Domain-initial effects on C-to-V and V-to-V coarticulation in French: A corpus-based study

Fanny Guitard-Ivent, Giuseppina Turco, Cécile Fougeron

To cite this version:

Fanny Guitard-Ivent, Giuseppina Turco, Cécile Fougeron. Domain-initial effects on C-to-V and V-to-V coarticulation in French: A corpus-based study. Journal of Phonetics, 2021, 87, pp.101057. 10.1016/j.wocn.2021.101057 . hal-03497631

Please cite as

Guitard-Ivent, F., Turco, G., Fougeron, C. (2021) Domain-initial effects on C-to-V and V-to-V coarticulation in French: A corpus-based study. *Journal of Phonetics*, 87, 101057, https://doi.org/10.1016/j.wocn.2021.101057

Domain-initial effects on C-to-V and V-to-V coarticulation in French: A corpus-based study.

Fanny Guitard-Ivent[1], Giuseppina Turco[2], Cécile Fougeron[1]
[1] Laboratoire de Phonétique et Phonologie, UMR 7018, CNRS/Sorbonne Nouvelle, Paris, France
{fanny.ivent, cecile.fougeron}@sorbonne-nouvelle.fr
[2] Laboratoire de Linguistique Formelle, UMR 7110, CNRS/Université de Paris-Paris Diderot, Paris, France
gturco@linguist.univ-paris-diderot.fr

Abstract

This corpus-based study investigates whether and how coarticulation is modulated domain-initially in French. Post-boundary IP-initial sequences were compared to the same sequence in word-medial position according to the composition and structural link between the elements of that sequence. We examined 4 cases of coarticulation: C-to-V coarticulation within syllable in #CV and #VC. sequences, across syllables in #V.C sequences, and anticipatory V_2-to-V_1 coarticulation in? #V_1.C(C)V_2 sequences. Coarticulation was determined by the amount of spectral change in the formant of the vowel /a/ according to its context: an alveolar vs. uvular C context for C-to-V coarticulation, and a high vs. low V context for V-to-V coarticulation. The analysis of 15,000 tokens of /a/ revealed that coarticulation can be modulated by prosody but this effect appeared to depend on the structural relationships (the relation of the sounds relative to each other and to the syllable structure) between the elements in the IP-initial sequence. A reduction of coarticulation of IP-initial sequence was only found for the two cases of anticipatory C-to-V coarticulation (VC and V.C sequences). Interestingly, this reduction of coarticulation was larger in the heterosyllabic V.C sequence. These results are discussed in terms of difference of the coupling between the elements in the sequence and the proximity of the vowel to the boundary. They suggest that coupling relations constrain prosodically induced variation.

Keywords: C-to-V, V-to-V, coarticulation, French, prosodic boundaries, syllable structure

1. Introduction

Domain-initial effects (DI effects) refers to phonetic change undergone by consonants or vowels when produced in initial position in a prosodic domain -- i.e. after a prosodic boundary, as compared to other positions. DI effect is proportional to the strength of the boundary such that the stronger the boundary, the stronger the effect on the initial segment (see Cho, 2011, 2016 among others).

Previous results have suggested than DI effects target contrastive phonetic properties of the domain-initial segments. Phonetic contrasts have been found to be maximized for segments realized after strong prosodic boundaries in two ways: by enhancing the phonetic properties making the domain-initial segments more distinct from its neighbors (syntagmatic contrast), and/or by enhancing the contrastive phonetic properties within a paradigm (paradigmatic contrast). For instance, Cho & Jun (2000) have shown that Korean initial consonants located in an Intonational Phrase (IP) are more distinct in the phonetic features marking contrast between stops (here VOT and airflow at release), such that the phonetic contrast between aspirated and fortis consonants is enhanced. Moreover, these IP-initial consonants, in Korean and in many other languages, are overall less sonorous, with a stronger oral constriction for oral and nasal stops, and with less nasal flow or nasal energy for nasals, such that they better contrast with their neighboring vowels (see, for instance, the following articulatory studies: Fougeron and Keating, 1997, Cho and Keating, 2009 for English; Fougeron 2001 for French). Similarly, DI-induced variations on vowels also provide evidence for enhancement of phonetic contrasts. For example, in French all front vowels are produced with a larger lip opening in IP-initial position, which makes them overall more sonorous. Nonetheless, the distinction in terms of lip opening between the rounded and unrounded counterparts is maximized in that position, due to a larger increase of lip opening for the unrounded vowel than for the rounded one (Georgeton & Fougeron, 2014). Acoustically, IP-initial French vowels are also found to be more peripheral compared to IP-medial vowels, which are more centralized (Georgeton 2014; Georgeton, Kojanćić-Antolik & Fougeron 2016).

These results converge towards the idea that DI effect modulates the distinctive phonetic properties of domain-initial segments making them more distinct in their phonetic form compared to when they are embedded within a prosodic domain, where they appear to be more reduced. However, we still do not know *how* these phonetic properties are modulated by DI effect.

An interesting approach, originally proposed by Byrd and Saltzman (2003), explains segmental variation at the vicinity of prosodic boundaries via the action of a prosodic (π)-gesture. This modulation gesture operates at prosodic boundaries in order to instantiate prosodic structure. The action of π-gestures is to slow down the temporal unfolding of (segmental constriction) gestures that are co-active with it. While the interval of activation of the π-gestures and its coordination with co-active gestures remain to be more precisely defined (in particular when other prosodic events such as prominence need to be instantiated -- see Byrd and Riggs 2008, Katsika 2016), it is proposed that its effect extends over an interval in both (i) a *progressive* manner, i.e. with a stronger effect on gestures located under its peak of activation and with a fading effect on the more distant gestures, and (ii) a *cumulative* manner, i.e. with an activation strength related to the strength of the prosodic boundary, such that the degree of modulation of the π-gestures on the other gestures depends on the boundary strength.

Empirical evidence supporting the π-gesture is provided through the lengthening of the segments around prosodic boundaries: the well-known acoustic lengthening of domain-final segments found in most of the world languages (see Krivokapić 2007, among others), but also the lengthening of articulatory movements reported for domain-initial segments in several languages, such as Korean (Cho & Keating 2001), English (Cho & Keating 2009), German (Bombien, Mooshammer, Hoole, Kühnert, 2010) or Greek (Katiska 2016). Byrd & Choi (2010), for instance, found a gradient lengthening of domain final CC# coda clusters and of initial #CC onset clusters according to the strength of the prosodic boundary (#). Specifically, they found that the lengthening was stronger for the C immediately adjacent to the boundary, i.e. the consonant located under the scope of the π-gesture's maximal level of activation, while the lengthening effect was smaller or absent for the consonants further away from the boundary (i.e. C1 in C1C2# and C2 in #C1C2) (see also Byrd, Krivokapic & Lee 2006).

Interestingly, in the case of French, the DI effect on the duration of domain-initial segments has always been questioned. No clear relationships have been reported between the observed non-temporal DI effects and lengthening. In Georgeton & Fougeron (2014), articulatory and spectral changes for IP-initial vowels were not systematically accompanied by an acoustic lengthening of the vowels. They were found to be either longer, shorter or of similar duration compared to IP-medial vowels according to vowel categories and speakers (see also Fougeron 2001 for similar results on consonants). Moreover, in studies comparing French to other languages, the pairing of temporal and spatial/spectral changes was always weaker in French: more acoustic lengthening of IP-initial vowels (compared to IP-medial) in German compared to French (Gendrot, Gerdes & Adda-Decker, 2011), stronger correlation between acoustic duration of IP-initial consonants and linguopalatal contact in Korean than in French (Keating, Cho & Fougeron, 2004).

One possible explanation for language-specific (and probably speaker-specific) consequences of DI effect on the presence or not of lengthening could be related to its consequences on the coordination between the gestures close to prosodic boundaries. In the π-gesture approach, for instance, if constriction gestures adjacent to prosodic boundaries can be lengthened by the slowing down of their clock, they are also said to unfold further apart, i.e. to be less overlapped. This idea of prosody acting on the temporal arrangement between speech units has been proposed since a while to account for the change in acoustic or articulatory properties of vowels in prominent syllables. In De Jong, Beckman and Edwards (1993), for instance, the larger jaw opening observed for pitch accented American English vowels (as compared to non-accented vowels) is explained by a smaller truncation of this opening gesture by the adjacent consonantal closing gesture compared to what is found for non-accented syllables. Less overlap between gestures in prominent position could also explain change in the degree of coarticulation (i.e. contextual variability) in the acoustic or articulatory signal. In a series of acoustic studies by Cho & Kim and colleagues, accented vowels under focus are found to be less nasalized by a neighboring nasal consonant in American and Australian English, in Korean and in Mandarin (Cho, Kim & Kim, 2017; Jang, Kim & Cho, 2018; Joo, Jang, Kim, Cho & Cutler, 2019; Li, Kim & Cho, 2019).

Unlike the prominence-related effects discussed above, it is less straightforward to relate boundary-related effects to a reduction of overlap between gestures. Some studies have reported changes in the coordination within a prosodic domain-initial #CV syllable or within a domain-initial #CC cluster, but results were not systematic. For instance, in an electropalatographic investigation on French #CV sequences (with # indicating either a syllable-, an accentual phrase- or an IP-boundary and C=/k, t, l/), Meynadier (2003) found no changes in C-to-V carryover coarticulation related to boundary type. In Byrd and Choi (2010), inter-gestural timing within #CC English sequences was also not modified across prosodic positions. In an other EMA study, Byrd (2000) analyzed inter-gestural timing for three speakers within the syllable #mi (with # indicating either a word, a minor, or a major prosodic boundary). While two speakers did not show an effect of boundary strength, one of the speakers showed a reduction of overlap at the major prosodic boundary, with a delayed, and therefore less overlapped, tongue body gesture for the /i/, relative to the labial gesture of the domain-initial /m/. More recently, in the above-mentioned acoustic studies on nasal C-to-V coarticulation by Cho and colleagues, carry-over nasalisation in NV syllables was found to reduce in IP-initial position in American and Australian English (Cho et al. 2017, Joo et al. 2019) and in Mandarin (Li, et al. 2019). Taken together, these studies showed that whenever the coordination between the elements following a prosodic boundary (i.e. initial syllable or initial cluster) is found to be sensitive to the strength of the boundary, it is in the direction of a reduction of their overlap (and consequently, a reduction in the degree of coarticulation or contextual variability in the signal).

In the present study, our objective is to question further how prosodic phrasal organization modulates the coordination between speech gestures at the beginning of strong prosodic domains. More specifically, we examine **whether DI effects on overlap/coarticulation depend on the structural relationship between the elements *following* the prosodic boundary**. By 'structural relationship' we mean the relation of the sounds relative to each other and to the syllable structure.

We extend here Byrd and Choi (2010)'s idea that DI effects on overlap may depend on the specified coordination patterns of gestures forming an onset cluster. They explain why /s+stop/ clusters in English, known to have a characteristic and stable pattern of coordination, are less sensitive to prosodic variation, as compared to /kl/ onset clusters. In our study, DI-related changes in overlap will be assessed indirectly by looking at the amount of contextual coarticulation in the acoustic signal, in sequences where the structural links between the elements can be modeled in terms of different coordination specifications. Browman & Goldstein (2000) and Nam, Goldstein and Saltzman (2009) have proposed that syllable structure relates to specific coupling modes between gestures: the onset-nucleus intergestural relationship is modeled by an in-phase coupling, allowing a synchronous initiation of the C onset and V nucleus gestures and this mode of coupling is said to be strong and stable. In contrast, gestures related to a nucleus and a coda, or related to two consonants in a cluster, are coupled anti-phase (are 180° out of phase) and thus are initiated sequentially. The sequentiality of this anti-phase mode, as well as the potential competition between coupled gestures (e.g. between a complex onset and the following vowel), makes the temporal relationship between these anti-phased coupled gestures more variable.

In the current study, we investigate whether DI effects modulate coarticulation in a domain-initial sequence according to the structural links between the elements of that sequence. Since these links are supposed to be instantiated in terms of coordination between the elements, we thereby test whether different coordination patterns are sensitive to prosodically-driven variation. Acoustic consequences of coarticulation are examined in four different contexts. Cases of coarticulation within syllables are observed in #CV sequences (carryover C-toV coarticulation) and #VC. sequences (anticipatory C-to-V coarticulation) (with '.' indicating a syllable boundary and '#' as a place holder for the prosodic boundary we will manipulate). Cases of coarticulation across syllable boundaries are observed in #V.C and #V_1.C(C)V_2 sequences: anticipatory C-to-V coarticulation in #V.C sequences, and anticipatory V_2-to-V_1 coarticulation in #V_1.C(C)V_2 sequences.

In the in-phase tautosyllabic #CV sequence, overlap is expected to be more stable. Therefore, acoustic consequences of C-to-V carry-over coarticulation are expected to be weakly influenced by DI effects: changes in V acoustics which can be attributed to the type of C onset should be similar across prosodic positions. This condition is similar to the one used in Cho and colleagues' studies (Cho et al. 2017, Joo et al. 2019, Li, et al. 2019) on N-to-V acoustic nasalization; however, here we look at the overlap between potentially competing gestures linked to the lingual articulation of the elements. In #VC syllables, C-to-V coarticulation is predicted to be more variable (i.e. variable across prosodic position), since coda C gestures are coupled with the nucleus gestures in a more flexible anti-phase mode. In the heterosyllabic cases, we predict that the temporal organization between the elements should be even more flexible, although the specific coupling relations at play across a syllable boundary are yet little understood (Turvey, 1990). Consequently, the acoustic effects of C-to-V anticipatory coarticulation in #V.C sequences and V-to-V anticipatory coarticulation in #V_1.C(C)V_2 sequences are expected to be more flexible and potentially more sensitive to DI effects.

2. Method

2.1 Speech material and acoustic measures

In order to test how coarticulation is modulated by prosodic position on a large set of data, the first methodological choice made is to work on naturalistic speech produced by a large number of speakers. The speech material examined comes from two publicly available French corpora: ESTER (Gravier et al., 2006) and NCCFr (Torreira et al., 2010). ESTER corpus is based on broadcasted news speech; NCCFr represents conversational speech from friends' (guided and free) discussions on several social topics. The two corpora have been automatically aligned (see Gauvain et al., 2002 for the system used on NCCFr and Galliano et al., 2005 for ESTER) and the selection of the material was done based on these alignments.

2.1.1 Selection of the test sequences

Four types of sequences are extracted in these corpora: (i) tautosyllabic /CV_1/, (ii) tautosyllabic /V_1C./, (iii) heterosyllabic /V_1.C/, and (iv) /V_1.C(C)V_2/ (with '.' marking a syllable boundary). V_1 is the vowel on which coarticulation is measured and this vowel is always /a/. In C-to-V_1 coarticulation, V_1 spectral properties are influenced by the place of articulation of C; in V_2-to-V_1 coarticulation, V_1 spectral properties are influenced by V_2's height.

The analysis of C-to-V coarticulation (in CV and VC sequences) is based on 7,000 tokens of the vowel /a/. Selection of these tokens was done according to the following criteria: V_1 had to be adjacent to either an alveolar (C_{ALV}=/t, d, z, s, l, n/, e.g. *dépanner* /dep**a**ne/ - "to help") or a uvular (C_{UV}=/ʁ/, e.g. *app**a**reil* /apaʁɛj / - "device") consonant. The other adjacent consonant in the opposite context (left for V_1C and right for CV_1), if any, had to be a labial consonant (among /p, b, f, v, m/). Once all VC sequences candidates were extracted, a further division was made between tautosyllabic

V_1C. sequences (e.g. *arbitre* /aʁ.bitʁ/ - “referee”) and heterosyllabic $V_1.C$ sequences (e.g. *arrange* /a.ʁɑ̃z/ - he arranges).

The analysis of V_2-to-V_1 coarticulation is based on 8,000 tokens of /a/ vowel. This vowel was chosen because its high frequency in the corpus and because it occurs in all the tested conditions. We selected words that include a $V_1C(C)V_2$ sequences with V_2 being either a high vowel (VHIGH=/i, y, u/, e.g. ***a**mi* /ami/ - “friend”) or a low vowel (VLOW=/a, ɑ̃/, e.g. *également* /eg**a**lmɑ̃/ - “also”), and with 1-to-3 consonants between V_2 and V_1.

C-to-V coarticulation is captured in terms of acoustic change between the two C contexts. As illustrated in Figure 1a and 1b, the contextual influence of the alveolar consonant on the open vowel /a/ should manifest itself as an attraction of F2 towards a locus of 1800Hz and as a restriction of the increase of F1 due to the impedance of the consonant on the opening of the vocal tract. Conversely, the coarticulatory influence of the uvular consonant should surface as a lowering of F2 and a raising of F1. V_2-to-V_1 anticipatory coarticulation is captured as a decrease of F1 when the aperture of /a/ is reduced by a following V_2HIGH (figure 1c), as compared to when V_2 is LOW (figure 1d). Note that we decided not to examine V-to-V coarticulation on F2 because the anteriority/posteriority of V_2 was not balanced (the V_2HIGH context is mainly composed of anterior vowels).

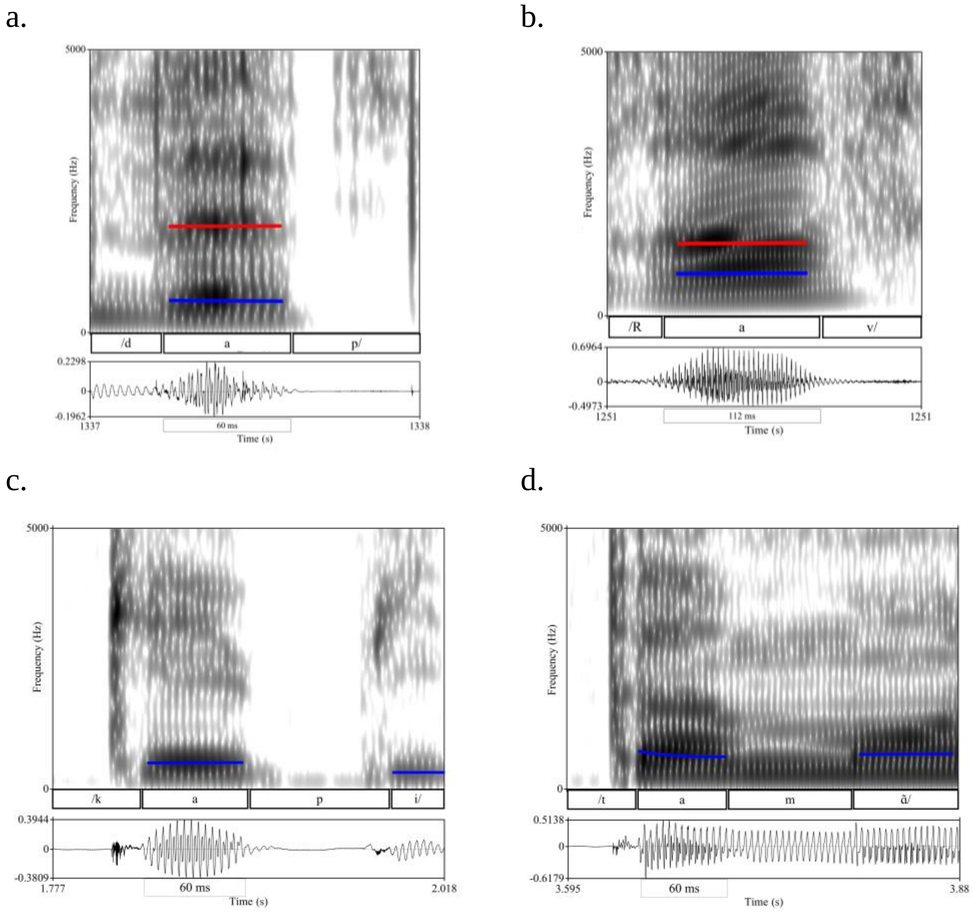


Figure 1. Spectrograms illustrating contextual influences on the formant values of the V_1 /a/ vowel. Examples of carryover C-to-V coarticulation in CV_1 sequences with (a) an alveolar context in /dap/ extracted from the word *adapter* “to adapt”, and (b) a uvular context in /ʁav/

extracted from the word *grave* “serious”). The opposite C (the right consonant) is always a labial. Examples of anticipatory V_2-to-V_1 coarticulation with (c) a V_2HIGH context in /kapi/ extracted from the word *capital* “capital”, and (d) a V_2LOW context in /tamɑ̃/ extracted from the word *notamment* “especially”.

2.1.2 Selection of prosodic positions

Coarticulation within the four sequences is compared in two prosodic positions: the selected sequence occurs either (i) at the beginning of a word in an IP-initial position and (ii) in a medial position of a word (non-initial, non-final) and therefore also in an IP-internal position (Wm). Figure 2 illustrates the criteria defining the two prosodic positions. With this definition, the selected tokens have the test vowel V_1 in strictly initial position in the IP for the $V_1C(C)V_2$ sequence and the $V_1(.)C$ sequences, but in the CV_1 sequence, V_1 is further away from the boundary since it is the C that is strictly adjacent to the boundary.

Prosodic phrasing and IP boundaries have been determined using automatable methods adapted to large corpora analysis. As done in previous studies on the same corpus (Gendrot, Gerdes and Adda-Decker, 2011), the presence of a pause was considered as a cue to IP boundaries, and thus post-pausal elements were considered as IP-initial. Pauses were detected on the automatic alignment of the corpus, with no minimum threshold. In the selected material, pauses have an average of 80 ms (314 ms standard deviation). Although this method is easily applicable on such a large corpus, it has the disadvantage of underestimating the number of IP boundaries since it does not capture IPs not delimited by a pause. For this reason, the second prosodic position (Wm) was selected in the middle of a word so to ensure that the sequence is not initial or final in those IPs not delimited by a pause.

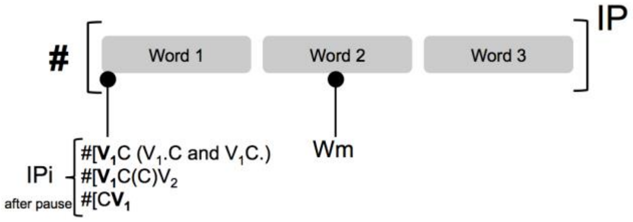


Figure 2. Description of the two tested prosodic positions IPi (IP-initial) and Wm (word-medial) according to the type of sequence: V_1C, CV_1 and $V_1C(C)V_2$. The “#” symbol represents the presence of a pause; ‘.’ represents the presence of a syllable boundary; V_1 in bold is the test vowel on which coarticulation is measured.

Table 1 gives an overview of the distribution of the selected speech material, per sequence types, prosodic positions and coarticulatory contexts.

	CV_1		V_1C				$V_1C(C)V_2$	
			$V_1.C$		$V_1C.$			
	ALV	UV	ALV	UV	ALV	UV	HIGH	LOW
IPi	148	56	1526	134	112	199	100	480
Wm	1454	435	1697	769	55	428	3705	3694
total	*1602*	*491*	*3223*	*903*	*311*	*483*	*3805*	*4174*
N of speakers	*371*		*487*				*540*	

Table 1. Number of /a/ token per condition: for each sequence types (CV_1, $V_1.C$, $V_1C.$ and $V_1C(C)V_2$) as a function of prosodic position and place of articulation of the adjacent consonant (alveolar (ALV) *vs.* uvular (UV)), for CV_1, $V_1.C$ and $V_1C.$) or V_2 height of (HIGH *vs.* LOW, for $V_1C(C)V_2$). The last row indicates the number of different speakers from which the material has been selected.

2.2 Acoustic measures and statistical analysis

F1 and F2 of the 15,000 V_1 /a/ were extracted from the test sequences using the Praat software's Burg algorithm (Boersma & Weenink, 2014) with standard settings (5 formants between 0 and 5kHz for male speakers, and 0 and 5.5kHz for female speakers). Formants were measured in consecutive 5ms frames, and we computed a single formant value per token by averaging the values taken at 1/3, 1/2 and 2/3 of the vowel duration in order to compensate for probable imprecision on the vowel onset and offset in the automatic alignment of the vowels. An automated procedure was further applied to filter out outliers by applying the threshold defined by Gendrot & Adda-Decker (2005) for French spontaneous speech (F1/a/ <1000 Hz or <1100 Hz and F2/a/ 800 Hz< >2300Hz or 900Hz< >2300Hz for male and female speakers respectively). Following Traunmüller (1997), formants were Bark-transformed as a way to attenuate speaker and sex related differences since by speaker normalization was not feasible (not all speakers have the same amount of vowels nor are they represented in all subsets).

The interaction between prosodic position and coarticulation was tested on each of the four sequences and for each formant (F1 and F2 for C-to-V coarticulation and F1 only for V-to-V coarticulation) using linear mixed-effects models in the R software (R development Core Team, 2008, with the “lme4” (Bates et al., 2015) and “afex” packages (Singmann et al., 2020).

For the cases with C-to-V carry-over coarticulation (CV_1 sequences), the models tested the relationship between formant values (Fn: F1 and F2) of the target V_1 as function of the CONSONANTAL CONTEXT (C: alveolar *vs.* uvular), PROSODIC POSITION (IPi *vs.* Wm), and the interaction between these factors. For the cases with V_2-to-V_1, the same model was tested on the F1 of V1, but with V_2 HEIGHT (high *vs.* low) instead of C context. For the cases with to C-to-V anticipatory coarticulation, the models were different since they included the relationship between formant values (F1, F2 in Bark) of the target V_1 in relation to SYLLABIC STRUCTURE (V.C *vs.* VC.), besides the adjacent CONSONANTAL CONTEXT (C: alveolar *vs.* uvular) and PROSODIC POSITION (IPi *vs.* Wm). In all models, by-speaker random intercepts and slopes were included (see the R-syntax in the Appendix section).

P-value estimates for linear regressions were based on *Satterthwaite* approximations through the *lmerTest*()-function (Kuznetsova *et al.,* 2013). The threshold was set to $p<.05$. Likelihood ratio tests as implemented in the *anova*()-function were performed to check main effects of each fixed factor and interactions. Finally, R^2 values associated to each model were calculated by using the *r.squaredGLMM*()-function within the library ‘MuMIn’ (Bartón, 2014). A posteriori analysis of contrasts was performed using *lsmeans*()-function from the library “emmeans” (Lenth et al., 2018) with Tukey’s *p*-value adjustments.

3 Results

3.1 DI effects on anticipatory C-to-V coarticulation according to syllabic structure in V_1C. & V_1.C sequences

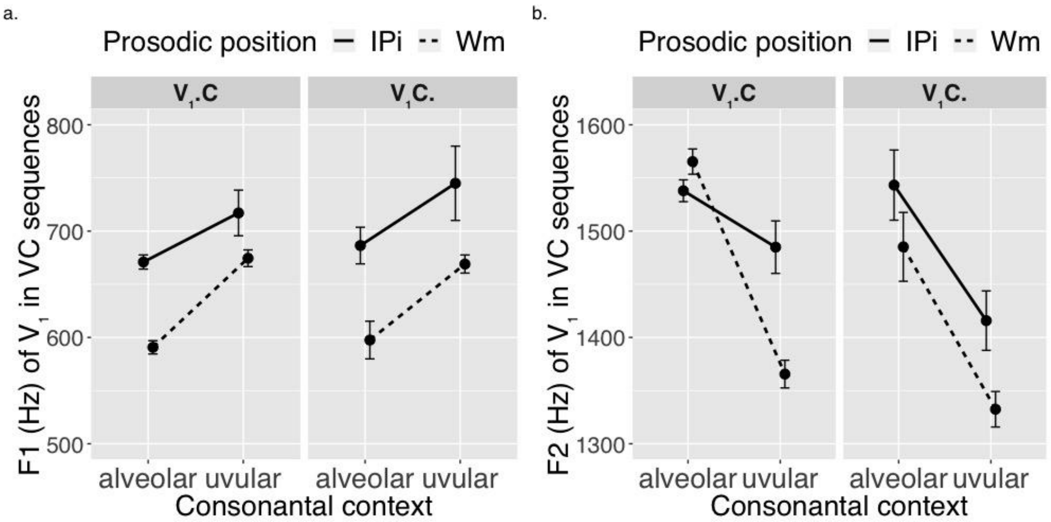


Figure 3(a)-(b). Effects of CONSONANTAL CONTEXT (alveolar, uvular) on F1 (a) and F2 (b) of V_1 in V_1.C (left panels) and V_1.C (right panels) sequences as a function of PROSODIC POSITION (IP initial – “IPi”, Word medial – “Wm”). In the plots, F1 and F2 are expressed in Hz (y-axis). Statistical analyses are performed on bark transformed values.

Figures 3(a)-(b) show F1 and F2 of V_1 as a function of consonantal context, prosodic position and syllabic structure (see also plots for each factor in the Appendix). The output of the models for F1 and F2 are presented in Table 2.

First, as expected, spectral changes are observed according to CONSONANTAL CONTEXT in all prosodic positions: F1 of /a/ is raised and F2 is lowered by the uvular - compared to the alveolar - context. Second, an effect of PROSODIC POSITION is found, in interaction with SYLLABLE STRUCTURE: in VC and V.C sequence the F1 of /a/ is higher in IP-initial position compared to Wm position, and this difference is larger in the VC. sequence. F2 is also higher in IP-initial position in VC. but not in V.C sequences (due to the alveolar context).

Crucially, an interaction between CONSONANTAL CONTEXT and PROSODIC POSITION is found for both F1 and F2. The amount of spectral change between the alveolar and uvular-context is found to reduce in IP-initial position for both heterosyllabic #V.C and tautosyllabic #VC.

Finally, results show an interaction between CONSONANTAL CONTEXT, PROSODIC POSITION and SYLLABLE STRUCTURE for both F1 and F2. Post-hoc analyses of three-way contrasts reveal that the reduction of coarticulation in IP-initial position is larger in the heterosyllabic #V.C sequence. As Figures 3(a)-(b) illustrate, the amount of spectral change according to CONSONANTAL CONTEXT (i.e. the raising of F1 by the uvular context and the lowering of F2 by the uvular context) in IPi position is smaller in the heterosyllabic (V.C) condition than in the tautosyllabic (VC.) position (for F1: β= -0.41 vs. β= -0.69 all *p*-values <.0001; for F2: β= -0.40 *vs.* β= 0.73 all *p*-values <.0001). However, the opposite is observed in Wm position, in which a greater amount of spectral change according to the two contexts is found in heterosyllabic (VC.) position compared to tautosyllabic (V.C) position (for F1: β= -0.93 vs. β= -0.60 all *p*-values <.0001; for F2: β= 1.29 vs. β= 0.89 all *p*-values <.0001).

Main effects & interactions	χ^2	**Effect on F1 of /a/**	β	**SE**	**t**
CTXT	209.8^{***}	ALV < UV	-0.66	0.04	-17.8
PP	205.1^{***}	IPI > WM	0.71	0.04	15.6
SYLL	30^{***}	VC. < V.C	-0.24	0.04	-5.62
CTXT*SYL	ns.	CTXT EFFECT: VC. = V.C	--	--	--
PP*SYL	6^{*}	PP EFFECT: VC. > V.C	-0.18	0.07	-2.46
CTXT*PP	**6.4^{*}**	**CTXT EFFECT: IPI > WM**	**0.21**	**0.08**	**2.56**
CTXT*PP*SYLL	**21.3^{***}**	**CTXT*PP EFFECT: V.C > VC.**	**0.61**	**0.13**	**4.64**
Main effects & interactions	χ^2	**Effect on F2 of /a/**	β	**SE**	**t**
CTXT	265.7^{***}	ALV > UV	0.80	0.04	22.2
PP	98.6^{***}	IPI > WM	0.40	0.04	10.6
SYLL	49.4^{***}	V.C < VC.	0.28	0.04	7.30
CTXT*SYL	ns.	CTXT EFFECT: VC. = V.C	--	--	--
PP*SYL	13.68^{***}	PP EFFECT: V.C > VC.	-0.27	0.07	-3.83
CTXT*PP	**41.70^{***}**	**CTXT EFFECT: IPI > WM**	**-0.52**	**0.07**	**-6.77**
CTXT*PP*SYLL	**41.72^{***}**	**CTXT*PP EFFECT: V.C > VC.**	**0.73**	**0.14**	**5.38**

Table 2. Likelihood ratio test of linear mixed effects models testing main effects of the predictors CONSONANTAL CONTEXT (CTXT), PROSODIC POSITION (PP) and SYLLABIC STRUCTURE (SYLL) and their interactions on F1 and F2 of /a/ in Bark. χ_2 values (with one degree of freedom) are reported as effect size estimates (p<.05 = *, p<.001 = **, p<.0001 = ***). "ALV" and "UV" stand for alveolar and uvular, "IPi" and "Wm" for initial position of Intonational phase and word medial position, "V.C" heterosyllabic and "VC." Tautosyllabic sequences. ALV < UV, for instance, means that F1 of V_1 is lower when V_1 is in alveolar context than in uvular. Estimates: β-coefficients, standard errors (SE) and t-values of the model are also reported. In bold the crucial interactions showing the effect of prosodic position on coarticulation.

3.2 DI-induced effects on carry-over C-to-V coarticulation (CV_1 sequences)

Figure 4(a)-(b) illustrates F1 and F2 of /a/ as a function of CONSONANTAL CONTEXT and PROSODIC POSITION in CV sequences. The output of the models for F1 and F2 are presented in Table 4.

The contextual effects of the uvular and alveolar consonants on F1 and F2 of /a/ are found in the expected direction. Next, there is an effect of prosodic position on the F1 (but not on the F2) of /a/. However, the difference in Hz between IPi and Wm position is much smaller in the CV sequence than that one observed in the VC sequence (compare Figure 3a and 4a).

Crucially, there is no interaction between CONSONANTAL CONTEXT and PROSODIC POSITION (see output of the models in Table 4): coarticulation is not reduced in IPi. Changes in F1 and F2 of /a/ as induced by the preceding consonant (i.e. raising of F1 and lowering of F2 when preceded by an uvular consonant) are constant across the two prosodic positions.

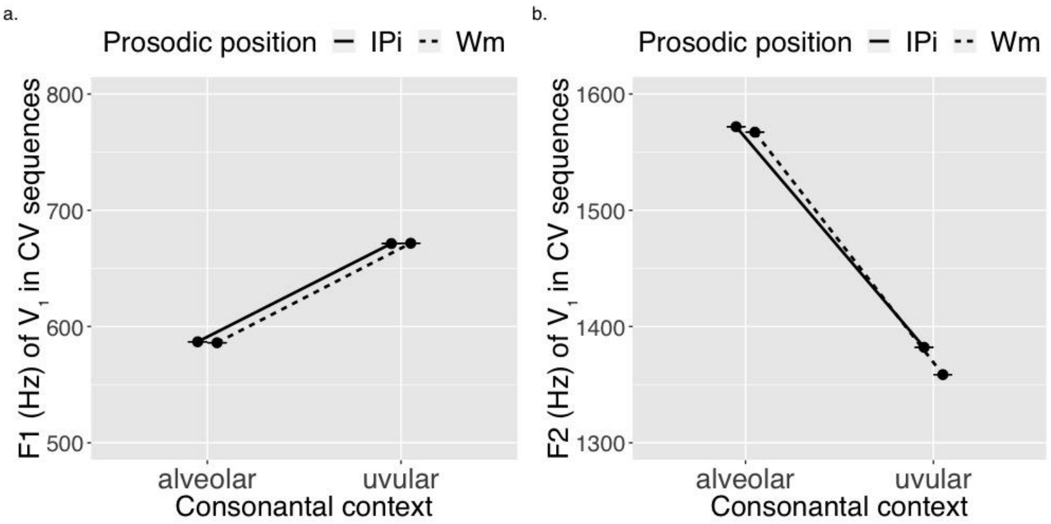


Figure 4(a)-(b). F1 (a) and F2(b) of V_1 as a function of CONSONANTAL CONTEXT (alveolar, uvular) and PROSODIC POSITION (IP initial – "IPi", Word medial – "Wm") in the CV_1 sequences. In the plots, F1 and F2 are expressed in Hz (y-axis). Statistical analyses are performed on bark transformed values.

F of /a/	**Main effects & interactions**	χ^2	**Effect on formants**	***β***	**SE**	**t**
F1	CTXT	222.7^{***}	ALV < UV	-1.06	0.04	-22.9
	PP	6.46^{*}	IPI > WM	0.15	0.06	2.60
	CTXT*PP	$\mathbf{0.27^{ns.}}$	**CTXT EFFECT: IPI = WM**	--	--	--
F2	CTXT	250.8^{***}	ALV > UV	1.41	0.05	26.64
	PP	$0.34^{ns.}$	IPI = WM	-0.36	0.06	-0.59
	CTXT*PP	$\mathbf{1.74^{ns.}}$	**CTXT EFFECT: IPI = WM**	--	--	--

Table 4. Likelihood ratio test of linear mixed effects models testing main effects of the predictors CONSONANTAL CONTEXT (CTXT), PROSODIC POSITION (PP) and their interactions on F1 and F2 of /a/ in Bark. χ_2 values (with one degree of freedom) are reported as effect size estimates ($p<.05 = ^{*}$, $p<.001 = ^{**}$, $p<.0001 = ^{***}$). "ALV" and "UV" stand for alveolar and uvular, "IPi" and "Wm" for initial position of Intonational phase and word medial position. For F2 for instance, ALV > UV, means that F2 of V_1 is higher when V_1 is in alveolar context than in uvular. Estimates: *β*-coefficients, standard errors (SE) and *t*-values of the model are also reported. In bold the crucial interactions showing the effect of prosodic position on coarticulation.

3.3 DI-induced effects on anticipatory V-to-V coarticulation (V_1CCV_2 sequences)

Figure 5 shows F1 changes of /a/ according to V2 HEIGHT (high vs. low) in the two prosodic positions. As expected (see Table 5), the first formant of /a/ raises (i.e. /a/ is more open) when followed by a low V_2, indicating the presence of anticipatory coarticulation. Moreover, there is an effect of PROSODIC POSITION: F1 gets significantly higher in IPi compared to Word medial position. Crucially, like for C-to-V, there is no interaction between V_2 HEIGHT and PROSODIC POSITION: V_2-to-V_1 coarticulation is not modulated by prosodic position.

Main effects & interactions	χ^2	**Effect on F1**	***β***	**SE**	**t**
V_2	205.7^{***}	HIGH < LOW	-0.41	0.05	-7.50
PP	168.6^{***}	IPI > WM	0.91	0.06	15.4
V_2*PP	**ns.**	**V_2 EFFECT: IPI= WM**	--	--	--

Table 5. Likelihood ratio test of linear mixed effects models testing main effects of the predictors V_2 HEIGHT (V_2), PROSODIC POSITION (PP) and their interactions on F1 of /a/ in Bark. χ_2 values (with one degree of freedom) are reported as effect size estimates ($p<.05$ = *, $p<.001$ = **, $p<.0001$ = ***). "IPi" and "Wm" for initial position of Intonational phase and word medial position. High<Low, for instance, means that F1 of V1 is lower when followed (V_2) by a high vowel than by a low vowel. Estimates: *β*-coefficients, standard errors (SE) and *t*-values of the model are also reported. In bold the crucial interactions showing the effect of prosodic position on coarticulation.

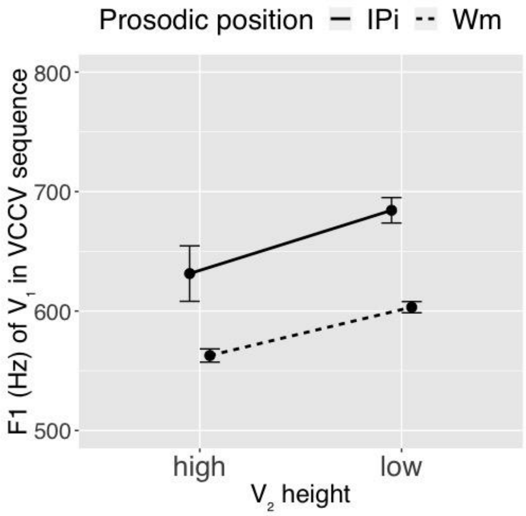


Figure 5. F1 of V_1 in $V_1C(C)V_2$ sequences as a function of V2HEIGHT (HIGH, LOW) and PROSODIC POSITION (IP initial – "IPi", Word medial – "Wm"). In the plots, F1 and F2 are expressed in Hz (y-axis). Statistical analyses are performed on bark transformed values.

4 Summary and general discussion

The current corpus-based study investigated whether and how coarticulation is modulated by prosodically induced variation occurring domain-initially. Coarticulation was examined acoustically as spectral changes in the formant of the vowel /a/ according to its context: an alveolar vs. uvular C context for C-to-V coarticulation, and a high vs. low vowel context for V-to-V coarticulation. As expected, spectral changes were observed according to the context in all prosodic positions: F1 of /a/ was raised and F2 was lowered in a uvular context, and F1 of /a/ was lowered by a following high vowel. More interestingly, both C-to-V coarticulation within a CV syllable and V-to-V coarticulation were similar in IP-initial and word-medial positions: i.e. they were not modulated by prosodic position. On the contrary, for C-to-V coarticulation in V.C and VC. sequences, the amount of spectral change between the alveolar and uvular context was found to be reduced in IP-initial position. This is interpreted as a reduction of coarticulation in IP-initial position. Interestingly, there was more reduction of coarticulation in an IP-initial heterosyllabic #V.C sequence (i.e. across a syllable boundary) compared to an IP-initial tautosyllabic #VC. sequence. Specifically, in #V.C, F1 and F2 of the IP-initial /a/ were less affected by the following uvular consonant: F1 was raised but to a smaller extent than it was in W-medial position (or than in IP-initial position in a tautosyllabic /aR/ sequence) and its F2 remained quite high compared to all other conditions (cf. Figure 1 and 2).

Overall, these results show that coarticulation can be modulated by prosody but this effect appears to depend on the structural relationships between the elements in the IP-initial sequence.

Previous studies on the effect of prosody on coarticulation (cf. Introduction) have reported a large amount of variability in the results, which has always put into question the systematicity and validity of the observations. For instance, Bombien et al. (2010) looked at articulatory overlap in #CC onset clusters in German (with # = word, minor or major boundary). Only four out of seven speakers reduced the overlap in /kn/ at the beginning of the higher constituent, while three (same or other) speakers did so for /kl/. In English, only two of the three speakers of Byrd and Choi (2010) showed less overlap in /sC/ clusters when initial in the tested highest constituent. Most of these articulatory studies have been limited to a restricted data set, due to feasibility reasons. One considerable advantage of a study based on acoustic data is that there is less restriction in terms of data collection constraints, and results can be obtained on a large set of materials and/or speakers (for instance see Cho et al. 2017, and follow-up studies). In the present study, we went a step further in this methodological consideration by examining a very large amount of data: the acoustic realization of more than 15,000 /a/ vowels, produced by a large cohort of 660 speakers and in a naturalistic speech setting. In the corpora used, the productions were extracted from broadcasted speech, which included prompted journalistic speech but also debates and radio shows, and casual face-to-face speech with free and guided discussions between friends. As in most large corpus-based study, it is assumed that the lack of control of the material or the speakers, and the possible unbalance between some categories, is compensated by the large amount of tokens. Again, this is what this study has shown.

The current results provide further support to the fact that prosodic structure does modulate the phonetic realization of the segmental content in the expected way: compared to the word-medial position, /a/ was found to be more open and sonorous with a higher F1, and a somewhat higher F2, when directly following an IPi boundary (#VC) or when placed in the IP-initial #CV syllable. More interestingly, studying this huge amount of naturalistically produced tokens allowed us to confirm that prosodic position modulates coarticulation between segments. In the present study, we have limited our investigation to the vowel /a/ since such a vowel is the most frequent in the corpus and testable in all conditions. A previous study, Guitard-Ivent (2018) also looked at other French vowels but in a limited set of conditions according to the tokens available in the corpus. C-to-V coarticulation with the vowels /i, e, ɔ/ could be tested only in VC sequences (merging both V.C and VC) in IP initial and W-medial position, following the same method as in this study. Like the /a/ vowel, also for these vowels, coarticulation was reduced in IPi: there was less change according to the uvular vs. alveolar context on F1 and F2 for /i, e/, and on F2 for /ɔ/. The reduction of coarticulation in IP initial VC sequences is therefore not limited to /a/.

One major drawback when looking at the manifestation of coarticulation on the acoustic output is that we are looking at the end point of a process which can result from vary different articulatory strategies (to which we are blind), such as blending, truncation, reduction, overlap. The remaining of this discussion will therefore be somewhat speculative when results are interpreted in terms of coordination of articulatory gestures. Further studies with direct observation of articulatory data will therefore be needed in order to confirm our hypotheses.

The main question to be addressed in this section is why coarticulation is modulated in some sequences (VC. and V.C) and not in others (CV and VC(C)V). Several possible interpretations are discussed below[1].

---

[1] It is also possible that the accentuation of V1 may explain some of the differences observed between our sequences (VC, CV and VCCV), however our data did not allow us to test this hypothesis. Well-known for its final accentuation, French also has an optional initial accent which can be realized for emphatic or rhythmic reasons at the beginning of an Accentual Phrases (AP; Jun

(a) *C-to-V coarticulation in CV sequences is not modulated by prosody because V is further away from the prosodic boundary.*

The CV and VC (either V.C or VC.) sequences differ with respect to the proximity of the vowel to the prosodic boundary in the IP-initial condition. In the #CV case, the vowel is part of the domain-initial syllable, but it is not immediately adjacent to the boundary (the consonant is post-pausal), whereas in #VC cases, the vowel is strictly initial. Depending on the scope of the DI effect and assuming that the DI effect is stronger when closer to the edge of the prosodic boundary, as predicted by the π-gesture approach (Byrd et al. 2006), one can expect a stronger DI effect on V when in VC sequences compared to CV sequences. This is indeed the case in our data: vowels are less coarticulated when they are closer to the IP boundary in VC sequence but not in the CV initial syllables. In that sense, DI effect translates as a greater resistance of the post-boundary vowel to the contextual influence of the following consonant, i.e. to the overlap of the upcoming C gesture.

The observation of the duration of the vowels brings a further argument for the consideration of the DI effect as an edge effect not affecting the distant vowel in the initial CV syllable. Indeed, while vowels directly following the IP boundary in all #VC., #V.C, #VC(C)V conditions were longer compared to W-medial vowels, this lengthening did not affect vowels in the CV syllables. This observation challenges our prior results showing that domain-initial lengthening was not systematic in French (see discussion in the Introduction). Here, IP-initial lengthening of vowels is attested in a wider dataset where speaker-specific patterns, which may have influenced previous results, are leveled out. This finding provides further arguments for the π-gesture approach, according to which, the unfolding of co-active gestures close to the boundary are predicted to slow down.

While domain-initial lengthening allows more time for the IP-initial vowel to achieve its target, differences in coarticulation between the tested sequences may not just be a matter of lengthening. First, IP-initial vowels in the least coarticulated V.C sequence were not longer than that in the more coarticulated VC sequence. Second, in a prior study (Guitard-Ivent, 2018) a subset of the tokens was analyzed with a control of the vowel duration in order to rule out this potential co-varying factor. Around 1,000 of /a/ tokens in VC sequences (V.C and VC. taken together) were chosen so to have a balanced set of tokens ranging from 70 to 80ms duration in both the IP and Wm categories. Coarticulation was again found to decrease in IP-initial compared to Wm position in such a set of tokens controlled for duration.

Nonetheless, it is not the case that all vowels adjacent to the IP boundary are more resistant to coarticulation. Indeed, vowels strictly adjacent to the boundary are less coarticulated with the following consonant in the VC sequences, but not with the following vowel V_2 in the $V_1C(C)V_2$ sequences. In the latter case, V_1 was influenced by V_2 to the same extent (i.e. same lowering of F1 with a high V_2), whether it is IP-initial or word-medial, and also whether it is domain-initially lengthened or not. In other

---

& Fougeron 2002). Interestingly, its location is quite flexible on one of the first three syllables of the AP according to various factors, such as the segmental composition of the first syllable. While the realization of the initial accent is usually preferred on an onset-full syllable (*arrange toi* /a.ˋ**ʁɑ̃z** twa/ - “arrange yourself”; Pasdeloup 1990), it can also fall on the onset-less first syllable (/ˋ**a**.ʁɑ̃z twa/) especially when it is in a strong initial prosodic position (such as an IPi) where it is often accompanied by glottalization at the onset of the vowel (/ˋ**ʔa**.ʁɑ̃z twa/). An interaction between boundary effects and prominence effects has been shown in several other languages, such as English, Korean or Greek (Byrd & Riggs 2008, Cho 2001, Krivokapić 2014, Katsika 2016). It is unfortunately impossible to assess on such a large data set whether the token selected in one or the other conditions were more frequently accented on their first syllable, but this should be kept in mind for a more controlled experimental setting.

terms, it was not more resistant to coarticulation even though directly under the scope of the boundary effect. This difference between C-to-V and V-to-V anticipatory coarticulation effect on vowels adjacent to IP boundary suggests that the coordination of the elements in the sequence may play a role, as will be discussed below.

(b) *C-to-V coarticulation in CV sequences not modulated by prosody because its coupling is less flexible.*

As mentioned in the Introduction, the elements in the CV syllable are coupled in-phase so that the tongue body gesture for /a/ is initiated synchronously with that of C (see e.g. Goldstein, Nam, Saltzman & Chitoran, 2009). As a result, the gestures in the CV sequence overlap quite extensively. In our acoustic data, this is shown by the large contextual effect of the uvular /R/ on /a/'s F1 and F2, both in IPi and in Wm for the CV sequence, as we can see in Figure 2 compared to the cases with less coarticulation (in IPi) in Figure 1. The cohesion between the onset C and the nucleus V in the syllable thus appears as a tightly coordinated sequence, where there seems to be little flexibility available for prosodically induced variation.

A larger amount of coarticulation in CV vs. VC sequences has been reported in many studies (among others: Kuehn and Moll 1976; Gay 1981; Krull, 1989, Sussman, Bessell, Dalston, and Majors, 1997). However, in our results, coarticulation for the VC. and V.C sequences (Figure 1) seemed as extensive as in the CV case, when they are word-medial. In other words, there was less coarticulation in a VC than in a CV sequence only if it was in IP-initial position. In the case of the tauto-syllabic VC. sequence, the coda consonant is said to be coupled anti-phase with the nucleus and therefore is sequentially initiated after the vowel. As such, this mode of coordination is considered to be less stable than the in-phase CV coordination, and we indeed see that coarticulation is more variable from one prosodic condition to the other.

In the heterosyllabic V.C sequence, we do not know whether and how the two elements are coordinated with each other. It seems reasonable to assume that within a lexical word, syllables are somehow glued together to maintain within-word cohesiveness; however, this coordination has not been explicitly modeled yet. For instance, in their discussion of the relationship between adjacent consonants across syllables, Goldstein and Pouplier (2014) suggest that cross-syllable coupling between consonants may be required in some languages, but not in others. A stronger reduction of coarticulation in the IP-initial V.C sequence compared to the VC sequence could result from the fact that V and C are indeed not coupled together, or because their coupling allows variation in overlap with the preceding syllable.

It is also the case that in V.C sequences, C is coupled with the element in its own syllable, and this tautosyllabic link may also constrain its overlap with V.

(c) *V_2-to-V_1 coarticulation in $V_1C(C)V_2$ sequences is not modulated by prosody because V_2 is too far from the boundary and/or because its coupling with the elements in its own syllable makes it less flexible.*

In his seminal work, Öhman (1966) proposed that in a VCV sequence, the movements of the tongue body are organized from one vowel to the next, as a single continuous trajectory on which consonantal gestures are superimposed. In this view, the vowels are linked together trans-consonantally and their articulatory specifications interact with each other. Furthermore, since trans-consonantal V-to-V coarticulation vary according to the language (Manuel 1999, Farnetani & Recasens 2010), it is coherent to assume that this coordination between the vowels is controlled. In French, this process has even been phonologized in the case of mid-vowels (/ɛ-e/, /ɔ-o/, /œ-ø/) whose alternation, in some regional varieties and for some speakers, can follow a rule of vowel height harmony. Specifically, mid-closed V_1 variants occur in $V_1C(C)V_2$ when followed by closed and mid-closed V_2 (e.g. *aimer* [eme] "to love") and mid-open variants occur when followed by an open and mid-open V_2 (*il aimait* [ɛmɛ] "he loved"). In an earlier study, we showed that V_1 mid-vowels were less subject to V_2 height harmony in IP-initial position than in word-medial position (Turco, Fougeron and Audibert, 2016).

In the current study, even though we were not dealing with the attested case of French vowel height harmony (i.e. the target was not a mid-vowel), we were expecting to replicate this effect since the origin of the process was similar. Also, earlier studies in English suggested that vowels were more resistant to V-to-V coarticulation in prosodically strong accented position (Fletcher, 2004; Cho, 2004). In our study, V_1 /a/ is found to coarticulate as much with the following V_2 in IP-initial condition than it does in the Word-medial condition. Therefore, V_1 is not more resistant to trans-consonantal V_2 influence when IP-initial, and V_2-to-V_1 coordination does not seem to be modulated by prosody. One possible explanation which follows the idea developed in (a) is that V_2 is too far from the boundary to see its timing relative to V_1 being influenced by the DI effect (i.e. it is not under the scope of the π-gesture). A further explanation would be that the coupling V_2 has with the elements in its own syllable provides sufficient stabilization for it to resist the (somewhat distant and therefore weak) boundary effects (D. Byrd, p.c.). Mok (2012) showed that V-to-V coarticulation in English do depend on the syllabic affiliation of the intervocalic consonants. She showed more V-to-V (acoustic) coarticulation with intervocalic /st./ than /.st/, and almost none with /s.t/. This result suggests that the relationship between the vowels across the syllable boundary depends on their respective coupling with the elements of their own syllable. Unfortunately, in our data it was not possible to obtain balance subsets of the $V_1C(C)V_2$ sequences according to the number of intervocalic consonants and the syllabic affiliation of these consonants. Future studies should collect articulatory data to test how multiple coupling links constrain variation in the timing of the elements.

Taken together, our results have suggested that the modulation of coarticulation by prosodic effects do occur when the timing relationship between the elements is flexible enough to allow variation in overlap. This happens in the case of the sequential organization of nucleus and coda gestures in the VC syllable, and in the case where V and C are not in the same syllable. In IP-initial position, this variation goes toward less coarticulation, so less overlap. Following the π-gesture approach, the gestures of both the IP-initial vowel and the following consonant would be under the scope of the modulation gesture in order to unfold further apart. Whether the reduction of overlap is indeed due to longer delayed initiation and/or longer activation of some gestures, this has to be checked against articulatory data. In the case of the heterosyllabic V.C sequence, however, it remains to be clarified how exactly the π-gesture can modulate the coordination between gestures that are probably not coupled.

To conclude, the current study conducted on a large quantity of data has opened up interesting food for thought about how linguistic constraints interact with motoric ones. Different responses to the influence of prosody has shown that there is not just one type of coarticulation but multiple ones. It will be interesting to develop more research in the future especially in relation to the way these types of coarticulation are handled in speech planning and processing.

Acknowledgments

This research was partly supported by Swiss National Science Foundation grants no. CRSII5_173711 awarded to the third author and by the program "Investissements d’Avenir" ANR-10-LABX-0083 (Labex EFL). It was also supported by the European Union Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 661530 awarded to the second author. We deeply thank Dani Byrd and the two anonymous reviewers for their inspiring comments on the manuscript.

References

Bartón, K. (2014). MuMIn: multi-model inference, R Package Version 1.9. 13. Available at: https://CRAN.R-project.org/package=MuMIn.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48.

Boersma & Weenink, (2014). Praat: doing phonetics by computer [Computer program]. Version 6.1.16, retrieved 21 July 2020 from http://www.praat.org/

Bombien, L., Mooshammer, C., Hoole, P. & Kühnert, B. (2010). Prosodic and segmental effects on EPG contact patterns of word-initial German clusters. *Journal of Phonetics*, 388-403.

Browman, C. P., & Goldstein, L. (1990). Tiers in articulatory phonology, with some implications for casual speech. *Laboratory phonology*, 1: Between the grammar and physics of speech, 341-376.

Byrd, D. (2000). Articulatory vowel lengthening and coordination at phrasal junctures. *Phonetica*, 57(1), 3-16.

Byrd D., Krivokapić J., & Lee S. (2006). How far, how long: on the temporal scope of prosodic boundary effects. *Journal of Acoustic Society of America*, 120(3), 1589-99.

Byrd, D., & Riggs, D. (2008). Locality interactions with prominence in determining the scope of phrasal lengthening. *Journal of the International Phonetic Association*, 38(2), 187-202.

Byrd, D., & Saltzman, E. (2003). The elastic phrase: Modeling the dynamics of boundary- adjacent lengthening. *Journal of Phonetics*, 31, 149–180.

Byrd, D., & Choi, S. (2010). At the juncture of prosody, phonology, and phonetics - The interaction of phrasal and syllable structure in shaping the timing of consonant gestures. *Laboratory Phonology*, 10, 31-59.

Cho, T. (2004). Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics*, 32(2), 141-176.

Cho, T. (2011). Laboratory phonology. The continuum companion to phonology, 343-368..

Cho, T. (2016). Prosodic boundary strengthening in the phonetics–prosody interface. *Language and Linguistics Compass*, 10(3), 120–141.

Cho, T., & Jun, S.-A. (2000). Domain-initial strengthening as featural enhancement: Aerodynamic evidence from Korean. *Chicago Linguistics Society*, 36, 31–44 (Also appeared in UCLA working papers in phonetics, 99, 57–70).

Cho, T. & Keating, P. A. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics*, 29(2), 155-190.

Cho, T., & Keating, P. (2009). Effects of initial position versus prominence in English. *Journal of Phonetics*, 37(4), 466-485.

Cho, T., Kim, D., & Kim, S. (2017). Prosodically-conditioned fine-tuning of coarticulatory vowel nasalization in English. *Journal of Phonetics*, 64, 71-89.

De Jong, K.J., Beckman, M.E., & Edwards, J. (1993). The interplay between prosodic structure and coarticulation. *Language and Speech*, 36, 197–212

Farnetani, E., & Recasens, D. (2010) Coarticulation and connected speech processes. In William J. Hardcastle, John Laver & Fiona E. Gibbon (Eds), *The Handbook of Phonetic Sciences*, 316-32. Chicherster : Wiley-Blackwell.

Fletcher, J. (2004). An EMA/EPG study of vowel-to-vowel articulation across velars in Southern British English. *Clinical Linguistics & Phonetics*, 18(6-8), 577–592.

Fougeron, C. (2001). Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics*, 29(2), 109-135.

Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *The Journal of the Acoustical Society of America*, 101(6), 3728-3740.

Galliano, S., Geoffrois, E., Mostefa, D., Choukri, K., Bonastre, J.-F., & Gravier, G. (2005). ESTER Phase II Evaluation Campaign for the Rich Transcription of French Broadcast News. Proceedings of *Interspeech*, Lisbonne, Portugal.

Gauvain, J.L., Lamel, L. and Adda, G. (2002). The Limsi Broadcast News Transcription System. *Speech Communication*, 37(1-2), 89-108.

Gay, T. (1981). Mechanisms in the control of speech rate. *Phonetica*, 38(1-3), 148-158.

Gendrot, C., Gerdes, K., & Adda-Decker, M. (2011). Impact of prosodic position on vocalic space in German and French. Proceedings of *International Conference of Phonetics Science,* 731-734, Hong-Kong, China. (halshs-00691138).

Georgeton, L. (2014). *Renforcement des voyelles orales du français en position initiale de constituants prosodiques: interaction avec les contrastes phonologiques*. [Strengthening of oral French vowels in initial position of constituants : interaction with phonological contrasts]. PhD dissertation Université Sorbonne Nouvelle - Paris 3, France.

Georgeton, L., Antolík, T. K., & Fougeron, C. (2016). Effect of domain initial strengthening on vowel height and backness contrasts in French: Acoustic and ultrasound data. *Journal of Speech, Language and Hearing Research*, 59(6), 1575-1586.

Georgeton, L., & Fougeron, C. (2014). Domain-initial strengthening on French vowels and phonological contrasts: evidence from lip articulation and spectral variation. *Journal of Phonetics*, 44, 83-95.

Goldstein, L., & Pouplier, M. (2014). *The temporal organization of speech*. In Victor Ferreira, Matthew Goldrick and Michele Miozzo (eds.), The Oxford Handbook of

Language Production. Oxford: Oxford University Press.

Goldstein, L., Nam, H., Saltzman, E., & Chitoran, I. (2009). Coupled oscillator planning model of speech timing and syllable structure. *Frontiers in phonetics and speech science*, 239-249.

Gravier, G., Bonastre, J-F., Geoffrois, E., Galliano, S., Tait, K. Mc., & Choukri, K. (2006). Corpus description of the ESTER evaluation campaign for the rich transcription of French broadcast news. Proceedings of *European Conference on Speech Communication and Technology*, 139–142.

Guitard-Ivent, F. (2018). *Coarticulation C-à-V en français : interaction avec le type de voyelle, la position prosodique et le style de parole.* [C-to-V coarticulation in French: interaction with vowel type, prosodic position and speech style.], PhD dissertation, Université Sorbonne Nouvelle -Paris 3, France. tel-02147399.

Jang, J., Kim, S., & Cho, T. (2018). Focus and boundary effects on coarticulatory vowel nasalization in Korean with implications for cross-linguistic similarities and differences. *The Journal of the Acoustical Society of America*, 144(1), EL33-EL39.

Joo, H., Jang, J., Kim, S., Cho, T., & Cutler, A. (2019, August). Prosodic structural effects on coarticulatory vowel nasalization in Australian English in comparison to American English. Proceedings of 19th *International Congress of Phonetic Sciences* (ICPhS) 835-839. Australasian Speech Science and Technology Association Inc.

Jun S.A. & Fougeron C. (2002). The Realizations of the Accentual Phrase in French Intonation, *Probus* 14: 147-172, A special issue on Intonation in the Romance Languages. Ed. by J. Hualde.

Katsika, A. (2016). The role of prominence in determining the scope of boundary-related lengthening in Greek. *Journal of phonetics*, 55, 149-181.

Keating, P., Cho, T., Fougeron, C., & Hsu, C. (2004). Domain-initial articulatory strengthening in four languages. *Laboratory Phonology*, 6, 143–161.

Krivokapić, J. (2007). The planning, production and perception of prosodic structure. PhD dissertation University of Southern California.

Krivokapić, J. (2014). Gestural coordination at prosodic boundaries and its role for prosodic structure and speech planning processes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1658), 20130397.

Krull, D. (1989). Second formant locus pattern and consonant-vowel coarticulation in spontaneous speech. *PERILUS X*, University of Stockholm, 87 – 108.

Kuehn, D. P. & Moll, K. l. (1976). A cineradiographic study of VC and CV articulatory velocities. *Journal of Phonetics*, 4, 303-20.

Kuznetsova, A., Brockhoff, P.B., & Christensen, R. H. B. (2013). Lmertest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). R package version 2. 0-29.

Lenth, R., Love, J., & Lenth, M. R. (2018). Package ‘lsmeans’. *The American Statistician*, 34(4), 216-221.

Li, H., Kim, S., & Cho, T. (2019). Language-specific prosodic structural modulation of coarticulation vowel nasalization in #NV and CVN# in Mandarin Chinese. Proceedings of 19th *International Congress of Phonetic Sciences* (ICPhS-poster). Australasian Speech Science and Technology Association Inc.

Manuel, S. (1999). Cross-language studies: Relating language-particular coarticulation patterns to other language-particular facts. In W. Hardcastle & N. Hewlett (Eds.), *Coarticulation: Theory, data and techniques*. Cambridge: Cambridge University Press.

Meynadier, Y. (2003). *Interaction entre prosodie et (co)articulation linguopalatale en français [Interaction betwwen prosody and linguopalatal coarticulation in French]*. PhD dissertation, Université Aix-Marseille, France.

Mok, P. (2012) Effects of consonant cluster syllabification on vowel-to-vowel coarticulation in English. Speech Communication, 54: 946-956.

Nam, H., Goldstein, L. & Saltzman, E. (2009). Self-organization of syllable structure: A coupled oscillator model. In Pellegrino, F. (Ed.) *Approaches to phonological complexity*, 16, 299-328, Walter de Gruyter, Berlin.

Öhman S.E.G. (1966). Coarticulation in VCV utterances: spectrographic measurements, *Journal of the Acoustical Society of America*, 39, 1, 151–168.

Pasdeloup, V. (1990). *Modèle de règles rythmiques du français appliquées à la synthèse de la parole.* [Rythmical rule model of French applied to speech synthesis]. PhD dissertation, Université d’Aix en Provence.

R Development Core Team. (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. ISBN 3-900051-07-0, retrieved from http://www.R- project.org/.

Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2016). afex: Analysis of Factorial Experiments. R package version 0.16-1. https://CRAN.R-project.org/package=afex

Sussman, H., Bessell, N., Dalston, E., & Majors, T. (1997). An investigation of stop place articulation as a function of syllable position: A locus equation perspective. *Journal of the Acoustical Society of America,* 101, 2826 – 38.

Torreira, F., Adda-Decker, M., & Ernestus, M. (2010). The Nijmegen corpus of casual French. *Speech Communication*, 52, 201–212.

Turco, G., Fougeron, C., Audibert, N. (2016). The Effects of Prosody on French V-to-V Coarticulation: A Corpus-Based Study. Proceedings of *Interspeech* 2016, 998-1001.

Turvey, M. (1990). Coordination. *American Psychologist 45,* 938-953.

Appendix

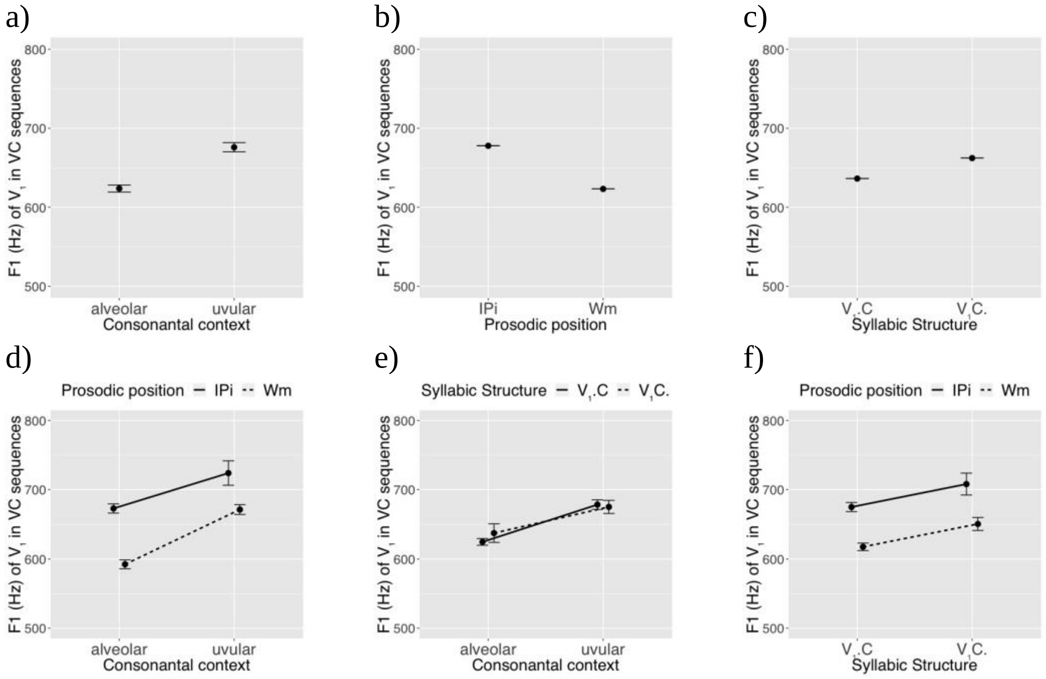
1. R- syntax of the models

seq.	Formants	R-syntax
CV	Fn (F1,F2)	lmer_alt (Fn ~ CONSONANTAL CONTEXT * PROSODIC POSITION + (CONSONANTAL CONTEXT * PROSODIC POSITION |speaker), data=CV, REML=FALSE)
VC	Fn (F1,F2)	lmer_alt(Fn ~ CONSONANTAL CONTEXT * PROSODIC POSITION * SYLLABIC STRUCTURE + (CONSONANTAL CONTEXT * PROSODIC POSITION * SYLLABIC STRUCTURE ||speaker), data=VC, REML=FALSE)
VV	F1	lmer_alt(F1 ~ V_2 HEIGHT * PROSODIC POSITION + (V_2 HEIGHT * PROSODIC POSITION ||speaker), data=VV, REML=FALSE)

R-syntax of the 5 linear mixed models.

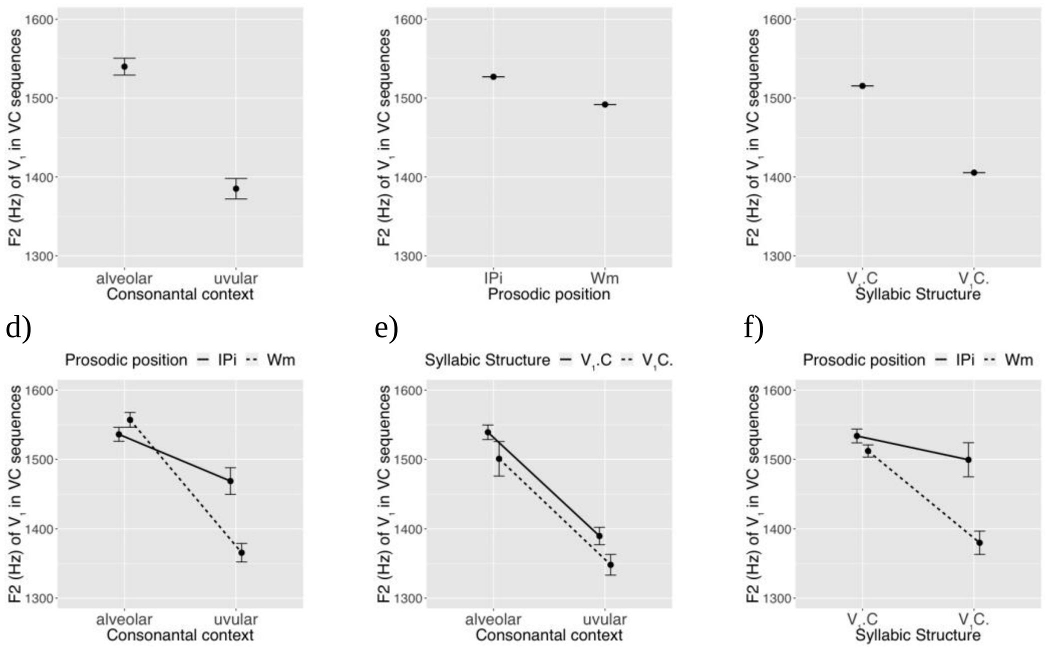
2. Figures

A. Visualization of the effects of each factor on the F1 of / a / for the C-to-V anticipatory coarticulation in VC sequences: a) CONSONANTAL CONTEXT, b) PROSODIC POSITION, c) SYLLABIC STRUCTURE; and their two-way interactions d) CONSONANTAL CONTEXT and PROSODIC POSITION; e) CONSONANTAL CONTEXT and SYLLABIC STRUCTURE, f) PROSODIC POSITION and SYLLABIC STRUCTURE. “IPi” and “Wm” for initial position of Intonational phase and word medial position, “V.C“ heterosyllabic and “VC.“ Tautosyllabic sequences.



B. Visualization of the effects of each factor on the F2 of / a / for the C-to-V anticipatory coarticulation in VC sequences: a) CONSONANTAL CONTEXT, b) PROSODIC POSITION, c) SYLLABIC STRUCTURE; and their two-way interactions d) CONSONANTAL CONTEXT and PROSODIC POSITION; e) CONSONANTAL CONTEXT and SYLLABIC STRUCTURE, f) PROSODIC POSITION and SYLLABIC STRUCTURE. “IPi” and “Wm” for initial position of Intonational phase and word medial position, “V.C“ heterosyllabic and “VC.“ Tautosyllabic sequences.

a) b) c)


F2 (Hz) of V1 in VC sequences
1600
1500
1400
1300
alveolar
uvular
Consonantal context
IPi
Wm
Prosodic position
V1.C
V1C.
Syllabic Structure
d)
e)
f)
Prosodic position — IPi -- Wm
Syllabic Structure — V1.C -- V1C.
Prosodic position — IPi -- Wm