CV COORDINATION: THE CASE OF ENCHAÎNEMENT AND LIAISON IN FRENCH

Sejin Oh, Cécile Fougeron, Philipp Buech, Anne Hermes

Laboratoire de Phonétique et Phonologie, UMR 7018, CNRS/Sorbonne Nouvelle (France)
se-jin.oh@cnrs.fr, cecile.fougeron@sorbonne-nouvelle.fr, philipp.buech@sorbonne-nouvelle.fr, anne.hermes@sorbonne-nouvelle.fr

ABSTRACT

The current study explores how different types of external sandhi in French affect the temporal coordination of gestures. Specifically, we examine the temporal coordination of three different types of CV in French (Onset CV, Enchaînement CV, and Liaison CV) at normal and fast speech rates. The results from this articulatory study revealed that the three different types of CV display the same pattern with respect to their temporal measures (C duration, gestural overlap between C and V, lag between C and V onsets, and lag between C and V targets), at both speech rates. Moreover, the effect of C duration on target-to-target lag is uniform across the different types of CV in French. These results suggest that the CV coordination driven by external sandhi in French exhibits the same coordination (in-phase) as true onset CV.

Keywords: CV coordination, liaison, enchaînement, speech rate, French

1. INTRODUCTION

Syllable structure is hypothesized to be associated with a characteristic pattern of temporal coordination [1-4]. For example, when a gesture occurs in a syllable onset position, it is coordinated in-phase with the following vowel, with the two gestures activated at the same time. In contrast, a coda gesture is hypothesized to be coordinated anti-phase with the preceding vowel, showing a sequential timing between the two gestures. However, a consonant which was an underlyingly coda can also be resyllabified to the following syllable when it is followed by a vowel. Fundamental questions are whether a resyllabified consonant is a 'true' onset, and whether it is shown in the coordination pattern.

In French, there are two different types of resyllabification: 1) an underlying word-final coda consonant can be resyllabified with the following word-initial vowel (Enchaînement CV); 2) a word-final liaison consonant, i.e., a latent consonant surfacing in some French words only when the following word starts with a vowel, can be resyllabified with the following word-initial vowel in what is called 'liaison enchaînée' (Liaison CV). As shown in Table 1, for example, these resyllabification cases and the case with a true word-initial onset consonant (Onset CV) are distinct underlyingly but are said to be homophonous in French.

As with other external sandhi processes, enchaînement and liaison processes have been extensively studied because of their potential cost on word recognition. However, previous perception results have shown that these processes do not generate lexical ambiguities (see e.g., [5, 6]). Acoustic studies have also shown that the phonetic neutralization between forms like the ones presented in Table 1 is incomplete: liaison and enchaînement consonants can be shorter in acoustic duration, the vowel preceding them can be longer, and/or they can preserve some specific allophonic properties associated with their word-final position [5-10]. However, arguments related to the preservation of phonetic contrasts between sequences created by enchaînement or liaison processes and sequences containing underlying onsets, do not rule out the process of resyllabification per se. Indeed, the sequence can preserve some phonetic properties related to its underlying lexical form (with an underlying word-final consonant) and at the same time adopt a 're-syllabified' organization, that could surface in the way the C and the V are temporally coordinated.

Onset	**Enchaînement**	**Liaison**
/CV#**<u>CV</u>**/	/CV.**<u>C#V</u>**/	/CV.**<u>CL#V</u>**/
petit **t**amis /pəti # tami/ [pə.ti.**t**a.mi]	peti**t**e amie /pətit # ami/ [pə.ti.**t**a.mi]	peti**t** ami /pəti # ami/ [pə.ti.**t**a.mi]

Table 1: Examples of three types of CV in French.

An EMA study by [11] examined two types of CVs in Korean, one with a true onset C and the other with a coda C, which turned out to be a resyllabified onset. They reported that there was no difference between the true onset CV and the resyllabified CV, in their coordination with the following vowel and in their stability. However, it might be attributable to surface ambiguity due to a consistent +/- lag between coordinated landmarks [12]. That is, a lag may cause surface ambiguity between in-phase and anti-phase timing, suggesting that overlap patterns alone may not be enough to determine which landmarks show coordination.

Recently, [13] hypothesized that variation in the duration of the first gesture (G_1), e.g., in CC, influences the G1 onset to G2 onset lag and depends on its coordination relations: for complex segments, such as palatalized /p^j/, the onset of G_2 (the palatal gesture) is temporally coordinated with the onset of G_1 (the labial gesture), while for segment sequences, such as /pj/, the onset of G_2 is temporally coordinated with the offset of G_1. These coordination relations were explored by investigating how the lag between the onset-to-onset varied with G_1 duration. The key finding involved differences between English consonant-glide sequences, e.g., [pj], [bj], [mj], [vj], and Russian palatalized labials, e.g., [p^j], [b^j], [m^j], [f^j], [v^j]. They found that for palatalized consonants in Russian, variation in G_1 duration had little effect on lag, which is consistent with the hypothesized temporal coordination for complex segments. In contrast, for English sequences, as C_1 duration increased, so too did the lag between consonant and glide gestures, showing a strong positive correlation.

Exploiting this temporal diagnosis, the current study aims to understand the temporal coordination of three different CV types (Onset CV, Enchaînement CV, and Liaison CV). The present study asks the following research questions:

- Is there a temporal difference among the three different types of CVs and does this difference hold at different speech rates?
- Do the different types of CVs affect the temporal coordination of gestures?

2. METHODS

2.1. Participants

Three female French native speakers (ages: 27, 42 and 51 years old) who speak Metropolitan French participated in the EMA experiment.

2.2. Speech materials and procedure

Articulatory data was recorded using Electromagnetic Articulography (EMA, AG501) at a sample rate of 1250 Hz and filtered afterwards with a Butterworth low pass filter (cutoff frequency: 20 Hz, order 5). The acoustic signal was recorded with a head-mounted microphone at a sample of 48 kHz and 16-bit resolution. To collect articulatory data, eight sensors were attached to the participants: three sensors midsagittal for tongue movements (tongue tip, blade and dorsum), two for lip movements (upper and lower lips), and two for reference (left/right mastoids). The data was corrected for head movements using the AG501's built-in head-correction procedure.

The data analyzed and presented here were part of a larger study of the temporal organization of gestures in French. Relevant to our purposes here were the three minimal and near-minimal triplets, where each triplet consisted of an Onset CV, Enchaînement CV, and Liaison CV (see Table 2). The target sequences were produced within a carrier sentence (Il y a la carte <X> avant la carte <Y> ici. ‘There is an <X> card before the <Y> card here.’), where each sequence presented in the current study was located in <X> in the sentence.

Onset /CV#CV/	**Enchaînement /CV.C#V/**	**Liaison /CV.CL#V/**
un petit tamis	une petite amie	un petit ami
un petit tapis	une petite abyme	un petit habit
un maudit tanneurs	une maudite année	huit années

Table 2: Stimulus items.

To test the stability of our coordination measures against variation in speech rate, participants were asked to read aloud each given sentence at their normal rate and then self-selected fast speaking rates in two successive blocks. Each sentence was produced 7 times per speaker at each rate. Of the 378 tokens collected, 22 tokens (5.8%) were discarded.

2.3. Data analysis

WebMAUS v3.12 [14] was used for acoustic alignments. Afterwards, the onset, target, release, and offset of the consonantal gesture for /t/ (TT sensor) and of the vocalic gesture for /a/ (TD sensor) were automatically identified using a custom script written in Python v3.10.8 [15]. These landmarks were detected within a window of the acoustic boundaries ±100 ms of [t] and of the sequence [ta] for the vocalic gesture, using 20% threshold of the maximum and minimum peak velocity (target, release) and the first zero crossing of the velocity before/after the peak velocities (onset, offset). As shown in Figure 1, the four key temporal intervals computed from these articulatory landmarks were (1) C DURATION = $C_{RELEASE} - C_{TARGET}$ (consonant plateau duration); (2) OVERLAP = $C_{RELEASE} - V_{ONSET}$; (3) ONSET LAG = $V_{ONSET} - C_{ONSET}$; (4) TARGET LAG = $V_{TARGET} - C_{TARGET}$. We also analyzed the correlation between C DURATION and TARGET LAG.

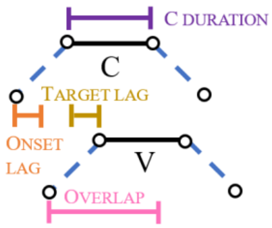

Figure 1: Four key temporal intervals: (1) C DURATION, (2) OVERLAP, (3) ONSET LAG, (4) TARGET LAG

2.4. Statistical models

The *brms* package v2.18.0 [16] in R v4.2.2 [17] was used for the statistical analysis. Bayesian linear mixed models were performed for C DURATION, OVERLAP, ONSET LAG, and TARGET LAG. The model structure was the same for all response variables, that is CV TYPE (Onset CV, Enchaînement CV, and Liaison CV), SPEECH RATE (normal, fast) and their interaction CV TYPE*SPEECH RATE as fixed parameters, and SPEAKER and ITEM as random intercepts. An additional model, for TARGET LAG as the outcome variable, the interaction between C DURATION and CV TYPE as fixed factors and SPEAKER and the ITEM as random intercepts, was performed. All models were run with standard priors and 4 chains with 6,000 samples each from which 3,000 were for warm-up, making a total of 12,000 post-warmup samples. Strong evidence for an effect being non-zero was determined if 0 was not included in the 95% Credible Interval (CI) of the posterior distribution. We report the posterior means (β) and the CI in square brackets.

3. RESULTS

3.1. Temporal differences

Figure 2 shows the distribution of C DURATION (A), OVERLAP (B), ONSET LAG (C), and TARGET LAG (D) across CV TYPE. Normal and fast speech rates are shown on the left and right panels, respectively. The three types of CV exhibit the same pattern for all four temporal measures for both speech rates.

First, the statistical analysis from Bayesian regression models reveals that there was no evidence for C DURATION of Enchaînement CV or Liaison CV being different from that of Onset CV (Enchaînement β=-1.45 [-5.79, 2.89]; Liaison β=-1.71 [-6.03, 2.59]). Moreover, although we found evidence for C DURATION being shorter for a faster speech rate compared to a normal speech rate (β=-10.21 [-13.59, -6.88]), we have no evidence for an effect of the interaction of speech rate and CV type on C duration (Enchaînement_fast β=2.64 [-2.06, 7.33]; Liaison_fast β=3.79 [-0.90, 8.60]). That is, within each speech rate, C DURATION of Enchaînement CV or Liaison CV is not systematically different from that of Onset CV.

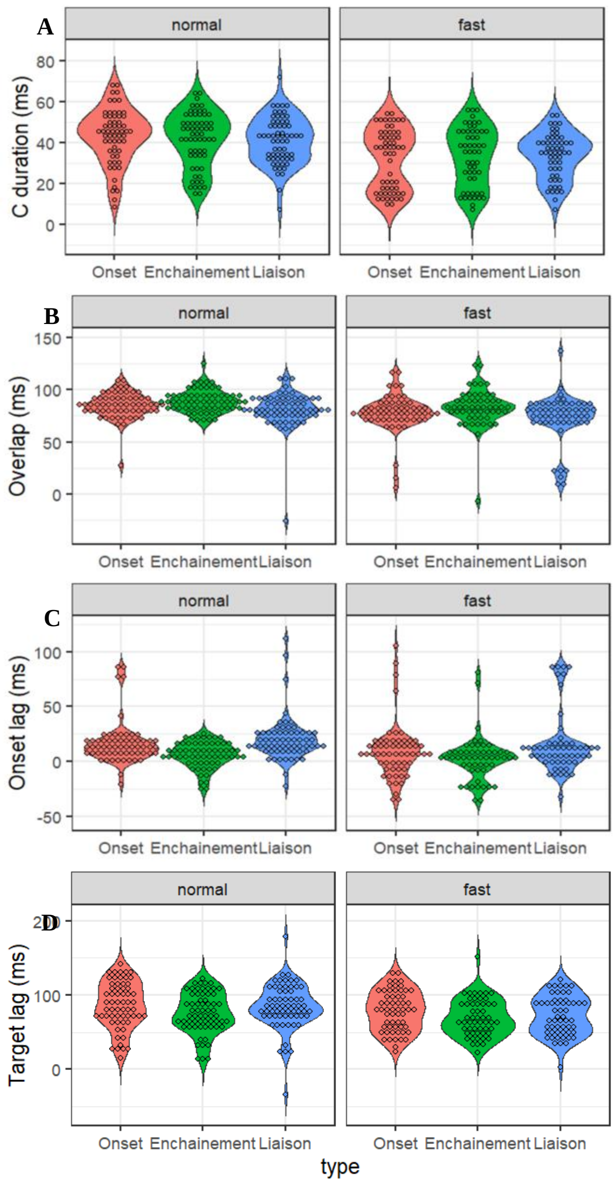

Figure 2: C duration (A); Overlap (B) and Onset lag (C); Target lag (D) at normal speech rate (left column) and at fast speech rate (right column).

We found the exact same pattern for OVERLAP and ONSET LAG. The statistical results reveal that different types of CV are consistent in their OVERLAP (Enchaînement β=3.85 [-4.31, 11.76]; Liaison β=-3.85 [-11.63, 4.04]) and in their ONSET LAG (Enchaînement β=-10.51 [-22.50, 1.54]; Liaison β=4.99 [-6.09, 16.45]). In addition, a faster speech rate also exhibits shorter OVERLAP and ONSET LAG compared to a normal speech rate (OVERLAP β=-7.12 [-13.02, -1.42]; ONSET LAG β=-7.81 [-14.56, -1.14]). However, we have no evidence for an impact of the interaction of speech rate and CV type on OVERLAP (Enchaînement_fast β=-0.22 [-8.07, 8.48];

Liaison_fast β=-1.76 [-10.11, 6.59]), and so is an effect of the interaction of speech rate and CV type on ONSET LAG (Enchaînement_fast β=5.75 [-3.64, 15.19]; Liaison_fast β=4.86 [-4.68, 14.33]).

Regarding TARGET LAG, Enchaînement CV or Liaison CV is not systematically different from that of Onset CV (Enchaînement β=-10.55 [-30.03, 8.68]; Liaison β=-1.61 [-20.48, 17.61]). Also, no evidence was found that normal and fast speech rates were systematically different (β=-6.05 [-15.30, 3.30]). Moreover, we have no evidence for an effect of the interaction of speech rate and CV TYPE on TARGET LAG (Enchaînement_fast β=0.34 [-12.74, 13.52]; Liaison_fast β=-4.56 [-18.04, 8.71]).

3.2. Gestural coordination of CVs

Figure 3 plots the relation between C DURATION (pooled duration at both rates; x-axis) and TARGET LAG (y-axis) across CV TYPE. Changes in C DURATION have little effect on TARGET LAG, showing only a slight upward trend. Notably, we observe the same pattern for all CV TYPE. To assess these results, we fit a series of Bayesian regression models to the data. We found evidence that variation in C DURATION impacts TARGET LAG, but the impact is small (β=0.82 [0.50, 1.15]). Crucially, however, we also found evidence that the way that variation in C DURATION impacts TARGET LAG is uniform across different types of CV in French (Cdur_Enchaînement β=-0.25 [-0.68, 0.18]; Cdur_Liaison β=0.40 [-0.09, 0.89]). To sum up, the results indicate that Enchaînement CV and Liaison CV showed the same temporal coordination as Onset CV.

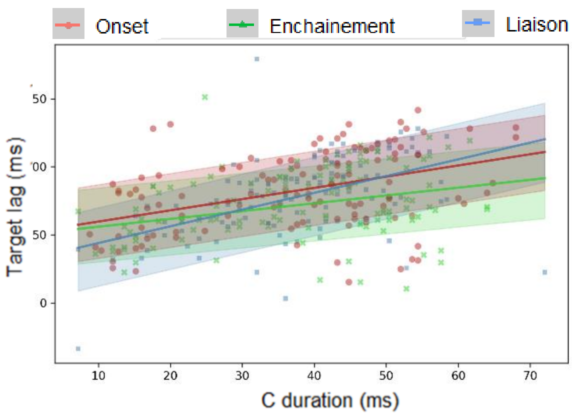

Figure 3: Target lag (ms) by C duration (ms) across CV types.

4. DISCUSSION & CONCLUSION

The present study investigated the temporal coordination between C and V in three different types of CV in French (Onset CV, Enchaînement CV, and Liaison CV) using kinematic data from Electromagnetic Articulography (EMA). We found that all CV types display the same pattern in their C duration, the overlap between C and V, the lag between C and V onsets, and the lag between C and V targets. Although, except for TARGET LAG, we found shorter duration on the other three measures for a faster speech rate compared to a normal speech rate, the interaction between CV types and speech rate on these measures was not significant. This suggests that Enchaînement CV and Liaison CV are not significantly different from Onset CV within each speech rate. Furthermore, we found a small impact of C DURATION on TARGET LAG. As [12] noted, this may be attributable to local variation in speech rate that independently influences both C duration and the lag. Crucially, however, the effect of C DURATION on TARGET LAG is uniform across different types of CV. This suggests that Enchaînement CV as well as Liaison CV in French show the same in-phase timing as a true Onset CV.

The results of the present study provide evidence that resyllabified consonants are also timed in-phase with the following vowel, as Cho and colleagues [11] have found for Korean. This resyllabification across words raises many questions—to be resolved—on the process of inter-gestural coordination. If the coordination between gestures within a syllable is specified in a word-size gestural score, how could it be modified in running speech [18]? Reversely, should this so-called resyllabification be considered to be a simple syllabification, surfacing from the coordination of gestures planned over units spanning several connected words? If so, how can we account for the results of previous studies which reported that speakers, as well as listeners, preserve phonetic contrasts among sequences with lexical onsets and sequences with external sandhi in French (cf., introduction)? One of the more frequently reported acoustic differences is the shorter acoustic duration of enchaînement C and liaison C compared to a true onset C. While acoustic measurements remain to be tested on this data, in the current study where temporal cues are tested on a wide set of articulatory events, we did not replicate this difference: the three types of CV are similar in their constriction duration, as well as their temporal lags and gestural overlap.

Obtaining kinematic data from EMA allowed us to examine the temporal coordination of gestures involving the articulation of external sandhi in French. Moreover, we make use of the concept of coordination to relate speech kinematics to the syllable structure. In future work, we will further investigate the temporal relations of C with the preceding vowel, which was also reported to be one of the acoustic characteristics preserving the contrast between the three sequence types.

5. ACKNOWLEDGEMENT

This research was partly funded by ChaSpeePro (CRSII5_202228) of the Swiss National Science Foundation and by the French ANR (ANR-10-LABX-0083, LABEX-EFL).

6. REFERENCES

[1] Goldstein, L., Byrd, D., & Saltzman, E. 2006. The role of vocal tract gestural action units in understanding the evolution of phonology. *In: Arbib M, editor. From action to language: The mirror neuron system. Cambridge: Cambridge University Press; 2006. pp. 215–249.*

[2] Goldstein, L., Nam, H., Saltzman, E., & Chitoran, I. 2009. Coupled oscillator planning model of speech timing and syllable structure. In G. Fant, H. Fujisaki & J. Shen (Eds.) *Frontiers in phonetics and speech science: Festschrift for Wu Zongji* (pp. 239–249). Beijing: Commercial Press.

[3] Nam, H., Goldstein, L., & Saltzman, E. 2009. Self-organization of syllable structure: A coupled oscillator model. *Approaches to phonological complexity* (pp. 297-328). Berlin/New York: Mouton de Gruyter.

[4] Saltzman, E., Nam, H., Krivokapić, J., & Goldstein, L. 2008. A task-dynamic toolkit for modeling the effects of prosodic structure on articulation. *Proceedings of the 4th International Conference on Speech Prosody (Speech Prosody 2008), Campinas, Brazil*, 175-184.

[5] Gaskell, M., Spinelli, E., and Meunier, F. 2002. *Perception of resyllabification in French. Memory and Cognition*, *3*0, 798–810.

[6] Spinelli, E., McQueen, J., and Cutler, A. 2003. Processing resyllabified words in French. *Journal of Memory and Language*, 48, 233–254.

[7] Rialland, A. 1986. Schwa and syllables in French. *Studies in compensatory lengthening*, *23*, 187.

[8] Fougeron, C., Bagou, O., Stefanuto, M., & Frauenfelder, U. H. 2003. Looking for acoustic cues of resyllabification in French. In *Proceedings of the 15th International Congress of Phonetic Sciences (ICPHS)*.

[9] Fougeron, C. 2007. Word boundaries and contrast neutralization in the case of enchaînement in French. *Papers in laboratory phonology IX: Change in phonology*, 609-642.

[10] Nguyen, N., Wauquier-Gravelines, S., Lancia, L., & Tuller, B. 2007. Detection of liaison consonants in speech processing in French: Experimental data and theoretical implications. *In Pilar Prieto. Segmental and Prosodic Issues in Romance Phonology, John Benjamins, pp.3-23, 2007, Current Issues in Linguistic Theory.*

[11] Cho, T., Yoon, Y., & Kim, S. 2014. Effects of prosodic boundary and syllable structure on the temporal realization of CV gestures. *Journal of Phonetics*, *44*, 96-109.

[12] Shaw, J. A., & Gafos A. I. 2015. Stochastic time models of syllable structure. *PloS one*, *10*(5), e0124714. doi:10.1371/journal.pone.0124714

[13] Shaw, J. A., Oh, S., Durvasula, K., & Kochetov, A. (2021). Articulatory coordination distinguishes complex segments from segment sequences. *Phonology*, *38*(3), 437-477. doi:10.1017/S0952675721000269.

[14] Kisler, T., Reichel, U., & Schiel, F. 2017. Multilingual processing of speech via web services. *Computer Speech & Language*, *45*, 326-347.

[15] G. V. Rossum and F. L. Drake. 2009. *Python 3: Reference manual*. United States: SohoBooks.

[16] Bürkner, P.-C. 2017. brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

[17] R Core Team. 2022. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

[18] Kohler, K.J. "Gestural Reorganization in Connected Speech: A Functional Viewpoint on 'Articulatory Phonology'" *Phonetica*, vol. 49, no. 3-4, 1992, pp. 205-211. https://doi.org/10.1159/000261916