Temporal patterning in French initial C/l/-clusters

Barbara Kühnert[1] *et Phil Hoole*[2]
[1] Institut du Monde Anglophone & Laboratoire de Phonétique et Phonologie (UMR 7018 – CNRS / Paris 3)
5 rue Ecole de Médecine, 75006 Paris, France
[2] Institut für Phonetik und Sprachliche Kommunikation / Ludwig Maximilians Universität München
Schellingstrasse 3/II, 80799 München, Allemagne
barbara.kuhnert@wanadoo.fr hoole@phonetik.uni-muenchen.de

ABSTRACT

This work examines aspects of inter-consonantal cohesion within French word-initial C_1/l/-clusters in light of recent proposals of gestural coordination. Based on articulatory and acoustic events, the timing of tongue and lip movements in one subject was studied using an electromagnetic transduction device. More temporal overlap between C_1 and /l/ gesture onset as well as /l/ closure period was found for /pl/ than /kl/. Although matching similar patterns of overlap in initial stop clusters, this 'place of articulation' effect is attributed to low-level motor factors rather than to considerations of perceptual recoverability. An additional analysis of the overall C-centre of /pl/ and /kl/ showed a surprising temporal stability, confirming a relative constant phasing between initial consonant sequence and following vowel.

1. INTRODUCTION

1.1. General background

In dynamically defined speech production models, the relevant units of articulation are *gestures*, which supposedly capture both, the phonological and physical properties of speech [1, 2]. Crucial to this account is the notion of gestural co-production, that is, the way different autonomic gestures overlap in time and space to generate the smooth sequential movements characterizing spoken language.

Over the last few years, much work has been devoted to exploring the temporal organization of articulatory gestures. In summary, it has been shown that the timing properties of gestures are influenced by global factors, such as speaking rate and style, as well as by prosodic factors [for an overview, see 3 and references herein]. In addition, the temporal coordination between gestures seems to be position-dependent, with word-initial clusters tending to show less gestural overlap than the same clusters in word-medial position or across word-boundaries [4].

1.2. Intergestural coordination in consonant sequences

Within the framework of the dynamic speech production models mentioned above, the observed variations in gestural cohesion are assumed to emerge from differences in the underlying coupling between the gestural units. More specifically, Browman and Goldstein [5] hypothesize that gestures within a syllable are associated with different degrees of bonding (or coupling) strength. Consonant gestures in syllable initial position exhibit both, a phasing relation with the following vowel as well phasing relations directly to each other. It is the competition of these two phase relations, which results in the C-centre effect, a temporal anchor point within the consonant sequence which is assumed to preserve a stable timing with respect to the vowel, regardless of the number of consonants and the phonetic composition of the cluster. No such inter-consonantal phasing is supposed to occur in syllable-final position where inter-gestural timing may therefore show higher variability.

A further factor, which has been evoked to affect timing relationships, is the consideration of the perceptual recoverability of gestures. Essentially, less overlap between gestures should occur in clusters in which perceptual recoverability is at stake. This has been applied to the observation that consonant clusters show less overlap in initial position than elsewhere. It also has been applied to account for studies, which repeatedly have shown that in CC-stop clusters the degree of gestural overlap is related to the place of articulation. Essentially, back-to-front sequences (e.g. [kt]) exhibited less temporal overlap than front-to-back sequences (e.g. [pt]) [see 4 for English; 6 for Georgian].

In light of this context, the following pilot study extends the empirical study of temporal patterning to C/l/-clusters in French, investigating both, their intra- and inter-gestural timing properties.

2. METHOD

2.1. Speakers and speech material

One male native speaker of French without any known hearing or speech disorders participated in the experiment. The data is a subset of a more extensive corpus in which the subject produced multiple repetitions of utterances of the form « Je vois *word*$_1$ et *word*$_2$ et *word*$_3$ ». Word$_1$ and word$_2$ were the target words containing all French phonotactically legal initial C, CC and CCC-sequences; word$_3$ was a dummy. The vowels following the consonant sequences were, as often as possible, [i] and [a]. Here, we report the results for 10 repetitions of the consonant cluster productions [pl] and [kl] as uttered in the words *plaque* and *claque*.

The items were randomized in such a way that each target word occurred five times as word$_1$ and five times as word$_2$ in order to avoid potential position effects. The sentences were prompted on a computer screen and the subject was instructed to read them at a self chosen speed.

2.2. Recording procedure

Movements of the lips, mandible and four points on the tongue were captured using a 3-D electromagnetic transduction device. Additional sensors were used to serve as reference coils. The acoustic speech signal was recorded synchronously with the movement signals. A detailed description of data acquisition, normalisation and preparation procedures is outlined in Hoole [7].

Our focus is on data from the lower lip (UL), the tongue tip (TT) and rearmost tongue body (TB) transducer, associated with the production of the bilabial stop /p/, the liquid production /l/ and the production of the velar stop /k/, respectively. The label 'tongue tip' is used here for conventional reasons only. The front most tongue receiver was placed about 1 cm posterior to the tip of tongue, thus rather corresponding to the tongue blade. The rearmost coil was attached about 0.5 cm behind the place where the tongue touched the palate during velar occlusion. The two remaining transducers were placed in between TT and TB receivers at approximately equidistant intervals, but will not be considered in the further presentation.

2.3. Data analysis

The data were analyzed using MATLAB to identify the following kinematic parameters, using maxima in the vertical displacement and relative minima in the velocity signals of the EMMA recordings: the onset of the bilabial closing gesture and the onset and offset of the bilabial closure phase for /p/; the onset of the closing gesture and the onset and offset of the closure phase for /l/; and, the onset of the velar closing gesture and the onset and offset of the velar closure phase for /k/. An illustration of parameter extraction is given in Figure 1.

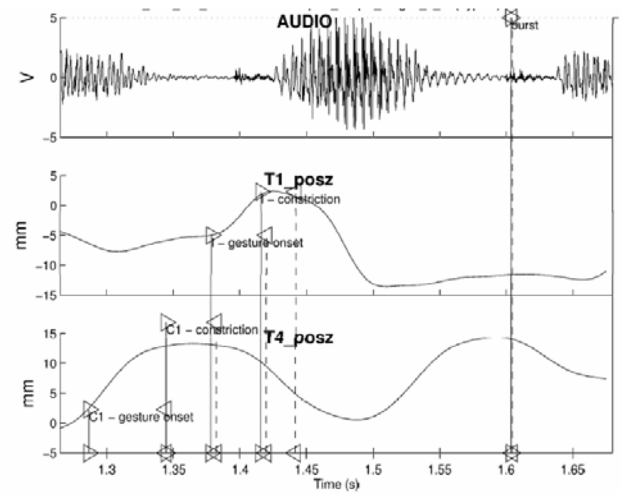

Figure 1: Illustration of kinematic parameter extraction (see text for details). Top panel: audio; middle panel: vertical TT-movement; bottom panel: vertical TB-movement.

As index of the temporal overlap between the plosive and the /l/ gestures, two measures were evaluated: (a) the interval between the release of the initial stop and the moment in time at which movement onset for /l/ is initiated, i.e. how early does movement onset of /l/ occur within C1; and (b), the interval between the release of the initial stop and the moment in time at which the target position for /l/ is reached, i.e., following the terminology of Chitoran et al. [5], the temporal lag between C1 constriction offset and C2 constriction onset.

3. RESULTS

3.1. Gestural overlap and intra-gestural properties

The results of the overlap measures are illustrated graphically in Figure 1, below. Note that the higher the positive value the earlier the /l/-gesture onset starts during the stop closure period; the higher the negative value the later the target position for /l/ is reached. The data show a clear effect of place of articulation with a prominently greater amount of overlap for the onset of the /l/ gesture in /pl/ than in /kl/-sequences. Likewise, the constriction plateau for /l/ is reached earlier when following /p/ than when following /k/. An unpaired two-tailed t-test calculated separately for gesture onset overlap and closure lag confirms that the difference for both parameters is significant (onset overlap: $p > 0.001$; closure lag: $p > 0.05$).

Implicit in the figure above is the observation that the intra-gestural timing properties of the /l/-productions, in

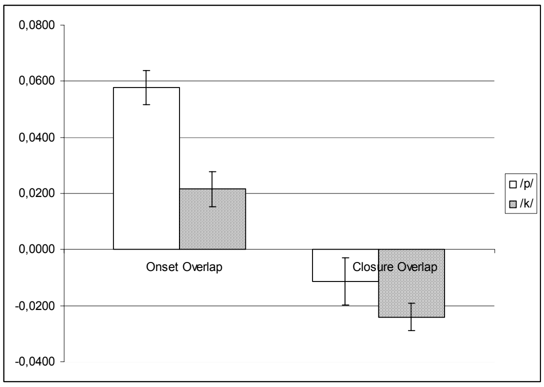

Figure 2: Mean values and standard deviation (ms) of overlap for onset of closing gesture (left) and of lag for closure onset (right) during /l/-productions in the context of /p/ and /k/ (n=10).

particular the time from its onset to target attainment, vary in function of C_1. Table 1 summarizes the durations of the different components of each gesture. As can be seen the two stops have fairly similar temporal properties, while /l/-closure is slightly, the intial /l/ closing period considerably lengthened in the context of /p/.

Table 1: Average durations of gesture onset movement and maximal constriction in ms (standard deviation in brackets; n=10).

	duration of closing gesture (ms)	duration of constriction plateau (ms)
/k/	63 (05)	59 (08)
/p/	67 (08)	55 (12)
/l/, C_1 = /k/	**45 (05)**	42 (10)
/l/, C_1 = /p/	**69 (06)**	47 (08)

3.2. Overall timing of the C-centre

As outlined previously, the C-Centre corresponds to the centre of the consonant sequence, which is hypothesized to represent the point in time that preserves a constant timing relationship with respect to the following vowel. Following Browman and Goldstein [5], the C-centre is calculated as the means of the centres of the target constrictions intervals of the individual gestures of C_1 and C_2, as indicated above. The location of the means here are aligned relative to the acoustic burst of the final consonant, assuming that the vowel gesture, whose temporal properties are difficult to measure directly, is coordinated equally with the final /k/ in both, *claque* and *plaque*.

Table 2 presents the mean location of the C-centre for the two clusters investigated relative to the vocalic alignment point. As can be seen this global timing measure appears to be fairly stable across both consonant sequences, despite the different phasing relations between the individual consonant gestures observed above (unpaired two-tailed t-test: p=0.695; n.s.).

Table 2: Means and standard deviations of the location of the C-Centre relative to the acoustic burst of the word-final /k/ (time in ms; n=10).

Cluster	mean (ms)	std (ms)
/pl/	268	20
/kl/	264	26

4. DISCUSSION

The patterns of overlap observed for French initial C/l/-clusters in this study show the same place of articulation effect as previously reported for in initial stop-stop sequences of other languages [e.g. 4, 6]. More overlap in direction front-to-back, i.e. /pl/, than in the direction back-to-front, i.e. /kl/. It seems to us, however, that this pattern is not only due to factors of perceptual recoverability, but also due to simple constraints of the executing motor system.

Firstly, since the liquid production /l/ does not involve a total constriction in the vocal tract it does not mask the possible perceptual information about the preceding consonant in the same way as a plosive would. Thus, if anything, problems of perceptual recoverability are weaker in the case of liquids and more overlap should be allowed.

The second argument relates to the fundamental difference in the duration of gesture onset to target achievement of /l/, being significantly longer when following /p/ than when following /k/. This difference cannot be attributed to possible differences in the underlying gesture specification, for example, in the parameter *stiffness* which would influence the relationship between gesture duration, velocity and displacement. However, the parameter settings in all /l/-productions are supposedly the same. Rather, it seems that the restricted overlap in the context of /k/ is caused by constraints on the overall tongue configuration during velar productions. In the context of a preceding /p/ with no lingual involvement, the tongue is simply free to move and anticipate the /l/-articulation already during the bilabial closure phase and consequently the constriction plateau can be reached earlier. Velar articulations, by contrast, require a holistic tongue movement, which also involves the tongue front. The specific properties of velar stop production, in particular, are known to highly constrain the articulator movement of the whole tongue during closure interval, as has been observed in many physiological speech production studies [e.g. 9]. Thus, in the context of /kl/ the tongue blade cannot execute the articulation of the /l/-gesture as early as for /pl/. Additional evidence comes from the subject's /skl/-productions, as in *sclérose*. A representative example is given in Figure 3, taken from the same recording session. As can be seen, the tongue blade (mid panel) does not move directly from the fricative articulation towards the lateral during velar occlusion, as one could expect, but rather exhibits a slight downward movement. The /l/ gesture is only initiated at the end of /k/-closure.

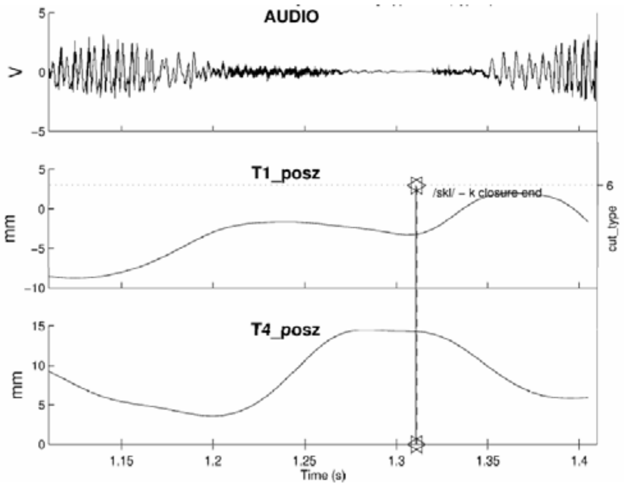

Figure 3: Example of /skl/-production. Top panel: audio; mid: vertical TT-movement; bottom: vertical TB-movement. The cursor is positioned at the kinematically defined end of velar closure.

Evidently, this data does not tell us anything about whether the present differences in overlap might also reflect some differences in the underlying sequencing of the individual gestures. However, they testify to the complex relationship between underlying control structures and observable kinematic behaviour. Even though velar stop productions are only specified for the underlying tract variables of tongue body, this does not imply that the gesture is unspecified for other articulators on the phonetic surface.

Finally, despite the apparent difference in inter-gestural phasing between /pl/ and /kl/ in the study, the stability of C-centre of the consonant-sequences supports the notion that word- (or syllable-)initial onset clusters are characterised by a special inter-consonantal cohesion, being partially responsible for the robustness of consonant clusters in this position.

5. CONCLUSION

Given the ubiquitous variability in speech production, it goes without saying that our results are only suggestive at this time, considering that the data of only one speaker has been analyzed. However, the results show that only a detailed investigation of many empirical data will allow us to sort out which of the gestural properties are due to the constraints of the executing motor system and which are the consequences of specifications in the central control structures. The results also show that the place of articulation effect of reduced temporal overlap in back-to front, i.e. velar-alveolar, sequences, is to some extent a reflection of the low-level dynamical properties of tongue back articulations.

We are presently finishing the analysis of a second speaker as well as extending the data analyses to further consonant sequences.

BIBLIOGRAPHIE

[1] E. Saltzman and K. Munhall. A dynamical approach to gestural patterning in speech production. *Ecological Psychology*, 1:333-382, 1989.

[2] C. Browman and L. Goldstein, 1992. Articulatory Phonology: an overview. *Phonetica*, 49:155-180, 1992.

[3] D. Byrd, 2003. Frontiers and Challenges in Articulatory Phonology. *Proceedings of ICPhS XV,* pages 89-92, 2003.

[4] D. Byrd. Influences on articulatory timing in consonant sequences. *Journal of Phonetics*, 24(2): 209-244, 1996.

[5] C. Browman and L. Goldstein. Competing constraints on intergestural coordination and self-organization of phonological structures. In *Bulletin de la Communication Parlée*, volume 5, pages 25-34, 2000.

[6] I. Chitoran, L. Goldstein and D. Byrd. Gestural overlap and recoverability: Articulatory evidence from Georgian. In: *Papers in Laboratory Phonology* VII, C. Gussenhoven, T. Rietfield and N. Warner (eds.), Mouton de Gruyter, pages 419-448, 2002.

[7] P. Hoole, A. Zierdt and C. Geng. Beyond 2D in articulatory data acquisition and analysis. *Proceedings of ICPhS XV,* pages 265-268, 2003.

[8] J. Perkell. Physiology of speech production: results and implications of a quantitative cineradiographic study. MIT Research Monograph 53, MIT Press, Cambridge, MA, 1969.