Gestural overlap and C-center in selected French consonant clusters
Barbara Kühnert, Phil Hoole, Christine Mooshammer

To cite this version:
Barbara Kühnert, Phil Hoole, Christine Mooshammer. Gestural overlap and C-center in selected French consonant clusters. 7th International Seminar on Speech Production (ISSP), Dec 2006, Ubatuba, Brazil. pp. 327-334, 2006. <halshs-00185737>

HAL Id: halshs-00185737
https://halshs.archives-ouvertes.fr/halshs-00185737
Submitted on 6 Nov 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.
Gestural overlap and C-center in selected French consonant clusters

Barbara Kühnert1,2, Philip Hoole3, Christine Mooshammer4

1Institut du Monde Anglophone, Université Paris 3 – Sorbonne Nouvelle, and 2Laboratoire de Phonétique et Phonologie (CNRS / UMR 7018)
5, rue de L’Ecole de Médecine, 75006 Paris, France

3Institut für Phonetik & Sprachliche Kommunikation, Ludwig-Maximilian-Universität München, Schellingstrasse 3 / II, 80799 München, Germany

4Institut für Phonetik & Digitale Sprachverarbeitung, Christian-Albrechts-Universität Kiel, 24098 Kiel, Germany

barbara.kuhnert@univ-paris3.fr, hoole@phonetik.uni-muenchen.de
timo@ipds.uni-kiel.de

Abstract. Inter-consonantal cohesion in French word-initial CC clusters is investigated in light of recent proposals of gestural coordination. Specifically, the timing of lip and tongue movements of C1/l/ and C1/n/ productions, with C1 being one of the consonants /p, f, k/, of two speakers were studied using electromagnetic articulography (EMA). In French, C/l/ clusters occur frequently in word-initial position, while C/n/ sequences have a limited distribution in less common words. The results provide evidence that liquid as well as nasal sequences consistently show the same ‘place of articulation’ effect of gestural overlap as previously reported for stop-stop clusters, with more overlap in front-to-back than back-to-front sequences. This suggests that the effect is partially due to low-level motor constraints rather than considerations of perceptual recoverability. Furthermore, the results show that the lexical frequency of the sequence does not influence the timing in a simple, categorical manner: stop-nasal clusters showed a strikingly different inter-gestural coordination in comparison with their stop-/l/ counterparts, while no such differences could be observed for the fricative pair. An additional analysis of the overall C-centre of the CC structures demonstrated a rather high temporal stability suggesting that there is, despite many timing differences emerging at the phonetic surface, a relative constant phasing between initial consonant sequence and following vowel.
1. Introduction

Over the last few years, much work has been devoted to exploring the temporal organisation of articulatory gestures in the production of consonant clusters. Specifically, word-initial consonant clusters have been shown to behave in many different ways when compared to their non-initial counterparts. Broadly speaking, they seem to be more “robust” in the sense that they tend to show less overlap between component gestures and they are less variable in their timing properties (e.g. Byrd, 1996). Even though, and since the production of a consonant cluster is not a particularly simple action, the detailed mechanisms of its temporal coordination entail many factors which remain unknown.

There are several proposals which have been put forward to influence the inter-gestural cohesion of word-initial consonant sequences. Within the framework of Articulatory Phonology, variations in gestural coordination are assumed to emerge from differences in the underlying coupling between gestural units. Browman & Goldstein (2000) hypothesize that gestures within a syllable are associated with different degrees of bonding 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-center 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 other syllable positions where inter-gestural timing may therefore show higher variability (cf. Nam & Saltzman (2003) for implementing inter-gestural coupling along these lines).

A further factor, which has been evoked to affect timing relationships, is the consideration of 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 as well as to account for studies, which repeatedly demonstrated a “place order” effect in stop-stop sequences. Back-to-front sequences (e.g. [kt, tp]) showed less temporal overlap than front-to-back sequences ([tk, pt]) (Chitoran et al., 2002; Kochetov & Goldstein, 2005). The argument is that a C2 constriction anterior to C1 masks the acoustic properties of C1 more than a C2 constriction posterior to it – thus, less overlap is allowed. In more recent pilot studies, the same effect has been observed for stop-liquid clusters. Since liquid productions, however, do no threaten the perceptual distinctiveness of C1 in the same way as a stop articulation, it has been suggested that the “place order” effect might be partially due to low-level motor constraints (Kühnert & Hoole, 2006) or rather reflect more fundamental underlying differences of lexically specified coordination patterns (Chitoran & Goldstein, 2006).

Gestural coordination might furthermore be affected by lexical characteristics of words, such as usage frequency, or sub-lexical characteristics, such as the phonotactic probability of the sequence (with both factors being, in fact, highly correlated). Both have been shown to affect production latency as well as production accuracy (Wright, 2004; Vitevitch et al., 2004). Wright argued that, even though word frequency appears to be encoded at the lexical level, it also affects more fine grained aspects of speech
production. In a unrelated, but similar, vein Davidson & Stone (2003) reported that speakers who have to produce consonant clusters which are phonotactically illegal in their native language often fail to employ the appropriate gestural coordination and have a tendency to pull the consonant gestures apart (see also Davidson, 2006).

In light of this context, the following study investigates both, the intra- and inter-gestural timing properties of word-initial C/l/ and C/n/ clusters in French in order to address the various potential factors influencing gestural orchestration in consonant sequences.

2. Methods

2.1 Subjects and speech material

Two male native speakers of French, exhibiting normal speech and hearing, served as subjects (henceforth S1 and S2). The data are a subset of a more extensive corpus in which the subjects produced multiple repetitions of utterances of the form « Je vois word₁ et word₂ et word₃ ». Word₁ and word₂ were one-syllable content words containing all French phonotactically legal initial C, CC and CCC-sequences; word₃ was a dummy. Here, we report the results for 10 repetitions of cluster productions in which C₁ was /p/, /f/, or /k/, and C₂ /l/ or /n/. Unlike C/l/ sequences, C/n/ sequences have a limited distribution in French and are constituents of only a few and rather less common words. We used the Lexique3 database (New et al., 2001; http://www.lexique.org/) to estimate the frequency of each test word relative to one million occurrences (see Table 1; in brackets in bold italics) and to estimate the frequency of the initial CC structure (given in normal font). Note that /fn/ in French only occurs in the name of a widespread book retailer, FNAC, and no estimation could be derived from the database. The items were randomized; the sentences were prompted on a computer screen and the subjects were instructed to read them at a self-chosen speed.

<table>
<thead>
<tr>
<th>C2</th>
<th>C1</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>bilabial</td>
<td>labio-dental</td>
<td>velar</td>
</tr>
<tr>
<td>/l/</td>
<td>plaque (46, 2620)</td>
<td>flaque (23, 1442)</td>
<td>claque (74, 1747)</td>
</tr>
<tr>
<td>/n/</td>
<td>pneu (17, 63)</td>
<td>fnac (--, --)</td>
<td>knout (0.2, 4)</td>
</tr>
</tbody>
</table>

Table 1. Test items with estimated frequencies (per million). In bold italics: word frequency; in normal font: C₁C₂ frequency (see text).

2.2 Recording procedure and data analysis

Movements of the lips, mandible and tongue were captured using a three-dimensional electromagnetic transduction device (Carstens AG500). Additional sensors served 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 et al. (2003). Our focus is on data from the
upper and lower lip (UL & LL), the tongue tip (TT) and rearmost tongue body (TB) transducer.

Fig. 1 provides an illustration of parameter extraction for one production of *claque* of S1. The data were analyzed using MATLAB routines to identify several kinematic parameters in the displacement and velocity signals of the EMMA recordings. Specifically, the on- and offset of: C₁ constriction plateau, C₂ closing gesture and C₂ constriction plateau. For bilabial stops, the tangential velocity of lip aperture, calculated as vertical distance between UL and LL, was used. For labio-dental fricatives, the tangential velocity of LL was evaluated, while for velar stops the velocity signal of the vertical movement of TB, and for /l/ and /n/ the tangential velocity of TT was used. On- and offsets were defined as points in time at which the velocity exceeded a 20 % threshold value of its maximal speed above zero velocity.

![Fig. 1. Kinematic parameter extraction (illustration only) in one production of *claque*. Top: audio; middle: vertical TT-movement; bottom: vertical TB-movement. 1 = C₁ constriction plateau; 2 = C₂ closing gesture; 3 = C₂ constriction plateau; C-center aligned relative to the acoustic burst of the word-final consonant (see text for details).](image)

As index of the temporal overlap between C₁ and C₂, two measures were evaluated: (a) the interval between the end of the constriction plateau of C₁ and the moment in time at which movement onset for C₂ is initiated, i.e. how early does gesture onset of C₂ occur within C₁ – this parallels Chitoran et al. (2002); and (b), the interval between the end of the constriction plateau of C₁ and the moment in time at which the constriction for C₂ is reached, i.e. the temporal lag between C₁ constriction off- and C₂ constriction onset. Both overlap indices are given as percentage relative to the overall cluster closure duration, defined as interval between C₁ constriction on- and C₂ constriction offset.

Finally, following Browman and Goldstein (2000), the C-centre is calculated as the means of the centres of the constrictions intervals of the individual gestures of C₁ and C₂, as indicated above. The location of the means is aligned relative to the acoustic burst of the final consonant (see right-most cursor in Fig. 1). The measurement is based
on the assumption that the vowel gesture, whose temporal properties are difficult to measure directly, is coordinated equally with the final plosive in all test items.

The statistical analysis included a separate two-way analysis of variance on each dependent variable, for each subject. The dependent variables are: duration of C1 constriction plateau, duration of C2 closing gesture, duration of C2 constriction plateau, and percentage of C1C2 onset overlap and C1C2 constriction lag. The independent variables are nature of C1 and nature of C2. When significant differences were found, Bonferroni’s tests were performed for comparison. The statistical results are summarized in Table 2; discussion of interactions, as far as significant, will be outlined in the text.

3. Results

<table>
<thead>
<tr>
<th></th>
<th>Speaker S1</th>
<th></th>
<th>Speaker S2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p f k l n</td>
<td></td>
<td>p f k l n</td>
<td></td>
</tr>
<tr>
<td><strong>C1_plateau</strong></td>
<td>(ms)</td>
<td></td>
<td>(ms)</td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>70 (12)</td>
<td></td>
<td>58 (16)</td>
<td></td>
</tr>
<tr>
<td>f</td>
<td>108* (23)</td>
<td></td>
<td>66 (15)</td>
<td></td>
</tr>
<tr>
<td>k</td>
<td>49* (12)</td>
<td></td>
<td>39* (15)</td>
<td></td>
</tr>
<tr>
<td>l</td>
<td>72 (18)</td>
<td></td>
<td>51 (16)</td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>79 (37)</td>
<td></td>
<td>46 (14)</td>
<td></td>
</tr>
<tr>
<td><strong>C2_closing</strong></td>
<td>(ms)</td>
<td></td>
<td>(ms)</td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>67 (15)</td>
<td></td>
<td>50 (16)</td>
<td></td>
</tr>
<tr>
<td>f</td>
<td>63 (16)</td>
<td></td>
<td>51 (14)</td>
<td></td>
</tr>
<tr>
<td>k</td>
<td>45* (05)</td>
<td></td>
<td>49 (17)</td>
<td></td>
</tr>
<tr>
<td>l</td>
<td>58 (19)</td>
<td></td>
<td>51 (14)</td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>59 (11)</td>
<td></td>
<td>50 (16)</td>
<td></td>
</tr>
<tr>
<td><strong>C2_plateau</strong></td>
<td>(ms)</td>
<td></td>
<td>(ms)</td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>58 (14)</td>
<td></td>
<td>51 (16)</td>
<td></td>
</tr>
<tr>
<td>f</td>
<td>50 (12)</td>
<td></td>
<td>43 (14)</td>
<td></td>
</tr>
<tr>
<td>k</td>
<td>55 (16)</td>
<td></td>
<td>41 (10)</td>
<td></td>
</tr>
<tr>
<td>l</td>
<td>46 (11)</td>
<td></td>
<td>62* (12)</td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>62 (19)</td>
<td></td>
<td>51 (14)</td>
<td></td>
</tr>
<tr>
<td><strong>onset_overlap</strong></td>
<td>(%)</td>
<td></td>
<td>(%)</td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>32 (16)</td>
<td></td>
<td>51 (15)</td>
<td></td>
</tr>
<tr>
<td>f</td>
<td>44* (12)</td>
<td></td>
<td>43* (08)</td>
<td></td>
</tr>
<tr>
<td>k</td>
<td>09* (09)</td>
<td></td>
<td>38* (15)</td>
<td></td>
</tr>
<tr>
<td>l</td>
<td>36 (20)</td>
<td></td>
<td>15* (15)</td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>21* (12)</td>
<td></td>
<td>80* (21)</td>
<td></td>
</tr>
<tr>
<td><strong>constrict_lag</strong></td>
<td>(%)</td>
<td></td>
<td>(%)</td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>-14 (10)</td>
<td></td>
<td>-01 (17)</td>
<td></td>
</tr>
<tr>
<td>f</td>
<td>03* (04)</td>
<td></td>
<td>15* (17)</td>
<td></td>
</tr>
<tr>
<td>k</td>
<td>-23* (08)</td>
<td></td>
<td>-08 (09)</td>
<td></td>
</tr>
<tr>
<td>l</td>
<td>-08 (12)</td>
<td></td>
<td>-14 (07)</td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>-14 (14)</td>
<td></td>
<td>-32 (09)</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Individual subjects means (s.d. in brackets) across C1 and C2. * indicate significance of the main effect at p < 0.01.

3.1. Kinematic durations

As expected, C1 constriction plateau differs significantly as a function of the nature of C1, being for both subjects significantly shorter in clusters starting with /k/ than starting with /p/ and /f/, respectively. Only S2 demonstrates at the same time a consistent influence of the nature of C2 on C1 constriction: C1-plateau is shorter, when followed by /n/ than when followed by /l/. For none of the speakers, the duration of the C2 closing gesture shows a significant influence of the nature of C2. For S1, however, the duration of the C2 closing gesture is significantly longer when the preceding consonant is anterior, i.e. /p/ or /f/, than when it is posterior /k/. The duration of C2 constriction plateau, finally, is not affected by the initial consonant, but significantly longer for /n/ than for /l/ for speaker S1.

3.2. Gestural overlap

The results of measurements of the overlap indices between the two consonants are listed in Table 2 and detailed in Fig. 2. Onset overlap and constriction lag were found to differ highly significantly depending on the nature of C1 and C2 for both subjects, though only S2 demonstrated a significant interaction between the two factors. While there are prominent differences in absolute timing, there are nevertheless some clear
patterns across speakers and conditions. Both speakers allow the highest overlap in /f/-
initial clusters, with S1 in this case forming C_2 closure still during C_1 constriction. /p/-
sequences, in turn, exhibit consistently higher onset and constriction overlap than the
Corresponding /k/-sequences. At the same time, the two stop sequences substantially
differ depending on the nature of C_2. For both /p/ and /k/, gesture onset and constriction
occur later for /n/ than for /l/. This timing difference is particularly substantial for S2,
who initiates C_2 gesture only after the release of C_1 in /pn/ and kn/ sequences.
Consistent differences in the two overlap indices between /n/ and /l/ in /f/-clusters are
not readily apparent.

Fig. 2. Mean values and standard deviation (in %) of C_1C_2 overlap
indices. Positive values indicate more overlap, negative values indicate
greater lag between C_1 and C_2 (see text for details).

3.3. C-center

As outlined above, we calculated the C-centre of the overall consonant sequence as the
means of the centres of the target constrictions intervals of the individual gestures of C_1
and C_2, aligned relative to the acoustic burst of the final consonant. Since this alignment
point was not available for pneu, this test item was removed from the following
analysis, a one-way Anova with C-center location as dependent variable and consonant
cluster as independent variable. As illustrated in Fig. 3, the phonetic composition of the
initial cluster had no significant effect on the location of the C-center for S1
[F(4,44)=5.149; n.s], while for S2 the statistical analysis proved to be significant
[F(4,49)=5.532; p<0.001]. As the results of the Bonferroni’s multiple comparison test
revealed this was due to the fact that only the C-center location of /kn/ was significantly
different from all other clusters.

4. Discussion

Looking first at the gestural coordination in the stop-C_2 sequences, the overlap patterns
observed in this study confirm earlier data reporting that the “place order” effect is not
restricted to stop-stop sequences, but also surfaces in stop-liquid and, as here, in stop-
nasal sequences (Kühnert & Hoole, 2006; Chitoran & Goldstein, 2006). Both speakers
allow consistently more overlap in the direction front-to-back than back-to front, i.e.
/pl/ > /kl, and /pn/ > /kn/. Most of all, this kind of order applies regardless of the fact
that the internal coordination between C_1 and C_2 was strikingly different in C/l/ and C/n/
sequences.
As argued previously, it seems to us that this pattern is not only due to factors of perceptual recoverability, which, if anything, are weaker in stop-liquid and stop-nasal productions since they do not mask the possible information about the preceding consonant in the same way as a plosive would. Rather it seems that this pattern is also caused by simple constraints of the executing motor system. The significantly longer duration of gesture onset to target achievement of /l/ and /n/ when following consonants with no lingual involvement for subject S1, seems to indicate that here the tongue is simply free to move and anticipate the upcoming articulation much earlier during the bilabial closure phase. Velar stop articulations, by contrast, require a holistic tongue movement, which are known to tightly constrain the articulator movement of the whole tongue, including the tip/blade system (e.g. Perkell, 1969). Thus, in the context of /k/C₂ productions the tongue blade cannot execute the articulation of any other lingual gesture as early as in /p/C₂ productions.

As far as the influence of the lexical frequency of the words on inter-gestural coordination is concerned, the data leave a rather mixed impression. Essentially, it cannot be said that lexical frequency by itself affects inter-gestural cohesion in a straightforward manner (cf. Davidson, 2006). For both speakers, there were no systematic differences with regard to any of the kinematic parameters investigated between /fl/ and /fn/ sequences - although /fn/ is just as limited in its occurrence as /pn/ and /kn/ and all three onsets are, for example, regarded by Rialland (1994) as being extra-syllabic in French.

The data of both speakers, however, revealed a fundamental difference in inter-gestural timing between stops followed by /l/ and stops followed by /n/, with the gestures much more spread apart in the latter case. As such, this pattern is similar to the temporal coordination for gestures that were not only phonotactically rare but illegal for speakers (Davidson & Stone, 2003). One factor causing the lesser overlap in C/n/ groups might be that nasals, unlike laterals, involve the coordination of two gestures, tongue front movement and velic opening. On the one hand, a plosive requires precise velic closure for the build up of oral pressure behind its constriction. On the other hand, the main audible trace of a word-initial stop is at its release and it is thus important that the distinctiveness of the burst is maintained. Too early nasal venting would violate both of these conditions and thus threaten the perceptual salience of initial plosives –
much more so than the incomplete constriction of /l/ does. The continuous noise of the fricative /f/, by contrast, might be perceptually somewhat more robust against some degree of velar opening and therefore tolerates more overlap with the nasal, in particular since the noise source is in front of the C₂ constriction.

Finally, despite the different coordination patterns between C₁ and C₂ outlined above, both speakers showed a surprising stability in the location of the overall C-centre across the different CC-sequences. As it were, the component gestures of the cluster productions were adjusted in such a way that the relationship between initial consonant cluster and following vowel remained globally the same. This can be seen most clearly in speaker S2’s data in which the pulling-apart of the gestures in stop-nasal productions was compensated for by a parallel shortening of C₁ constriction duration. As such, the data are compatible with Browman & Goldstein’s (2000) suggestion that there is a special coupling relationship between consonant gestures in initial CC-structures in such a way that an underlying global CV-cohesion is maintained.

References