GESTURAL OVERLAP, PLACE OF ARTICULATION AND SPEECH RATE
AN X-RAY INVESTIGATION

*Béatrice Vaxelaire **Rudolph Sock °Pascal Perrier,

*Institut de Phonétique de Strasbourg – Université Marc Bloch (UMB)
22, rue Descartes-67084 Strasbourg Cédex France

°Laboratoire de Psychobiologie du Comportement Moteur et des Sports – UMB

Institut de la Communication Parlée, UMR CNRS 5009–INPG & Université Stendhal - Grenoble – France

ABSTRACT
This investigation deals with the production of consonant sequences in French, with particular focus on overlap of labial, apical, tongue-dorsum and velar gestures. X-ray and acoustic data are obtained for two speakers, at two speaking rates, normal-conversational and fast. Speech rate is varied in order to explore gestural overlapping possibilities, when places of articulation differ between plosives. Results show that: anticipatory coarticulation is not systematically predominant; anticipatory or perseverative coarticulation may or may not occur depending on similarity/dissimilarity in place of articulation between contiguous consonants; vowels are less resistant than consonants to velar coarticulation; speech rate facilitates gestural overlap. Findings are explained in terms of biomechanical and viable linguistic constraints.

1. INTRODUCTION
This research is of a programmatic character and is preoccupied with anticipatory and carryover gestures in speech production and perception (see, for e.g., Vaxelaire, 1995, Vaxelaire & Sock, 1996; Vaxelaire, 1997). Particular attention is paid to velar activity in the present study, as velar lowering could be linguistically relevant in French. Velar coarticulation, simply considered here as the spreading or the extension of certain effects to adjacent or neighbouring segments, should be regarded as of utmost importance in fluent speech (see, for e.g. Benguerel et al., 1977; Bell-Berti & Krakow, 1991). Indeed, listeners exploit early cues related to upcoming speech elements (see Sock et al., 1999, for the perceptual effects of anticipatory labial activity in French). It is also known that some coarticulatory effects in synthetic speech improve perception of certain segments (Maeda, 1999). Considering coarticulation as overlap or blending of gestures (Brownman & Goldstein, 1990; Munhall & Löffqvist, 1992), it could be explained in terms of production mechanisms related to the co-ordination of relatively sluggish physiological speech structures, also constrained by the perceptual requirements of the linguistic code (see, for e.g., Perrier et al., 1996, for a biomechanical modelling of coarticulatory effects, following the Equilibrium Point Hypothesis). The main thrust in this work is therefore to try to explain gestural overlap, with reference to linguistically viable configurations, that largely depend on factors such as the nature of the adjacent or neighbouring segments, the articulators coming into conflict in producing a specific sequence of sounds, speaker-specific strategies, etc.

Varying consonantal place of articulation, should allow testing the degree of overlap as a function of similarity or proximity in contact locations. Increasing speech rate should provoke a speeding-up of articulatory activity, thus enhancing gestural overlap and compression of acoustic durations. This investigation may also offer insights into preferred coarticulatory directions (anticipatory or carryover), articulatory and acoustic timing constraints and phasing relations between articulatory and acoustic events.

2. METHOD
The corpus consisted of 58 short sentences that embedded the target words. Some of these words were specifically chosen to study the production of consonant sequences in a V1C1C2V2 context, where V1 and V2 are always the low vowel /a/, C1 is the nasal plosive /m/ followed by C2, one of the oral plosives /p, t, k/ or /b, d, g/, and symmetrically, where C1 is one of these oral plosives followed by C2 the nasal plosive. Place of articulation and voicing were thus also varied. The following sentences are examples from the corpus: “Ma cape marron.” /apma/ vs. “Ça me paraît bien.” /ampa/; “Des pattes magiques.” /atma/ vs. “Les lames tachées.” /atma/; “Tes sacs marrons.” /akma/ vs. “Ça me calmera.” /akma/; “Des crabes marrons.” /abma/ vs. “Ça me balance loin.” /amba/; etc. The data analysed for the present study is based on 12 (out of 58) sentences produced by two speakers, at a normal conversational speaking rate and at a self-selected fast rate. There were 5 conditions in all: 2 speakers (one male, AE and one female, MM) x 6 consonant types (/p, t, k, b, d, g/) x 2 speech rates (normal and fast) x 2 voicing contrasts (voiced and voiceless) x 2 nasal positions (C1 vs. C2). In the present study, only data from the voiceless condition and for a single speaker will be presented here. Speaker differences are not significant in this particular research, and results obtained in the voiceless condition will be reported in a future work.

With the help of a grid, measurement parameters for vocal tract configurations (Figure 1) were determined. Upper lip, lower lip, tongue-tip and tongue-dorsum movements were tracked and measured (in the axis were contact would be made subsequently), from the preceding “vocalic” oral opening configuration, through the “consonantal” obtrusive gestures, to
the subsequent “vocalic” oral opening configuration. As concerns the behaviour of the velum, both velum height, at the highest point, and velo-pharyngeal opening were measured. Previous studies (cf., e.g., Vaxelaire & Sock, 1996) have shown that measurements acquired at the highest point on the velum were better correlated with velic aperture. Thus maximum velic lowering and raising correspond to maximum velo-pharyngeal cross sectional opening and closing, respectively. Consequently, the terms velic height and port opening, following the above criteria, may be used interchangeably to better reflect acoustic nasalization.

A simultaneous audio recording of the speakers’ productions was obtained. Temporal events were detected on the audio signal and specific timing relations between these events allowed determining, in the VCCV domain, acoustic durations that correspond to articulatoriy opening and closing gestures, and also onsets and offsets of laryngeal activity.

Results obtained from a frame-by-frame analysis (50 frames per second) of X-ray profiles, together with the synchronised audio, are presented below. Since X-ray images are only available every 20 milliseconds, analyses of articulatory-acoustic timing relationships will always take this factor into account.

3. RESULTS AND DISCUSSION

3.1. Timing of Labial and Velar Gestures

In the /apma/ sequence, at a normal speaking rate, overlap of /p/ and /m/ configurations is revealed by the following: the velo-pharyngeal port is open during bilabial closure for /p/. That anticipatory opening of the port (Figure 2) does begin during the obstructive phase of the /p/ closure, is revealed by the expected absence of voicing on the acoustic signal for /p/. In fact, consonantal voicing for /m/ only begins when maximum opening of the port (3 mm) is obtained, presumably indicating that coupling of the pharynx and the nasal cavities may enhance vocal fold vibrations.

These timing relationships are confirmed in fast speech. Whereas anticipatory opening of the velo-pharyngeal port also occurs in the /p/ configuration, voicing onset for the nasal consonant is only observed well after a 2 mm opening of the port is attained. It is also interesting to note that in both prosodic conditions, the voiced acoustic closure duration, usually associated with the emergence of the nasal consonant, does not necessarily require an articulatory closed-lip configuration. Relative lip opening is indeed observed within this acoustic consonantal phase, thus revealing varying degrees of vocal tract obstruction during this acoustic closure, going from a tight occlusion to an increase in bilabial stricture. Moreover, the port remains open well into the late configurations of the subsequent vowel /a/, also showing a perseverative influence of the nasal consonant on the vowel. In fact, the data for velar activity, in general, confirm previous findings (Vaxelaire & Sock, 1996) indicating that, in a vocalic context, the velum is “free” to either anticipate its lowering movement or delay its raising movement.

In an /ampa/ sequence, at a normal speaking rate, there is temporal overlap of the /m/ and /p/ gestures, as opening of the velo-pharyngeal port begins during vocal tract configurations for /a/, is maximal during the bilabial contact for /m/, and continues into early configurations of the /p/ consonantal production, i.e. after offset of voicing buzz for /m/ on the acoustic signal. For the same sequence in fast speech, this gestural overlap is maintained, as the port remains open in an early configuration of the /p/ bilabial contact, at offset of /m/ consonantal voicing.

Comparison of the two sequences /apma/ and /ampa/ seems to suggest that the extent of spreading of the velar gesture of the nasal consonant (opening of the port) is, however, more pronounced in the anticipatory than in the carryover direction.
3.2. Timing of Labial, Apical and Velar Gestures

In the /atma/ sequence, in normal speech, opening of the velo-pharyngeal port occurs after release of the apical contact; both gestures are followed by lip contact for production of the nasal consonant. Thus, only the transitional phase between /t/ and /m/ seems to be affected by a reduced coarticulatory effect of port opening. The apical gesture for the /t/ and the labial and velar gestures for the /m/ are therefore highly sequential here. In fast speech, anticipatory opening of the port, together with bilabial contact overlap with apical contact for the /t/ consonant (Figure 3), although acoustic voicing onset for the nasal consonant only begins much later (i.e. approx. 30 ms later); at this point velic aperture is at its maximum (3 mm).

Figure 3: Anticipatory opening of the port, together with bilabial contact, overlap with apical contact for the /t/ consonant, in an /atma/ sequence; fast speech.

In an /amta/ sequence, results show, in normal speech, that gestures overlap more readily than in the previous context. A simultaneous occurrence of lip contact and port opening for the production of /m/, together with apical contact for /t/ are observed through three frames. The acoustic obstruent portion, associated with the production of the voiceless plosive /t/, is partly voiced. This voiced portion coincides with delayed opening of the velo-pharyngeal port during apical contact. In fast speech, the data show a phase with overlap of the three gestures for /m/ and /t/ and a phase with overlap of only the labial and the apical gestures. Although the port is closed at these latter stages, the corresponding portion of the acoustic signal still reveals voicing.

In the /akma/ sequence, overlapping gestures are observed in normal speech: although tongue-dorsum contact is queued after bilabial release, the lingual contact coincides with port-opening. Again, acoustic voicing continues after closing of the port, into a phase in the acoustic signal that corresponds to closure of the velar consonant /k/. When speaking rate is increased, all three gestures overlap towards the final configurations of the nasal plosive /m/. The acoustic closure for the plosive /k/ is also partly voiced, even after port closure.

3.3. Timing of Labial, Tongue-Dorsum and Velar gestures

In the /akma/ sequence, in normal speech, the different gestures are quite sequential. Port-opening and bilabial contact are queued after release of tongue-dorsum contact. However, there is an anticipatory opening (of approx. 2 mm) of the port, before bilabial contact and onset of vocal fold vibrations for the nasal consonant, on the acoustic signal. A simultaneous occurrence of gestures becomes possible in fast speech, when bilabial contact, port opening and tongue-dorsum contact coincide (Figure 4). Although bilabial contact and port-opening seem to emerge simultaneously, onset of consonant vibrations on the acoustic signal is delayed, here also, with respect to these articulatory events.

Figure 4: A simultaneous occurrence of bilabial contact, port opening and tongue-dorsum contact, in fast speech, for an /akma/ sequence.

In the /amka/ sequence, overlapping gestures are observed in normal speech: although tongue-dorsum contact is queued after bilabial release, the lingual contact coincides with port-opening. Again, acoustic voicing continues after closing of the port, into a phase in the acoustic signal that corresponds to closure of the velar consonant /k/. When speaking rate is increased, all three gestures overlap towards the final configurations of the nasal plosive /m/. The acoustic closure for the plosive /k/ is also partly voiced, even after port closure.

4. SUMMARY

4.1. Articulatory Timing Constraints

On a whole, these results, related to interarticulatory (labial, velar and lingual) gestures, confirm one of our initial hypotheses, i.e. overlap would be more readily achieved when separate, but co-ordinated, structures are involved. However, such a statement needs to be further refined with regards to various factors such as proximity in place of articulation, consonant position in the sequence, predominance of the direction of the coarticulatory effect, speech rate, etc.

When interarticulatory gestures — flanked by the low vowel — present a homorganic place of articulation (m/p or p/m), overlap is possible even at a normal speaking rate. Co-ordination is comparatively simple, as only labial and velar gestures are recruited. The anticipatory effect of the velar gesture seems to be more predominant.
In the case of heterorganic places of articulation that recruit labial, apical and velar gestures (tm/mt), co-ordination becomes more complex. The consequence is a more problematic condition for overlap, when tongue-tip contact has to be made before bilabial closure (and velic lowering), as in /tm/, at a normal speaking rate. The result is that no overlap is observed at all. However, when bilabial closure precedes apical contact, as in /mt/, it seems that the simultaneous emergence of the apical contact, within a closed-lip configuration with a lowered velum, is not a conflicting gesture as in the former case. Previous findings (Vaxelaire, 1995) have shown that in the oral consonant sequence (/pt/), the bilabial contact does not necessarily have to be released before implementation of the apical one. Including velum control in the present context does not seem to significantly modify the timing relationship between the labial and the lingual gestures. In fast speech, coarticulatory differences between the two sequences, that seemed to depend, at a normal speaking rate, on order of sequencing, are neutralised: overlap, due to anticipatory or perseverative effects, becomes possible in both cases.

When heterorganic places of articulation recruit labial, tongue-dorsum and velar gestures (km/mk), the scenario is more or less similar to the previous case. No overlap is noticed in normal speech for the /km/ sequence, presumably due to constraints in obtaining a tongue-dorsum contact before lip closure, with partial overlap. In the /mk/ sequence, overlap is possible but only with two gestures, i.e. velic lowering and tongue-dorsum contact. The greater distance between the labial and the lingual gesture in this context, seems to account for difficulty in overlapping the three gestures. It is only speech rate that facilitates a simultaneous occurrence of these conflicting gestures, regardless of the sequence.

Taken together, these results also tell us that anticipatory coarticulation may not be systematically predominant in coarticulatory effects, as is usually reported in the literature. The direction and extent of spreading of a given characteristic (here nasality) depend on the nature of the adjacent sound, proximity in place of articulation and speech rate increase. Thus vowels are less resistant than consonants to nasal coarticulation: they are equally affected by anticipatory and carryover effects. A consonant is equally resistant to nasal anticipatory or perseverative effects; the direction will depend on dissimilarity in place of articulation between the contiguous consonants, and order of emergence in the sequence.

4.2. Articulatory-Acoustic Timing

Articulatory-acoustic relative timing between port opening, and the acoustic onset of voicing for the production of the same nasal consonant show remarkable phasing relations. Briefly, anticipatory velum opening during the preceding voiceless consonant is a general tendency (in one instance, its domain does not go beyond the burst transition of the /l/ consonant). This result suggest that early port opening may serve to enhance the production of the nasal consonant, which combines nasal resonance and modal vocal fold vibrations. Carryover coarticulation is only observed in the context containing homorganic consonants; the port is still open after voicing has ceased. Thus, whilst nasalization seems to affect the entire voiced phase of the nasal consonant in homorganic contexts, it does not spread to the terminal portion of the nasal consonant, in heterorganic sequences. The reason may here also be related to proximity in place of articulation.

5. REFERENCES