there is little variation in the way they coarticulate across sessions. These results mirror that of [18] who observed C-to-V coarticulation across 300 tokens of words produced by the same speaker recorded in three sessions. The vowel acoustic at its mid point and the formant trajectories showed a striking stability across repetitions, suggesting that both the vowel targets and the trajectories out and to the surrounding consonants are controlled and planned by the speaker. Our results and that of [11], who showed large individual variation in the COG trajectories of fricatives (especially for /s/) depending on the surrounding vowels, further support individual preferences in the coordination of speech units, whose implementation in the speech signal may index speaker-specific information.

Individual patterns of coarticulation, like other types of speaker-specific articulatory aspects, could be attributed to physiological and anatomical characteristics of the speaker (e.g. [13]), but also to speaker-specific strategies to attain a desirable speech rate. In a study of 246 speakers, [9] showed that anticipatory Vowel-to-Vowel coarticulation is stronger for speakers with the fastest speech rates, but only for younger speakers. We did not measure speech rate on each of the sentences produced in our data, however, post hoc observation of the degree of lowering of the COG of /s/ as a function of its duration for each speaker does not seem to indicate a link between the two. Indeed, speakers seem to be constant in their degree of coarticulation regardless of the duration of /s/. Moreover, our models take into account the duration of the consonant and for the majority of them (9 / 14) this parameter is not significant. In other studies with the speakers included in corpus A [19, 20] we found that speaker F07, who present the most stable coarticulatory pattern, is quite variable in speech rate across sessions, while speaker F02, who show more variable coarticulation, has a stable speech rate across sessions. Future studies are thus desirable to understand better how these individual patterns of labial coarticulation relate to articulation rate and to other speaker-specific articulatory and coarticulatory behaviors.

## 5. ACKNOWLEDGEMENTS

This study was funded by the subside CRSII5_202228 (ChaSpeePro) of the Swiss National Science Foundation, and by the programs ANR-10-LABX-0083 (Labex EFL) & ANR-17-CE39-0016 (VoxCrim).

## 5. REFERENCES


[1] Atkinson, A. 1976. Inter- and intraspeaker variability in fundamental voice frequency. *J. Acoust. Soc. Am.* 60, 440-445.

[2] Smith, R., Hawkins, S. 2012. Production and perception of speaker-specific phonetic detail at word boundaries. *J. of Phonetics*. 40, 213-233.

[3] Serrurier, A., Badin P., Lamalle L., Neuschaefer-Rube C. 2019. Characterization of inter-speaker articulatory variability: A two-level multi-speaker modeling approach based on MRI data. *J. Acoust. Soc. Am.* 145 (4), 2149-2170.

[4] Toda, M. 2009. Étude articulatoire et acoustique des fricatives sibilantes. PhD dissertation. Université de la Sorbonne Nouvelle, Paris.

[5] Zellou, G. 2017. Individual differences in the production of nasal coarticulation and perceptual compensation. *J. of Phonetics*. 61, 13–29.

[6] van den Heuvel, H., Cranen, B., & Rietveld, T. 1996. Speaker variability in the coarticulation of /a, i, u/. *Speech communication*. 18(2), 113-130.

[7] Guitard-Ivent, F. 2018. Coarticulation C-à-V en français : interaction avec le type de voyelle, la position prosodique et le style de parole. PhD dissertation. Université de la Sorbonne Nouvelle, Paris.

[8] Robert, V., Wrobel-Dautcourt, B., Laprie, Y., & Bonneau, A. 2005. Inter speaker variability of labial coarticulation with the view of developing a formal coarticulation model for French. In: *5th Conference on Auditory-Visual Speech Processing-AVSP*.

[9] D'Alessandro, D., Fougeron, C. 2021. Changes in Anticipatory VtoV Coarticulation in French during Adulthood. *Languages*. 6(4), 181.

[10] Zellou, G. 2022. Coarticulation in Phonology. In: Cambridge Elements in Phonology Series. Cambridge: Cambridge University Press.

[11] Carralero-Fernandez, S. 2022. Speaker Specific Information in the Acoustic Characteristics of English Fricatives. *Int. J. Forensic Linguistics.* 3, 105-115.

[12] Gordon, M., Barthmaier, P. & Sands, K. 2002 A cross-linguistic acoustic study of voiceless fricatives. *J. Int. Phonetic Association*. 32, 141-171.

[13] Fuchs, S. & Toda M. 2010. Do differences in male versus female /s/ reflect biological or sociophonetic factors?. In Fuchs, Zygis, Toda (eds) *Turbulent sounds: An interdisciplinary guide*. Mouton de Gruyter.

[14] Fougeron, C., Audibert, N., Gendrot, C., Chardenon, E., & Wohmann, L. 2022. PATATRA and PATAFreq: two French databases for the documentation of within-speaker variability in speech. *Proceedings of 13th LREC*, 1939-1944.

[15] D'Alessandro, D. 2022. Interspeaker variations in V-to-V coarticulation: effects of Motor Speech Disorders, age, speech tempo changes, and boundary type. PhD dissertation. Université de la Sorbonne Nouvelle, Paris.

[16] Wood, S. 2011. Fast stable restricted Maximum Likelihood and Marginal Likelihood Estimation of Semiparametric Generalized Linear Models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 73(1), 3–36.

[17] van Rij, J., Wieling, M., Baayen, R. & van Rijn, H. 2022. itsadug: Interpreting time series and autocorrelated data using GAMMs. R Package, version 2.4.1

[18] Whalen, D. H., & Chen, W. R. 2019. Variability and central tendencies in speech production. *Frontiers in Communication*, 4, 49.

[19] Fougeron, C., Audibert N. 2022. Variabilité intra-individuelle en parole lue. In: *32e Journées d'Études sur la Parole*. Noirmoutier, France

[20] Audibert, N., Fougeron, C. 2022 Intra-speaker phonetic variation in read speech: comparison with inter-speaker variability in a controlled population. In: *Proc. Interspeech 2022*, 4755-4759.