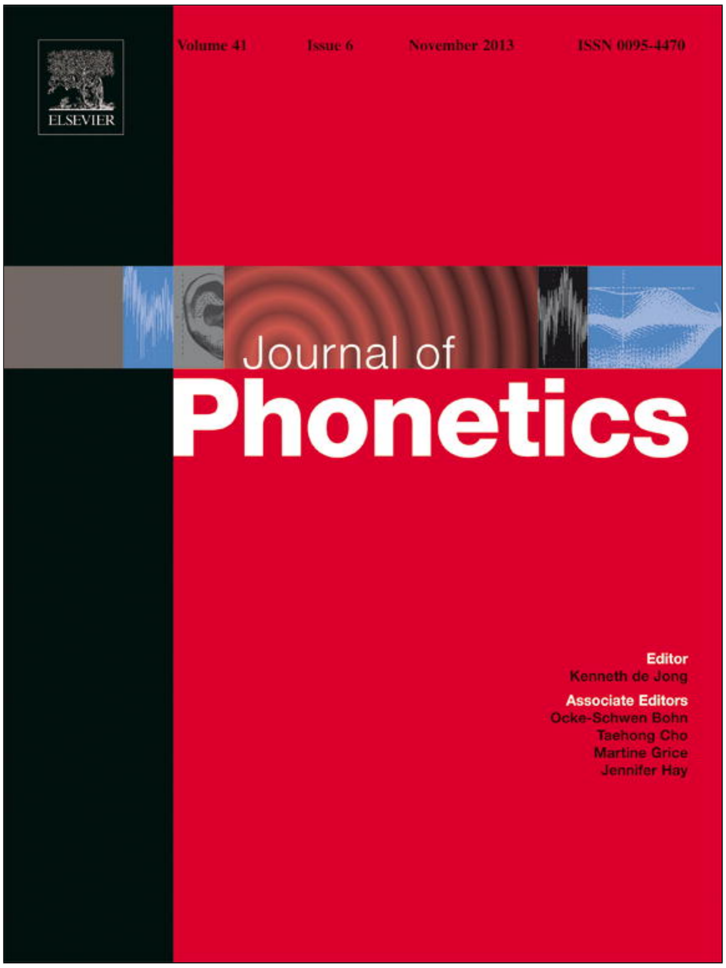
ELSEVIER
Volume 41
Issue 6
November 2013
ISSN 0095-4470
Journal of
Phonetics
Editor
Kenneth de Jong
Associate Editors
Ocke-Schwen Bohn
Taehong Cho
Martine Grice
Jennifer Hay

Journal of Phonetics 41 (2013) 546–561

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/phonetics

Articulatory coordination in word-initial clusters of German

Lasse Bombien [a,*], Christine Mooshammer [b,1], Philip Hoole [a]

[a] Institute of Phonetics and Speech Processing, Schellingstraße 3, D-80799 Munich, Germany
[b] Haskins Laboratories, 300 George St., Suite 900, New Haven, CT 06511, USA

A R T I C L E I N F O

Article history:
Received 29 December 2011
Received in revised form
11 July 2013
Accepted 26 July 2013
Available online 12 November 2013

A B S T R A C T

Intra-gestural and inter-gestural coordination in German word-initial consonant clusters /kl, kn, ks, pl, ps/ is investigated in four speakers by means of EMA as a function of segmental make-up and prosodic variation, i.e. prosodic boundary strength and lexical stress. Segmental make-up is shown to determine the extent of articulatory overlap of the clusters, with /kl/ exhibiting the highest degree, followed by /pl/, /ps/, /ks/ and finally /kn/. Prosodic variation does not alter this order. However, overlap is shown to be affected by lexical stress in /kl/ and /ps/ and by boundary strength in /pC/ clusters. This indicates that boundary effects on coordination are stronger for clusters with little inter-articulator dependence (e.g. lips + tongue tip in /pl/ vs. tongue back+tongue tip in /kl/). The results also show that the extent to which prosodic factors affect articulation interacts with the position of the affected segment in the sound sequence: In general, boundary strength strongly affects the cluster's first consonant while lexical stress influences the second consonant. This indicates that prosodic effects are strongest at their source (i.e. the boundary or the stressed nucleus) and decrease in strength with distance from their source. However, prosodic lengthening effects can reach the more distal consonant in clusters with a high degree of overlap and high inter-articulator dependence. Besides these aspects the discussion covers differences in measures of articulatory coordination.

1. Introduction

In a previous study of electropalatographic (EPG) data (Bombien, Mooshammer, Hoole, & Kühnert, 2010) we have shown that the segmental make-up of word-initial German consonant clusters plays a major role in determining the articulatory coordination of the individual consonant gestures. The most prominent differences emerged from the comparison of /kl/ with /kn/ clusters: speech movements of the tongue dorsum (for /k/) and the tongue tip (for /l/ and /n/) overlapped to a larger degree when the second consonant (C_2) was a lateral rather than a nasal (see also Hoole, Bombien, Kühnert, & Mooshammer, 2009; Kühnert, Hoole, Mooshammer, & Bombien, 2008). Furthermore, coordination is also influenced by the strength of a preceding prosodic boundary and by whether the vowel following the clusters is stressed or not: In clusters at strong prosodic boundaries and before stressed vowels articulatory overlap decreased compared to clusters at weaker boundaries and before unstressed vowels (for boundary effects see also Byrd & Choi, 2010). Crucially, however, the segmental make-up also acts as a constraint on prosodically conditioned effects: /kl/, which showed a large degree of overlap, was much less likely to undergo coordinative changes due to boundary strength or stress than /kn/, which overlaps to a much lower degree. This study aims to extend the findings of Bombien et al. (2010) by widening the range of clusters under analysis from purely lingual consonants to labial consonants as well. To accomplish this, the data were recorded by means of electromagnetic articulography (EMA) instead of EPG, since labial articulations cannot be captured with EPG pseudopalates. Again, we will address segmental make-up and prosodic variation as two influential aspects on articulatory coordination.

This study will furthermore address different means of analyzing temporal coordination in consonant clusters. In Bombien et al. (2010), we used plateau overlap as a measure of temporal organization. In order to facilitate better comparison to other studies, we will here also analyze the latency between target achievements of successive consonants. Since these measures are of very different nature we will offer a brief discussion of their benefits (see also Section 2.5).

1.1. *Considerations regarding segmental make-up*

Our expectations regarding the articulatory organization in consonant clusters are based on considering perceptual and articulatory constraints. While these constraints are at least partly conflicting, we do not aim to show that one is superior to the other. Instead we use them to formulate some

* Corresponding author. Tel.: +49 89 2180 6388; fax: +49 89 2180 5790.
E-mail addresses: lasse@phonetik.uni-muenchen.de (L. Bombien), tine@haskins.yale.edu (C. Mooshammer), hoole@phonetik.uni-muenchen.de (P. Hoole).
[1] Now at Department of German Studies and Linguistics, Humboldt University, Berlin.

0095-4470/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.wocn.2013.07.006

basic expectations as a foundation for the discussion. In the above EPG study in which /kl/ was found to be produced with a particularly high degree of overlap, prosodic variation did not influence the inter-consonantal coordination. An articulatory perspective on this might be that the tongue tip is to a certain extent independent of the tongue back such that tongue tip raising can occur relatively early during the velar occlusion. At first glance /kn/ clusters can be expected to behave likewise since both clusters involve the same lingual articulators. However, the data we presented in Bombien et al. (2010) and also in Hoole et al. (2009) showed that /kn/ exhibits substantially less overlap than /kl/, which enforces the consideration of other factors as well. We have therefore drawn on the concept of perceptual recoverability (e.g. Chitoran, Goldstein, & Byrd, 2002) which states that consonant sequences are coordinated in such a way that essential perceptual cues of one consonant will not be hidden by another consonant: The second consonant (C_2) in /kn/ is produced with a total oral occlusion and velic lowering as opposed to C_2 in /kl/ with a central occlusion and no velic lowering. The burst of the first consonant (C_1) in /kn/ must occur before oral occlusion and velic lowering of C_2 in order to prevent nasal venting during closure of C_1 and thus the loss of the burst characteristics, which are the major cue for stop place recoverability. Overlap is therefore dispreferred in this cluster. If, on the other hand, /k/ is released into the lateral /l/, both the burst characteristics and the laterality are preserved. In Bombien et al. (2010), /kn/ overlap was also found to be significantly affected by prosodic variation in terms of boundary strength with longer lags between /k/ and /n/ at higher prosodic boundaries, whereas overlap for /kl/ was stable across prosodic conditions. We have argued that higher articulatory complexity (addition of the velic gesture) might play a role in why /kn/ displays little overlap and why /kn/ is more prone to coordination changes due to prosodic variation. However, due to the limitations of the recording method (electropalatography, EPG) clusters with bilabial consonants could not be studied and only /kl/ and /kn/ could enter the systematic analysis of stop+sonorant clusters. Therefore, differences might be explained by idiosyncratic characteristics of these particular clusters: Both clusters consist of lingually articulated consonants with high biomechanical linkage.

It should be interesting to see how strong overlap is in a cluster whose articulators are more independent of each other. The most suitable candidate for such a comparison with /kl/ is /pl/ since it is the only other stop + lateral cluster in German (apart from clusters with phonologically voiced stops, such as /gl/, /bl/, which, however, appear to have a different internal structure, see Bombien & Hoole, 2013; Hoole et al., 2009; Kühnert et al., 2008). It is not entirely clear from production-based accounts, how coordination should differ between /kl/ and /pl/. On the one hand, /p/ should be expected to exhibit less coarticulatory resistance than /k/ due to stronger dorso-palatal constraints in the latter (e.g. the DAC model, Recasens, Pallarés, & Fontdevila, 1997) therefore allowing for earlier onset of C_2 and hence more overlap. Another scenario is, however, that inter-articulator independence in /pl/ allows for higher degrees of freedom regarding coordination whereas coordination relations in /kl/ are highly constrained since both consonants are lingual. This scenario predicts in general more coordination variability in /pl/ than in /kl/ which could just as likely result in a higher or a lower degree of overlap in /pl/ compared to /kl/. It is at present not clear which of these scenarios will be supported by the data.

In the comparison of /kl/ and /pl/, the order of the place of articulations in the clusters changes from back-to-front in /kl/ to front-to-back in /pl/. In the context of gestural recoverability, previous work has reported more overlap in front-to-back clusters than in back-to-front clusters (*place order effect*, Chitoran et al., 2002) and the general preference of front-to-back clusters (Byrd & Tan, 1996; Hardcastle & Roach, 1979; Kochetov, Pouplier, & Son, 2007; Zsiga, 2000) in which C_1 cues are less likely to be masked be C_2. However, the evidence provided for this preference is based on stop-stop sequences and we believe it is only applicable there since in other stop+/C/ clusters C_2 does not involve complete occlusion. Thus, there is no perceptual motivation to expect coordination differences in /kl/ vs. /pl/.

The choice of considering another C_1 place of articulation also allows for an attempt to answer the question of whether /l/ is particularly suited for overlap due to its articulatory properties. To this end two additional clusters are selected with C_2 replaced by another alveolar (/s/) such that the complete set of clusters under analysis is: /kl kn ks pl ps/. Both perceptual and articulatory considerations lead us to assume that stop+/s/ should be produced with less overlap than stop+/l/ clusters. On the perceptual side, less overlap in stop+/s/ is necessary in order to ensure a sufficiently salient stop burst, and also because /s/ requires a certain amount of stationary frication (50 ms, Jongman, 1989) to be reliably recovered. Articulatorily, the requirement of a precise formation of a medial groove (see Stone, Faber, Raphael, & Shawker, 1992) poses a sufficient constraint to make /s/ highly resistant to coarticulatory effects (Recasens et al., 1997), which might also be reflected in the temporal domain.

1.2. Prosodic variation

Prosodic theories often focus on f_0 variation and temporal changes, e.g. autosegmental metrical phonology (Beckman & Pierrehumbert, 1986; Grice & Ladd, 2000; Ladd, 2008). However, in recent years there have been attempts to integrate changes of segmental properties as a function of prosodic phrasing and of prominence into prosodic theories (e.g. Bombien, Mooshammer, Hoole, Rathcke, & Kühnert, 2007; Bombien et al., 2010; Byrd & Choi, 2010; Cho & Keating, 2009; Hoole et al., 2009; Kuzla, Cho, & Ernestus, 2007; Pierrehumbert & Talkin, 1992). Prosodic effects due to phrasal accent have been reported by quite a number of studies (e.g. Cambier-Langeveld, 2000; Cho & Keating, 2009; Cho & McQueen, 2005; Turk & White, 1999). This study does not investigate phrasal accent directly and this section will therefore focus on lexical stress and phrase boundaries. Pierrehumbert and Talkin (1992) observed that for simple CV syllables preceding phrase boundaries strengthen the syllable onset, e.g. French nasals have been shown to be produced with lower air-stream adjacent to boundaries resulting in more obstruent-like productions (Fougeron, 1998). Prominence, on the other hand, strengthens the nucleus, i.e. longer and less reduced vowels. As for boundaries, numerous studies have at least partially confirmed this view. Kuzla et al. (2007) showed that underlyingly voiced fricatives were produced with less voicing after strong boundaries than after weak boundaries. In EPG studies of several languages, Fougeron and Keating (1997) and Keating, Cho, Fougeron, and Hsu (2003) showed that many lingual consonants are produced with stronger linguopalatal contact and longer contact durations when adjacent to phrase boundaries. Some findings, however, are not easily integrated with Pierrehumbert and Talkin's view: e.g. Cho and McQueen (2005) report less VOT in fortis stops at strong prosodic boundaries than at weak. Few studies have systematically investigated the effect of more than one prosodic parameters at a time. An exception to this is an EPG study by Cho and Keating (2009) which systematically varied boundary strength, lexical stress and phrasal accent. In a variety of acoustical and articulatory measures only a small number turned out to be consistently affected by both boundary strength and prominence. The most important measure among these – and also the most relevant for this study – is the EPG seal duration within the initial consonant of the target word. Lengthening of the seal duration was observed at strong boundaries as well as in stressed and accented positions. A central concern of their study was to investigate whether effects due to accent, stress and phrasal position add up cumulatively, e.g. whether a segment lengthened by a prosodic boundary will undergo additional lengthening when subject to lexical stress and/or phrasal accent. No cumulative effect was found, i.e. the magnitude of the effect was similar regardless of whether all of the prosodic factors or only one of them was contributing. In another EPG study, Bombien et al. (2010) systematically vary the prosodic boundary preceding a consonant cluster (/#kn/ /#kl/ /#ks/ /#sk/) and the lexical stress of the syllable containing the cluster. Their results extend the findings of Cho and Keating (2009) in that they show that boundary effects most strongly affect the first consonant while stress effects are more pronounced on the second consonant. This indicates that prosodic effects induced by both boundary strength and lexical stress are of a graded nature.

The π-gesture approach (see Byrd, Kaun, Narayanan, & Saltzman, 2000; Byrd, Krivokapić, & Lee, 2006; Byrd & Saltzman, 2003) explicitly incorporates the notion of gradedness of prosodic effects. Associated with the framework of Articulatory Phonology (e.g. Browman & Goldstein, 1992), this approach formalizes effects of boundary strength in terms of a so-called *π-gesture* (the *prosodic* gesture) which acts to slow the clock-rate of an utterance. Articulatory gestures co-active with a π-gesture are slower and – as a consequence – can be stronger. The π-gesture itself is associated with prosodic boundaries and its size correlates with the strength of the prosodic boundaries. Articulatory gestures at utterance boundaries should therefore be under the influence of a strong π-gesture and should display considerable lengthening and strengthening as opposed to articulatory gestures remote from any prosodic boundary. Crucially, π-gestures are not active or inactive in a binary fashion but wax and wane continuously (Byrd & Saltzman, 2003). Peak activation co-occurs with the location of the boundary and the graded nature of prosodic effects results from the rising and falling of activation around the prosodic boundary. Evidence for gradedness has been presented using tongue tip EMA data in Byrd et al. (2006), where the closing movement for a post-boundary consonant was affected to a larger extent than the opening movement. The reverse was found for pre-boundary consonants, where the opening movement was affected more strongly than the closing movement. In both directions, the lengthening effect decreases with distance from the boundary.

As for coordination, the π-gesture approach makes predictions not only for single sounds but, importantly, also for the coordination of consecutive articulations. The slowing of the clock driving the gestural score is also hypothesized to delay the activation of speech gestures that are initiated under the influence of π-gestures. As a result, consonant clusters are expected to exhibit less articulatory overlap at strong prosodic boundaries than phrase medially (Byrd & Saltzman, 2003). Byrd and Choi (2010) were the first to investigate the paradigm of articulatory lengthening in consonant clusters. The results confirm the prediction of the π-gesture approach in that boundary-induced articulatory lengthening effects wane with distance from the boundary, i.e. in the case of initial clusters the first consonant displayed stronger lengthening effects than the second. Very similar findings have been discussed for German EPG data in Bombien et al. (2010): There, however, effects on the second consonant were only observed consistently in a cluster showing an especially high degree of overlap, /kl/. It is interesting to note that Fougeron (1998) found no effect on the second consonant in /kl/ clusters in an EPG study of French. This crosslinguistic difference is possibly due to the fact that French /kl/ clusters overlap to a significantly lower degree than German /kl/ clusters (Bombien & Hoole, 2013; Hoole et al., 2009; Kühnert et al., 2008) such that the activation of the π-gesture does not last into the second consonant. Both Byrd and Choi (2010) and Bombien et al. (2010) confirm that prosodic boundary strength not only affects segmental properties such as duration and strength of articulatory gestures, but also inter-segmental properties: Some consonant clusters at strong prosodic boundaries displayed a lower extent of gestural overlap than identical clusters in phrase medial positions.

Effects on articulation triggered by prosodic prominence are not explicitly covered by the π-gesture approach.[2] There is, however, some evidence which allows for the assumption that effects of prosodic prominence are graded in a similar fashion to that predicted by the π-gesture for boundaries. Assuming peak impact of the effect of stress to be located on the stressed syllable's nucleus (see also Mooshammer & Geng, 2008), Bombien et al. (2010) varied lexical stress (confounded with accent) in disyllabic words with initial consonant clusters (this study also varied prosodic boundary strength). Syllabic prominence resulted in lengthening effects on the second onset consonant (which is adjacent to the nucleus) rather than on the first. Furthermore, de Jong, Beckman, and Edwards (1993) demonstrated a decrease of coarticulation or an increase in hyperarticulation in CV syllables bearing accent. This study will therefore also address the question of whether effects of stress can be characterized as graded.

1.3. Predictions

Based on the models presented in the introduction the following predictions are made. In shorthand the C_1 stops /p/ and /k/ will be represented as /C/ (e.g. /Cs/ vs. /Cl/ clusters) and so will the alveolar C_2 consonants (e.g. /pC/ vs. /kC/ clusters).

1.3.1. Segmental make-up

(S1) For the difference between /kl/ and /kn/ it is assumed in accordance with both perceptually and articulatorily driven approaches (see 1.1) and the results from our EPG study (Bombien et al., 2010) that /kn/ should be produced with less overlap than /kl/.

(S2) (a) For the four-way contrast of /kl/ – /pl/ – /ps/ – /ks/, we expect larger overlap in /Cl/ than in /Cs/ clusters.
(b) Regarding the effect of different C_1 places of articulation, we expect that /pC/ clusters have more degrees of freedom regarding their coordination due to higher inter-articulator independence. Based on this assumption we expect greater variability in the degree of overlap. We have, however, no clear prediction as to whether this will be accompanied by overall more or less overlap.

1.3.2. Prosodic variation

(P1) Prosodic boundaries preceding the cluster will have the strongest impact at their centers, i.e. the peak of the π-gesture (Byrd & Saltzman, 2003). A strong effect of boundary strength on C_1 is therefore assumed. C_2, being further removed from the center of the effect, is less affected, if at all, since the effect of boundaries is considered to be of a graded nature. In terms of gestural coordination, it is expected to find less overlap (higher latency) in clusters adjacent to a strong prosodic boundary (i.e. phrase initial clusters) which is in accordance with Byrd and Saltzman (2003), Byrd and Choi (2010) and Bombien et al. (2010).

(P2) Lexical stress is assumed to be centered on the stressed syllable's nucleus, which renders C_2 closer to the point of impact than C_1. Consequently, a stronger effect is expected on C_2 than on C_1. In terms of gestural coordination, effects of stress on overlap are interpreted in the same way as effects of boundary strength. This prediction is made with some reserve, since there was only a consistent effect of stress for

[2] However, it has previously been proposed by Byrd and Saltzman (2003) and Saltzman, Nam, Krivokapić, and Goldstein (2008) that π-gestures might be applicable for prominence as well and thus account for findings related to stress and accent as demonstrated in several studies (for example Cambier-Langeveld & Turk, 1999; Cho & Keating, 2009; Turk & White, 1999).

a subset of the data in Bombien et al. (2010). However, effects of stress on the syllable onset have been observed in the literature (e.g. Bombien et al., 2007; Cho & McQueen, 2005).

1.3.3. Segmental make-up and prosodic variation

(SP1) Byrd and Choi (2010) observed that the extent to which the coordination of clusters is sensitive to prosodic variation depends on their segmental make-up. Specifically, /#sC/ (C in C_2 position = any consonant) clusters were found to be less affected than /#kl/ clusters. However, /sC/ clusters are not included in this study. Bombien et al. (2010) found stronger effects of prosodic variation on overlap in /kn/ than on /kl/ clusters. This difference was attributed to the closer timing in /kl/ clusters. In accordance with prediction (S1), it is therefore assumed that /Cl/ (C in C_1 position = p/k) clusters will exhibit less sensitivity to prosodic variation than /Cs/ clusters.

2. Method

2.1. Speech material

This study investigates the articulatory timing of the sequences /kn kl pl ks ps/ in word-initial position of one or two disyllabic German target words: one with lexical stress on the first syllable (henceforth: the stressed condition) and the other with lexical stress on the second syllable (henceforth: the unstressed condition). Table 1 displays all target words with transcriptions and translations. The inclusion of /pn/ would be highly favorable in this analysis. However, to keep the corpus to a manageable size a number of clusters had to be left out, including /pn/. Further reasons for the exclusion of /pn/ are its very low functional load in German and the difficulties to find a pair of disyllabic words with a contrast in the position of lexical stress. Such pairs were generally too exotic or technical. For the same reason the unstressed condition for the cluster /ks/ remained empty after the candidate *Xanthan* /ksan.ˈtaːn/ was rejected. Hoole et al. (2009) present data on /pn/ taken from a different corpus. Their results are in line with the results presented in this study. All target words were embedded in four types of carrier phrases to elicit different levels of prosodic boundaries before the target word. In each case, the target word – and therefore the cluster – was preceded by an unstressed open vowel ([a, ɐ]). In the *utterance initial* condition (U), the target word is the first word in the second of two consecutive sentences. The *phrase initial* condition (P) has the target word as the first word in a sub-clause. In the *list* (L) condition the target word appears as the third of four items. Finally, in the *word* condition (W), the target word is preceded by a simple word boundary. The complete speech material can be found in the appendix in Tables 8–12. Stress and accent are confounded in the data, i.e. in the stressed condition the target syllable carries a pitch accent and in the unstressed condition the target syllable precedes the pitch accent. Sentences in the *word* condition were designed such that the target words bear pre-nuclear accent.

2.2. Recordings

EMA (Electromagnetic Articulography) and EPG (Electropalatography as used in Bombien et al., 2010) are very different systems which yield very different information. The following will point out that they are complementary rather than competing methods. EMA allows for position and motion tracking of sensor coils in 3D space whereas EPG measures the contact of a pseudopalate fitted to the hard palate and the tongue. One EPG data frame is in essence nothing more than a fixed number of binary states each indicating whether a corresponding electrode on the artificial palate is in contact with the tongue at a given point in time. Information about actual articulatory movements can only sparsely be deduced from the dynamics of contact strength and location. The distribution of contacts on the artificial palate and their number (62 in the Reading EPG3 System, Hardcastle, Gibbon, & Jones, 1991) allows for a detailed analysis of linguopalatal contact in both the anterior–posterior as well as the lateral dimension of the hard palate. It should be noted that EPG measures the contact location/area on the pseudopalate but not the part of the tongue that is engaged in this contact. EMA data often have higher temporal resolution (here 200 Hz) than EPG data (100 Hz in Bombien et al., 2010).

EMA data is not limited to lingual articulation and – most decidedly – it does not inform about contact but about position. EMA point coordinates are obtained by solving complex nonlinear equations whose input is the demodulated electrical signal induced into the sensor coil by six magnetic field transmitters (see Hoole & Zierdt, 2010). Barring computing errors, EMA allows for position tracking of sensors at any given time. The system used in this investigation (Carstens AG500) uses a maximum of 12 sensors which can be positioned anywhere within the recording space and – in the case of speech recordings – attached to any accessible part of the head. EMA sensors can interfere with each other if spaced too closely. In the case of lingual articulations it is therefore necessary to decide how to place sensors on the tongue. Data density for the tongue will accordingly always be lower in EMA recordings than in EPG. Apart from a custom fitted pseudopalate for each speaker EPG requires only little effort in both the recording as well as the analysis. EPG is therefore a good choice for investigations which focus on the amount and location of linguopalatal contact. For other articulations, e.g. labial articulations, EPG is insufficient and EMA is more promising in spite of a much larger effort during recording and data processing. EMA is also appropriate for investigations focusing on precisely those articulations where linguopalatal contact is not involved, e.g. tongue lowering

Table 1
All target words with transcripts and translations listed by onset cluster and stress of the first syllable.

/kn/	/kl/	/pl/	/ks/	/ps/
Stressed				
Kneipe	Claudia	Platin	Xaver	Psalmen
/ˈknaɪ.pə/	/ˈklaʊ.dia/	/ˈplaː.tin/	/ˈksaː.vɐ/	/ˈpsal.mən/
‘pub’	(name)	‘platinum’	(name)	‘psalms’
Unstressed				
Kneipier	Klausur	Plakat		Psalmist
/knaɪ.ˈpjeː/	/klaʊ.ˈzuɐ/	/pla.ˈkaːt/		/psal.ˈmɪst/
‘pub owner’	‘written exam’	‘poster’		‘psalmist’

kinematics in open vowels. But from EMA alone it cannot be detected whether contact with the palate was made. It also does not readily provide information about the distance from the palate to a sensor.

In this study, four speakers (3 female, 1 male; age: 20–25) were recorded by means of EMA (Carstens AG500) at a data rate of 200 Hz. EMA coils were attached to the tongue (tip, mid and back), the upper and lower lips, and the jaw. Additional coils for head movement correction were located above the upper incisors, on the bridge of the nose, and behind each ear. Acoustic data were sampled at 24,000 Hz. The female subjects are of urban Bavarian origin (Munich) with a standard-like German pronunciation. The male speaker originates from Ingolstadt with quite some dialectal coloring. No speech or hearing disorders were reported and the speakers were naïve as to the purpose of the study. The speech material was presented to the speakers for reading five times in randomized order. Speakers were instructed to read the presented sentence naturally after understanding it. This instruction was introduced mostly to ensure correct production of a contrastive focus condition which is not a part of this study. The prompting system used for the recordings also triggered the simultaneous acquisition of EMA and audio data. Raw EMA amplitudes were stored on hard disk for subsequent processing. Audio data were recorded on a multichannel DAT device (Sony PC208Ax) together with a synchronization impulse for later segmentation. EMA coils were attached to the tongue tip, the tongue back (as far as the speakers tolerated) and between them to the tongue mid. Further coils were placed on the vermillion border of the lower and upper lips, on the gums below the lower incisors (jaw) and above the upper incisors (for reference). Further reference sensors were located on the bridge of the nose and on the base of the skull behind the ears where the subjacent skin was not subject to movement due to facial expressions. The reference sensors were used to assess the quality of the recording trials and, more importantly, for head movement correction.

2.3. Prosodic grouping

In order to analyze the impact of varying the prosodic boundary preceding the target clusters, all utterances were assigned to one of three prosodic groups: (1) BiG boundary (BG), (2) SMall boundary (SM), and (3) prosodic WorD boundary (WD).

In the cases of BG and SM phrase boundaries, the strength of the boundary was determined by properties of the transition between the target phrase and the immediately preceding phrase as described in Peters (2006). On the part of the preceding phrase, these properties are the presence or absence of a pause and of final lengthening as well as the quality of the boundary tone (low or high). On the part of the target phrase, a parameter *step* was labeled as up, down or equal depending on the shift of the f0 onset relative to the f0 offset of the preceding phrase.

All utterances consisting only of one phrase (i.e. where no phrase boundary preceded the target word) were assigned to the WD category (prosodic word). In accordance with Peters (2006), utterances with boundaries involving either a pause or the combination of a low boundary tone plus final lengthening plus an up-step of f0 across the boundary were classed as BG (big boundary). All utterances with a boundary that did not meet the requirements for the BG group, i.e. all remaining utterances, were assigned to the SM group (small boundary).

In the end, however, very few tokens were categorized as big boundaries. Consequently, a binary opposition was established between phrase medial (Pm) consisting of all items categorized as WD and phrase initial (Pi) consisting of both the SM and BG categories. Table 2 gives an overview of how the positional categories map the syntactically defined utterance types as listed in Appendix A.

No such grouping was carried out for lexical stress. Instead, stress was analyzed in two categories according to the canonical form of the target word.

2.4. Data processing

The physiological data were analyzed in Matlab and Emu. Semi-automatic algorithms computed the time points of articulatory landmarks within a given interval using EMA-coil trajectories (or signals derived from such trajectories, see below) and their velocity signals, i.e. their first derivatives. Different types of velocity signals were used depending on the articulator in question. The following outlines which articulators, EMA coils, trajectories and velocities were involved in the analysis of the sounds at the focus of this work.

(TB) The articulatory trajectory resulting from the EMA coil glued to a position on the tongue dorsum is referred to as TB (tongue back). Its vertical (up-down) component was used for the analysis of the velar stops /k g/ because they are produced by lifting the tongue dorsum to the soft palate. Accordingly, the velocity signal was computed as the first derivative of the vertical component only.

(TT) The TT trajectory captures the movement of the tongue tip. It was utilized for the analysis of the coronal consonants /l n s/. Coronal constrictions can involve both tongue tip lifting and tongue tip fronting. Therefore, the tangential velocity was used for landmark detection. The tangential velocity (v_t) is defined as the square root of the sum of the squared first derivatives of the trajectory's anterior–posterior and vertical dimensions (v_x, v_y): $v_t = \sqrt{v_x^2 + v_y^2}$. The lateral component can be excluded here since it does not substantially contribute to articulations of the tongue.

(LA) The bilabial stops /p b/ are special in that two active articulators, the lower and the upper lip, are involved in two dimensions, vertical and anterior–posterior. This can best be captured by computing the Euclidean distance between the respective EMA coils as a measure of lip aperture (LA). In this analysis, the Euclidean distance was calculated in a two-dimensional space since only motion in the vertical and the

Table 2
Mapping of utterance types (*U*tterance-initial, *P*hrase-initial, *L*ist item, prosodic *W*ord) to prosodic groups (Pi: phrase-initial, Pm: phrase-medial; separated for both stress categories, stressed (S) and unstressed (U)).

Prosodic grouping	U	P	L	W	Total
S					381
Pi	92	65	23	1	181
Pm	0	29	74	97	200
U					309
Pi	78	55	26	1	160
Pm	0	20	50	79	149
	Total Pi: 341; Total Pm: 349				

anterior–posterior dimension but not in the lateral dimension was of interest. The Euclidean distance d between two points p, q in a two-dimensional space (x,y) is calculated as $d(p,q)=\sqrt{(p_x-q_x)^2+(p_y-q_y)^2}$. Landmarks were detected using the first derivative of the resulting signal.

Fig. 1 shows the positioning of articulatory landmarks by reference to a trajectory and its absolute velocity. Maximum onset velocity (2), maximum constriction (4) and maximum offset velocity (6) are easily detectable from the respective signal. The other landmarks, onset and offset of the gesture (1, 7) and begin and end of the constriction plateau (3, 5), are interpolated values and represent the 20% threshold of the difference between two adjacent extrema in the velocity signal, e.g. begin of constriction plateau (3) is positioned at the 20% threshold between maximum onset velocity and maximum constriction (where velocity is zero). The 20% threshold method has been found to yield the most stable results when compared to other static and dynamic thresholds. Using zero crossings (or local minima for tangential velocities) does not yield reliable landmarks because sometimes more than one zero crossing can occur during and after the target phase (see Kroos, 1996). Fig. 2 is a real example of the word "Claudia" uttered by speaker *f01*. In this study, only the constriction plateau durations will be used as measures of intra-gestural timing. It has to be noted that the notion of a plateau in the EMA data is different from that in EPG data. In the latter, the plateau boundaries were placed with respect to the time-course of the contact percentage in the anterior or the posterior region of the EPG palate. In the EMA data, plateaus were defined with respect to the velocities of the trajectories of coils attached to the articulators. In consequence, EMA plateaus in the case of lingual articulators can be expected to be shorter because linguopalatal contact may be present before EMA plateau attainment and may also extend until after EMA plateau release. Assuming that EMA plateaus and EPG plateaus are still centered around the same target, EMA plateaus of two subsequent consonants are less likely to overlap since they are shorter than EPG plateaus.

2.5. Measurements of inter-gestural timing

Effects of segmental make-up on the temporal coordination of consonant clusters are analyzed here in terms of gestural overlap. Oliveira, Yanagawa, Goldstein, and Chitoran (2004) compared various measures of articulatory timing in consonant clusters across a number of corpora. The

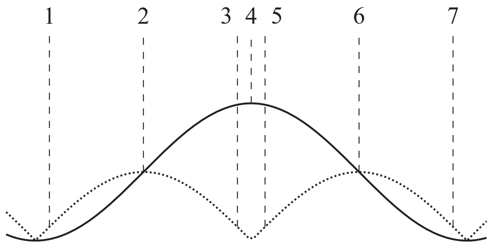

Fig. 1. Schematic display of landmark positioning. 1: onset of gesture; 2: maximum velocity in onset; 3: begin of constriction plateau; 4: maximum constriction; 5: end of constriction plateau; 6: maximum velocity in offset; 7: offset of gesture. 1, 3, 5, 7 positioned using 20% threshold, see text. Solid line: articulatory trajectory; dotted line: (absolute) velocity.

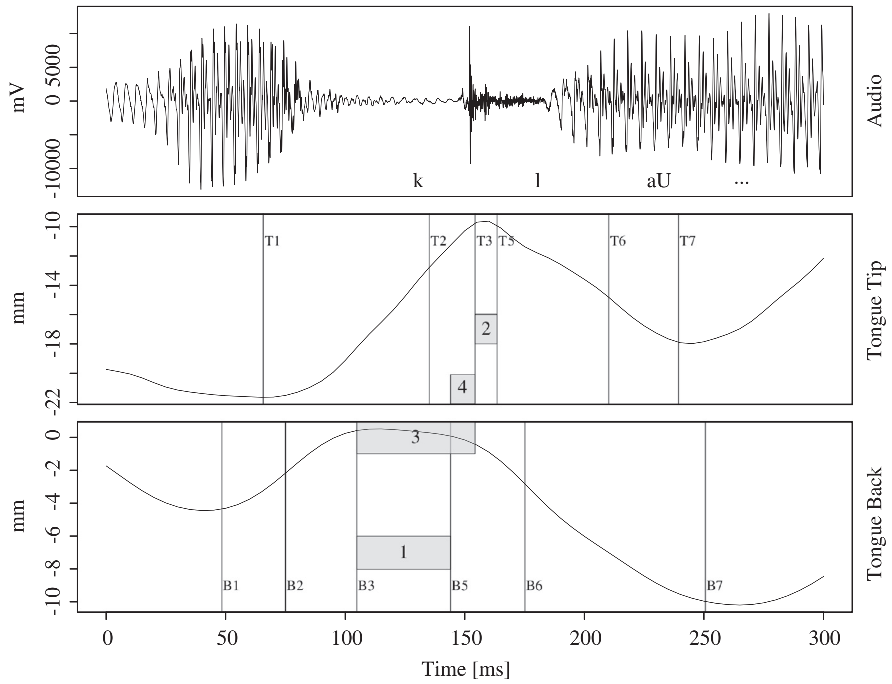

Fig. 2. Extraction of temporal parameters in a case of the word "Claudia" (phrase-medial, stressed). Vertical lines correspond to articulatory landmarks labeled according to articulator (T=tongue tip, B=tongue back) and landmark number as defined in Fig. 1 (max. constriction (T4/B4) omitted for clarity). Gray boxes mark the temporal parameters which where statistically evaluated: (1) C_1 constriction plateau, (2) C_2 constriction plateau, (3) target latency, (4) plateau lag.

most stable measures, i.e. the ones with the lowest variation coefficient, turned out to be measures of latency between the two consonants' gestures, specifically between their targets (begin of constriction plateau in Fig. 1) and usually normalized by C_1 formation duration, i.e. the interval from gestural onset to target attainment. Here, target latency will be defined as the difference between C_2 plateau onset and C_1 plateau onset (T3 and B3 in Fig. 2). A similar measure has been applied in EMA studies by e.g. Byrd and Choi (2010). The measure used in Bombien et al. (2010), plateau overlap, was defined as the interval between C_1 constriction plateau release (B5 in Fig. 2) and C_2 constriction plateau onset (T3 in Fig. 2), i.e. the overlapping section of both plateaus (positive) or the lag between them (negative). Both measures, plateau overlap and target latency, will be employed in this study for consistency with respect to Bombien et al. (2010) and for better comparability with other studies. Plateau overlap, however, will be denoted plateau lag and its sign inverted on account of mostly negative values. While both describe aspects of inter-gestural timing, plateau lag and target latency are not necessarily readily comparable. Plateau lag is based on the plateau offset of C_1 and the plateau onset of C_2 and indicates how close together the plateaus are (maybe even overlapping) or, conceivably, how close they may be. In that, plateau lag may reflect biomechanical constraints on inter-gestural timing. Target latency, on the other hand, is based on consecutive plateau onsets and disregards plateau offsets. It indicates when successive targets are achieved and may therefore reflect gestural phasing more generally. Both measures of articulatory timing will also be analyzed in normalized form. Normalized plateau lag is calculated as plateau lag divided by the interval from C_1 plateau onset to C_2 plateau offset, i.e. the entire cluster's constriction phase including a lag if present. The resulting value indicates the fraction of the constriction phase which consists of the lag (if positive) or which is overlapped (if negative). Target latency was normalized by division with C_1 formation duration as described above. In this way the speech rate can be factored out of the latency data.

2.6. Statistics

Table 3 lists the dependent variables under analysis in this study along with their predictors. The R environment (R Development Core Team, 2009) was used for statistical computing. Tests across speakers were conducted using an R extension for linear mixed-effects modeling (LME) by Bates and Maechler (2009) (see Baayen, 2008, for a detailed description) specifying subject as a random factor. Mixed modeling was preferred over repeated measures ANOVAs mainly because of two major disadvantages in the latter: (a) Repeated measures ANOVAs should be calculated on cell means to avoid inflated degrees of freedom, which requires data manipulation and entails loss of data. (b) Repeated measures ANOVAs rely on a balanced design of the data which is not the case in the present study since the unstressed condition for the cluster /ks/ is missing for reasons outlined in Section 2.1. Both these points can be disregarded with linear mixed-effects modeling.

In order to avoid collinearity between the predictors, they were coded and centered. All interactions between the predictors were included in the models if not explicitly stated otherwise. Log-likelihood tests for goodness of fit were applied in order to assess whether models were improved by allowing the slopes of the fixed effects to vary with the random factor subject. In the case of model improvement (indicated by increased log-likelihood) by inclusion of the slope for a fixed effect in the random term, additional analysis on the speaker level was conducted. In the present case, this was never necessary or justified. The actual set of necessary predictors was determined by the individual problems and will be presented and explained in the subsections of Section 3. Statistical results of the fixed effects are presented by the estimates of the regression coefficients of the model β (for the intercept, this is the grand mean in the case of centered factors) and the standard error of β.

A potential problem in LME modeling is that degrees of freedom in the denominator of the resulting F statistics cannot be estimated with sufficient reliability. Baayen (2008) proposes an anti-conservative approach which amounts to subtracting the number of levels in the relevant factors from the total number of observations. In the more complex cases in this study this number can easily reach values above 1500. Following Reubold, Harrington, and Kleber (2010), the degrees of freedom in the denominator were rather arbitrarily set to 60 to avoid obtaining significance for only small changes in the F value.

3. Results

The results will be presented in two parts. Section 3.1 will deal with the clusters /kl/ and /kn/ exclusively. Section 3.2 will then turn to the clusters /kl pl ks ps/ where C_1 place and C_2 manner of articulation are varied systematically. The reasons for this separation are twofold: (a) a separate analysis of /kl/ and /kn/ allows for a better comparison to the results in Bombien et al. (2010) and (b) dropping /kn/ in the second set of clusters enables a clean 2×2 design of the data set. Both parts will follow the same structure. In a first step, the temporal properties that are due to segmental make-up will be presented based on prosodically unmarked tokens, i.e. only tokens embedded in carrier phrase type W (i.e. simple word boundary) with stress on the first syllable of the target word are included. In the second step, the emphasis will lie on how these properties are affected by prosodic variation. The following temporal parameters will be analyzed: (1) the duration of C_1 plateau, (2) the duration of C_2 plateau, (3) the plateau lag (normalized where

Table 3
Dependent variables and predictors of the statistical models and their descriptions. See Figs. 1 and 2 for reference points relevant for measurement calculation.

Dep. var.	Description	Note
C_1 plateau duration	$=B5-B3$: C_1 plateau duration (constriction duration)	
C_2 plateau duration	$=T5-T3$: C_2 plateau duration (constriction duration)	
Target latency	$=B3-T3$: target latency: interval from C_1 plateau onset to C_2 plateau onset	Larger values mean longer latencies
Normalized target latency	$=(B3-T3)/(B3-B1)$: target latency normalized by interval from C_1 gestural onset to C_1 constriction onset (formation)	
Plateau lag	$=T3-B5$: plateau lag: lag between C_1 plateau and C_2 plateau or their overlap	Positive for lag, negative for overlap
Normalized plateau lag	$=(T3-B5)/(T5-B3)$: plateau lag normalized by the interval from C_1 plateau onset to C_2 plateau offset	
Predictor	**Description**	**Note**
C_1 place (c1p)	C_1 is always a stop but varies in place of articulation. Levels are velar /k/ or labial /p/	
C_2 manner (c2m)	C_2 is always a coronal varying in manner of articulation. Levels are fricative /s/, lateral approximant /l/ or nasal /n/	
Position (pos)	prosodic position, phrase-initial (Pi) or phrase-medial (Pm), i.e. adjacent to a phrase boundary or not	
Stress (stress)	lexical stress, on the target syllable (stressed, S) or on not on the target syllable (unstressed, U)	

applicable) and (4) the target latency (normalized where applicable). The results will be visualized as horizontal bar charts with the averaged durations of C_1 plateau and C_2 plateau juxtaposed in dark and light gray respectively. In all these figures, the averaged plateau lag is displayed as a white stretch between the plateau durations. The duration of target latency is the interval from C_1 plateau onset (left edge) to C_2 plateau onset. Each part of the results section will be headed by a table presenting the relevant statistics. These include only interactions which substantially add to the understanding of the data.

3.1. /kl/ vs. /kn/

In this section, the results for the comparison of /kl/ and /kn/ clusters as a function of segmental make-up (C_2 manner) and prosodic variation (position and stress) are presented.

3.1.1. Segmental make-up

As is evident from the overview given in Table 4, articulatory timing varies considerably as a function of C_2 manner. Both, the latency measure and the lag measure show significant responses to C_2 manner variation, with higher significance in the absolute measures as compared to the normalized data. /kl/ shows 21±4 ms less plateau lag than /kn/ ($F[1,60] = 28.47$, $p < 0.001$) or 15±4% in the relative measure ($F[1,60] = 11.00$, $p < 0.01$). Analogously, /kn/ has a longer target latency than /kl/: 26±5 ms ($F[1,60] = 26.16$, $p < 0.001$) or 21±3% in the normalized measure ($F[1,60] = 9.13$, $p < 0.01$). C_1 plateau duration is 5±3 ms longer in /kn/ than in /kl/ ($F[1,60] = 4.24$, $p < 0.05$). No such effect is found for C_2 plateau duration. Overall, this amounts to a shorter total duration in /kl/ than in /kn/ primarily due to more overlap in the former. These results are visualized in Fig. 3. They are consistent with the results presented in Bombien et al. (2010) as well as with the prediction made in S1.

3.1.2. Prosodic variation

Table 5 gives an overview of the effects of varying position and stress on the articulation of the consonant clusters /kl/ and /kn/. The prevalent coordination pattern of less plateau lag and shorter target latency observed in /kl/ clusters in Section 3.1.1 is resistant to prosodic variation. While no effect of position is detectable for plateau lag, stress significantly lengthens the lag by an average of 8±2 ms ($F[1,60] = 13.45$, $p < 0.001$). There is also an effect of

Table 4
Overview of segmental effects on articulatory measures in /kl/ vs. /kn/ clusters.

Measures	C_2 manner (l/n)
Normalized plateau lag	$l<n^{**}$
Plateau lag	$l<n^{***}$
Normalized target latency	$n>l^{**}$
Target latency	$n>l^{***}$
C_1 plateau duration	$l<n^{*}$
C_2 plateau duration	*n.s.*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05.

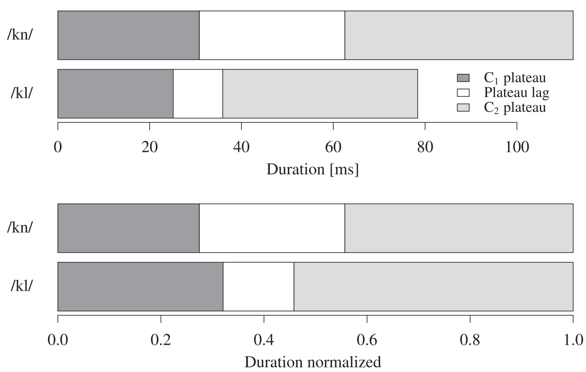

Fig. 3. Overlap pattern of /kl/ and /kn/ clusters averaged across subjects. C_1 plateau is displayed in dark gray, C_2 plateau in light gray. Plateau lag is expressed as the white stretch inbetween. Target latency corresponds to the sum of C_1 plateau and plateau lag.

Table 5
Overview of segmental effects as well as position and stress on articulatory measures in /kl/ vs. /kn/ clusters.

Measure	Position (Pi/Pm)	Stress (S/U)	C_2 manner (l/n)
C_1 plateau duration	$Pi>Pm^{***}$	*n.s.*	*n.s.*
C_2 plateau duration	Pi>Pm (/kl/) POS × C2M*	$S>U^{**}$	*n.s.*
Plateau lag	*n.s.*	$U<S^{***}$	$l<n^{***}$
Target latency	$Pi>Pm^{***}$	$S>U^{*}$	$n>l^{***}$

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05.

stress on target latency albeit not as clear as for plateau lag: stressed items have 8±3 ms longer target latency than unstressed items. Unlike plateau lag, however, target latency is quite sensitive to position: phrase-initially, the latency is 18±3 ms longer than phrase-medially ($F[1,60]=30.95$, $p<0.001$).

In Bombien et al. (2010), effects of prosodic position on plateau overlap were also found to be more pronounced in /kn/ than /kl/ with less overlap at higher boundaries. Here, there is no such effect on the plateau lag, i.e. the inverted plateau overlap. With regard to prediction P1 there appears to be more evidence for an approach favoring the specification of inter-gestural timing by segmental make-up. However, the results for target latency do not seem to fit into such an approach. The stress sensitivity of coordination measures as observed here is stronger in comparison to the findings in Bombien et al. (2010) where a consistent effect of stress on plateau overlap was only found in /kn/.

There is very clear evidence for articulatory lengthening induced by prosodic position as predicted in P1. C_1 plateau duration is on average 17±2 ms longer phrase-initially than phrase-medially ($F[1,60]=85.57$, $p<0.001$). For /kl/ but not for /kn/, the lengthening effect even spreads onto C_2 plateau duration, as evident from the interaction of position and C_2 manner ($F[1,60]=7.11$, $p<0.01$). Evidence for a graded nature of boundary effects is even stronger here based on EMA data than for the EPG data in Bombien et al. (2010). As for prediction P2, C_2 plateau durations in stressed items are indeed longer than in unstressed items (7±2 ms, $F[1,60]=9.35$, $p<0.01$). Effects of stress do not extend to C_1 plateau duration. Stress-induced lengthening of C_2 plateau duration could not be observed in the EPG data in Bombien et al. (2010). Effects of position, stress and C_2 manner are displayed in Fig. 4.

3.2. Stop + /l,s/ clusters

Similar to Section 3.1, this section describes the effects of segmental make-up and prosodic variation on consonant clusters. Here, the set of clusters consists of /kl/, /ks/, /pl/ and /ps/. Since not only C_2 manner but also C_1 place is varied, both factors will be analyzed in terms of segmental make-up.

3.2.1. Segmental make-up

Table 6 summarizes the effects of segmental make-up. In both the relative and the absolute measure, plateau lag is affected by C_2 manner. It is greater in stop+/s/ clusters (/Cs/) than in stop+/l/ clusters (/Cl/). The difference amounts to 13±5 ms ($F[1,60]=7.67$, $p<0.01$) and the relative difference is 10±5% ($F[1,60]=4.92$, $p<0.05$). In accordance with this, target latencies are 20±3 ms longer in /Cs/ clusters than in /Cl/ clusters ($F[1,60]=33.98$, $p<0.001$) or 26±7% in the normalized measure ($F[1,60]=15.56$, $p<0.001$). These findings are in agreement with prediction S 2a. The normalized measure also appears to be influenced by C_1 place of articulation: The latency is 35±7% longer in /pC/ than in /kC/ clusters ($F[1,60]=14.73$, $p<0.001$). This effect is not present for the absolute latency measure. With regard to prediction S 2b, there is thus some evidence that more degrees of freedom in /pC/ clusters appear to be accompanied by less rather than more overlap. Fig. 5 illustrates these results.

The plateau durations of both consonants depend on the segmental make-up to a high degree. C_1 plateau duration is on average 8±2 ms longer for /pC/ than for /kC/ clusters ($F[1,60]=13.38$, $p<0.001$). C_1 plateau duration is 8±2 ms longer when C_2 is the sibilant rather than the lateral ($F[1,60]=10.35$, $p<0.01$) but according to the interaction of C_1 place and C_2 manner of articulation, this is only true in /pC/ clusters ($F[1,60]=5.63$, $p<0.05$). The finding of longer plateau durations in /p/ than in /k/ is in agreement with previous findings as reported by e.g. Byrd (1993) and is considered a universal according to Maddieson (1997). More surprising is the fact that C_2 plateau duration in /pC/ clusters exceeds that

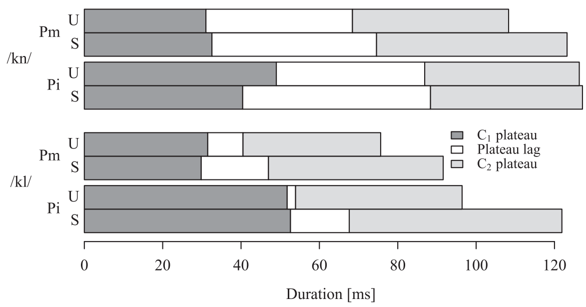

Fig. 4. Overlap pattern of /kl/ and /kn/ clusters averaged across all speakers as a function of prosodic position (phrase initial (Pi) or phrase medial (Pm)) and lexical stress (stressed (S) and unstressed (U)). C_1 plateau is displayed in dark gray, C_2 plateau in light gray. Plateau lag is expressed as the white stretch inbetween. Target latency corresponds to the sum of C_1 plateau and plateau lag.

Table 6
Overview of segmental effects on articulatory measures in /kl/, /pl/, /ks/ and /ps/ clusters. Terms with brackets indicate interactions which limit the effect, e.g. "/l/>/s/ (/k/)*" should be read as "greater for /l/ than for /s/ when C_1 is /k/."

Measure	C_1 place (p/k)	C_2 manner (l/s)
Plateau lag	*n.s.*	/l/</s/**
Normalized plateau lag	*n.s.*	/l/</s/ (/k/)* C_1 place × C_2 manner
Target latency	*n.s.*	/s/>/l/***
Normalized target latency	/p/>/k/***	/s/>/l/***
C_1 plateau duration	/p/>/k/***	/s/>/l/ (/p/)** C_1 place × C_2 manner
C_2 plateau duration	/p/>/k/*	/l/>/s/ (/p/)* C_1 place × C_2 manner

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05.

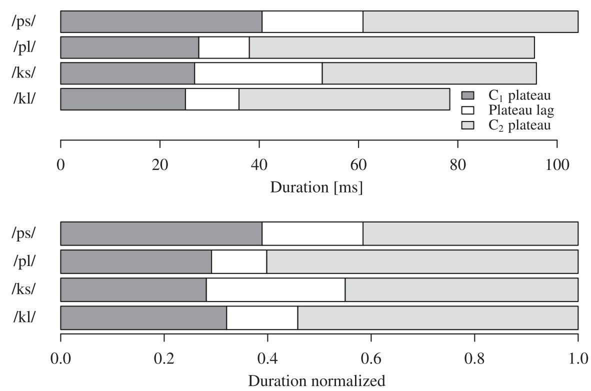

Fig. 5. Overlap pattern of /kl/, /ks/, /pl/ and /ps/ clusters averaged across subjects. C_1 plateau is displayed in dark gray, C_2 plateau in light gray. Plateau lag is expressed as the white stretch inbetween. Target latency corresponds to the sum of C_1 plateau and plateau lag.

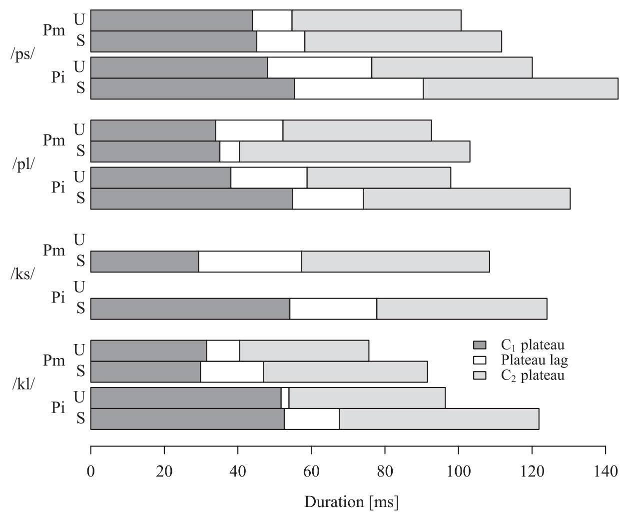

Fig. 6. Overlap pattern of /kl/, /ks/, /pl/ and /ps/ clusters averaged across all speakers as a function of prosodic position (phrase initial (Pi) or phrase medial (Pm)) and lexical stress (stressed (S) and unstressed (U)). C_1 plateau is displayed in dark gray, C_2 plateau in light gray. Plateau lag is expressed as the white stretch inbetween. Target latency corresponds to the sum of C_1 plateau and plateau lag.

of /kC/ clusters by 7±3 ms ($F[1,60]=5.15$, $p<0.05$). C_2 manner likewise affects C_2 plateau duration towards 7±3 ms shorter durations in the fricative than in the lateral ($F[1,60]=6.07$, $p<0.05$). Regarding the interaction of C_1 place and C_2 manner of articulation, the manner effect is only present in /pC/ clusters ($F[1,60]=6.89$, $p<0.05$). This behavior might be considered a perceptually conditioned effect: the /l/ plateau is shorter after /k/ than after /p/ in onset clusters because the velar's stop burst is filtered by the articulatory setup to contain laterality cues. This is not the case for /pl/ clusters since the labial stop burst does not transit the lateral constriction. Given that the timing in /pl/ and /kl/ is very similar, /l/ is lengthened after /p/ to ensure enough laterality cues in the signal. Details are presented in Fig. 5.

3.2.2. *Prosodic variation*

Details of the effects of segmental make-up in combination with variation of position and stress are presented in Fig. 6. The overview in Table 7 clearly indicates that the segmentally conditioned coordination patterns as observed in Section 3.2.1 are not affected by prosodic variation.

The variations of the plateau durations induced by segmental make-up presented in Section 3.2.1 are partially preserved and partially altered in prosodic variation. C_1 plateau duration remains longer in /ps/ than /pl/ across different prosodic conditions. However, the above observation of longer plateau durations in /p/ than in /k/ is now restricted to the phrase-medial position ($F[1,60]=18.44$, $p<0.001$). Phrase-initially the difference is not maintained.

As in Section 3.1.2, the data show very clear effects of prosodically conditioned lengthening. Phrase-initial clusters have on average 18±2 ms longer C_1 plateau durations than phrase-medial clusters ($F[1,60]=111.64$, $p<0.001$). There is also an effect of stress on the C_1 plateau duration but only in terms of an interaction with position ($F[1,60]=6.97$, $p<0.05$): Phrase-initially, stressed clusters have longer C_1 plateau durations than phrase-medial. This finding is incongruent with previous findings presented in this work. Neither Bombien et al. (2010) nor Section 3.1 of this study reports an influence of stress on C_1 plateau duration. Since the effects of stress and position add up rather than display a ceiling effect, this interaction can be interpreted as a cumulative effect of position and stress as discussed by Cho and Keating (2009).

C_2 plateau duration is more variable when viewed within prosodic variation as compared to Section 3.2.1: After /p/, C_2 plateaus are longer than after /k/ (4±2 ms, $F[1,60]=4.74$, $p<0.05$) but the C1P × POS interaction restricts this to phrase-initial occurrences ($F[1,60]=5.23$, $p<0.05$). The lateral's plateau duration is only shorter than the fricative's in unstressed clusters ($F[1,60]=5.20$, $p<0.05$). The considerable influence of stress on C_2 plateau duration found in Section 3.1.2 for the clusters /kl/ and /kn/ is confirmed here. C_2 plateau durations are on average 11±2 ms longer in stressed than in unstressed clusters ($F[1,60]=45.68$, $p<0.001$). This is in agreement with the findings in Section 3.1.2 and prediction P2.

Table 7
Overview of segmental effects as well as the effects of position and stress on articulatory measures in /kl/, /pl/, /ks/ and /ps/ clusters. Terms with brackets indicate interactions which limit the effect, e.g. "S<U (k)/(s)*" should be read as "less strong in stressed than in unstressed position when C_1 is /k/ and C_2 is /s/".

Measure	Position (Pi/Pm)	Stress (S/U)	C_1 place (k/p)	C_2 manner (l/s)
C_1 plateau duration	Pi>Pm***	S>U (Pi)** POS × STRESS	p>k (Pm)*** POS × C1P	s>l (p)** C1P × C2M
C_2 plateau duration	*n.s.*	S>U***	p>k (Pi)* POS × C1P	s>l (U)* STRESS × C2M
Plateau lag	Pi<Pm (p)*** POS × C1P	S<U (k)/(s)* C1P × STRESS C2M × STRESS	*n.s.*	l>s***
Target latency	Pi>Pm***	S>U (Pi) POS × STRESS	p>k*	s>l***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05.

There is no prosodic main effect in plateau lag. In /pC/ clusters – but not in /kC/ – there is more lag phrase-initially than phrase-medially ($F[1,60]=22.18$, $p<0.001$). Two two-way interactions emerge for the effect of stress on plateau lag: (1) Stress interacts with C_1 place of articulation such that stressed /kC/ clusters have 18±5 ms more lag than their unstressed counterparts ($F[1,60]=6.53$, $p<0.05$). This obviously only applies to /kl/ clusters since the unstressed condition is missing for /ks/ clusters. (2) Stress also interacts with C_2 manner ($F[1,60]=5.07$, $p<0.05$). Stressed stop+/s/ clusters are produced with 12±5 ms more lag than unstressed stop+/s/ clusters. Again, since the unstressed condition is missing for /ks/ clusters this can only apply to /ps/ clusters.

Prosodic effects on target latency do not interact with segmental factors. A simple main effect of position gives evidence that latencies are on average 21±3 ms longer phrase-initially than phrase-medially ($F[1,60]=70.43$, $p<0.001$). Stressed items have 8±3 ms longer latencies than unstressed items ($F[1,60]=7.90$, $p<0.01$). Due to an interaction with position, however, this effect is restricted to phrase-initial occurrences (position × stress): $F[1,60]=8.23$, $p<0.01$.

4. Summary and discussion

The results are summarized and discussed with respect to the predictions given in Section 1.3.

4.1. Segmental make-up

In S1 we expected to confirm the finding of Bombien et al. (2010) that /kl/ should be produced with more overlap than /kn/. This expectation is met in accordance with both perceptually driven as well as articulatory considerations. Longer plateau lags (i.e. the interval between C_1 and C_2 plateaus) and target latencies (i.e. the interval between C_1 and C_2 target achievements) in /kn/ indicate that in a plosive+nasal sequence the nasal is timed such that its oral occlusion and nasal venting do not prevent the emergence of crucial cues at the plosive's release. In the case of /kl/ clusters on the other hand, higher degrees of overlap are not critical since the lateral will shape the stop's release cues but not completely mask them while at the same time preserving its own place and manner cues.

From an articulatory perspective it can be argued that /kn/ is more complex in terms of articulatory coordination since it involves the velum as an articulator on another tier. While previous research on oral-velic coordination in initial nasals (e.g. Byrd, Tobin, Bresch, & Narayanan, 2009; Krakow, 1993) suggests that oral and velic gestures start simultaneously these works do not include initial plosive+nasal clusters. It can therefore not be ruled out that higher complexity is reflected in temporal coordination yielding higher latencies and plateau lags. Complexity might play a role in addition to the perceptual account offered above.

For the four-way analysis of /kl ks pl ps/ we assumed in prediction S 2a that overlap should be greater in stop+/l/ than in stop+/s/ clusters. This assumption, motivated from both perceptual and articulatory perspectives, was fully borne out in that the lag and latency measures (see Section 3.2.1) were greater for fricative than lateral C_2. Delaying the onset of the fricative serves the purpose of separating local frication of the stop's release from the fricative's stationary noise (Jongman, 1989). It further allows a sufficient time frame for the reorganization of the tongue shape in order to produce salient frication noise (Stone et al., 1992). This finding supports previous reports that /s/ is highly resistant to coarticulatory effects of neighboring segments (Recasens et al., 1997). At least in the temporal domain, however, there is no support for an inverse relation between coarticulatory resistance and aggressiveness (e.g. the window model, Keating, 1990, or the DAC model, Recasens et al., 1997).

The prediction made in prediction S 2b is that articulatory independence in /pl/ vs. /kl/ should allow for more degrees of freedom. In Section 1.1 we developed two scenarios: The first predicted shorter plateau lags in /pC/ than in /kC/ because of less coarticulatory resistance in /p/. The second allowed for both shorter and longer plateau lags and target latencies in /pC/ than in /kC/ due to a higher degree of inter-articulator independence and consequently more coordination variability. Normalized target latencies turned out to be larger in /pC/ than in /kC/ while the other coordination measures fail to reach significance. This contradicts the first scenario but is compatible with the second. The C_1 conditioned difference in normalized target latency in the analysis presented in Section 3.2.1 is not present in plateau lag. Since measures of inter-gestural coordination can obviously not be interpreted uniformly, a discussion of such measures is given below in Section 4.3. The following order arises from sorting the four clusters by the observed normalized target latency: /ps/>/pl/, /ks/>/kl/. With regard to the results presented for the two-way contrast of /kl/ and /kn/ it is interesting to see that overlap in stop+/s/ clusters patterns with the repeatedly observed overlap in /kn/ rather than /kl/. It was a conclusion of Bombien et al. (2010) that /kn/ might play a special role since it displays so much less overlap and more prosodically induced variation in overlap than /kl/. The current data allow for the reverse: stop+/l/ – and especially /kl/ – appear exceptionally stable in allowing overlap, i.e. having short plateau lags, to a rather large extent. The coordination patterns observed here suggest indeed that German /kl/ clusters are particularly well suited for parallel transmission (Mattingly, 1981) without mutual masking of acoustic cues: Both consonants can maintain their recoverability when the stop burst is filtered, but not obscured, by the lateral constriction. This notion finds support in the case of /pl/ clusters where the burst cues originate at the lips such that they cover laterality rather than being filtered by it. In this case, the lateral has to be lengthened in order to maintain recoverability.

4.2. Prosodic variation

The results concerning prosodic boundaries and lexical stress are summarized individually in this section. It was assumed in prediction P1 that strong *prosodic boundaries* exert lengthening on the adjoining segments in a graded manner, i.e. strongest on the immediately adjacent segments and less strong on more remote segments. In accord with this prediction the present data show that C_1 is significantly longer phrase-initially than phrase medially. This finding is robust across all clusters under analysis in this study. As for the graded nature of positional effects, boundary conditioned lengthening of C_2 is only found in /kl/ clusters. In all other clusters, C_2 appears to be insensitive to boundary strength. This will be further discussed below with regard to overlap differences between the clusters. Plateau lag was significantly affected by prosodic variation in /pC/ clusters (but not in /kC/ clusters). This is illustrated by the interactions of cluster type and position in Table 7 and the corresponding Fig. 6. Corresponding effects on target latency appear to be more general here since they also apply to /kC/ clusters. This issue will be discussed further in Section 4.3. Overall, these findings are in agreement with previous findings, especially as reported by Byrd and Choi (2010).

The impact of *lexical stress* was assumed to be centered on the stressed syllable's nucleus which renders C_2 closer to the point of impact than C_1 (prediction P2). Consequently, a stronger effect was expected on C_2 than on C_1. The data support this assumption across the board. C_2 plateau duration is significantly longer in stressed than in unstressed position. There is also a tendency for a graded effect of stress. C_1 plateau duration especially in /pC/ clusters exhibits some stress-induced lengthening in phrase-initial position. Effects on C_1 are overall less pronounced and less consistent than effects on C_2. Stress lengthens plateau lag in /kl/ clusters and in /ps/ clusters. In all clusters considered in Section 3.2, however, there is a significant main effect on target latency which is longer in stressed position phrase-initially. The expectations of prediction P2 concerning coordination were therefore not consistently met. This is another point to be discussed in Section 4.3.

The following paragraphs discuss the theoretical implications of the observed effects of position and stress. Fougeron (2001) found that effects of boundary strength were local to the segment immediately adjacent to the boundary. For the most part, this finding is supported by the present data as the strongest impact of prosodic position was observed on C_1 plateau duration. This is also in line with the findings presented in the EPG study in Bombien et al. (2010). The case of stress presents the mirror image in a sense. Here, too, the effect was local in being strongest on the segment closest to the stressed nucleus. This is somewhat in contrast with the findings of our EPG study (Bombien et al., 2010) as no significant effect of stress on C_2 plateau duration was found there. The only obvious reason for this difference is that articulatory trajectories are immediately available from EMA but not from EPG data. EMA data can therefore be regarded as more exact than the derived indices used in the EPG study. The results do, however, corroborate the findings of Cho and Keating (2009) and Bombien et al. (2007) where stress-induced lengthening of prenuclear consonants was observed. The π-gesture approach predicts that effects of boundary strength should be graded, i.e. waxing towards the boundary position and waning afterwards. This is confirmed by Cho and Keating (2009) in that not only parameters of the consonant but also of the vowel underwent some sort of strengthening or lengthening that could be attributed to a stronger preceding boundary. The present data does not include information about vowel production but at least for the cluster /kl/ there is evidence for the graded nature of boundary effects as not only C_1 plateau duration but also C_2 plateau duration were lengthened in phrase-initial position. This has interesting implications for the post-boundary temporal scope of the π-gesture as discussed by Byrd et al. (2006). In post-boundary /C_1 VC_2 VC_3 V/ sequences, Byrd et al. found prosodic lengthening only for the first consonant (while C_2 and C_3 undergo compensatory shortening). The strongest effect is found for C_1 closing movements followed by a less consistent effect on the opening movement of C_1. The authors conclude that boundary effects are strongest at the juncture and wane afterwards. It is not immediately clear from their study whether the π-gesture's influence spreads in terms of articulatory units (as in closing or opening movements) or in terms of time (as in ms). The current data appear to support a spreading in terms of time rather than articulatory units: Boundary effects that stretch beyond C_1 are only found in /kl/ clusters which also happen to have the shortest overall duration and the lowest values for plateau lag. In the other clusters, the C_2 plateau outreaches the scope of the π gesture and no lengthening effect is detectable.

Although its application to lexical stress is a challenge to the π-gesture (but see Byrd & Saltzman, 2003; Saltzman et al., 2008), we have assumed that effects of stress should be graded in the same way as effects of position. It is well established that the nucleus of a stressed syllable bears the strongest effect of stress since the nucleus is its primary domain. In finding such a consistent effect in C_2 plateau duration, the assumption of gradedness then finds considerable support, since it is definitely not only the nucleus which is affected. But there is some evidence that the domain of stress effects is even more widely spread since phrase-initially there is some lengthening of C_1 plateau duration in stressed clusters. The exception here is /kn/ which also happened to be the longest cluster. In this light there is less reason to assume a discontinuity of the spreading of stress effects as mentioned in the discussion of Bombien et al. (2010): In the EPG data, stress did not lengthen the plateau of C_2 but decreased the extent of overlap of C_1 and C_2. This is discontinuous in that a more distal property of the cluster was affected while the immediate neighbor was skipped. Assuming gradedness of prosodic effects, however, we would predict the opposite: a stronger impact on C_2 than on overlap. In the present study, lexical stress caused lengthening of C_2 plateau duration and plateau lags in /kC/ clusters such that a rather continuous pattern emerges now.

Finally, it should be recalled (see Section 2.1) that lexical stress and accent were confounded in the data for this study such that all target words bore a pitch accent. It is thus not certain from the data presented here which level of prosodic prominence – lexical stress or phrasal accent – is at work in the effects described above. However, preliminary results varying both accent and stress in Bombien (2011, pp. 87 & 95) indicate that stress effects as found in this study are present also in target words not bearing a pitch accent.

4.3. Implications for measures of temporal coordination

Especially in the analysis of prosodically driven variation, plateau lag and target latency yield very different results. Often, target latency varies according to the predictions – i.e. longer latencies in phrase initial position or under prominence – where no such change is visible in plateau lag. But a similar difference is also observed in the analysis concerning the segmental make-up in Section 3.2.1.

Considering C_1 plateau durations at this point sheds light on the differences between the measures for temporal coordination because usually target latency includes the entire duration of C_1 plateau. Target latency could conceivably contain only portions of C_1 plateau duration in the case of negative plateau lags. There were, however, hardly any such occurrences in the present data. Target latency is the interval of C_1 plateau onset to C_2 plateau onset while plateau lag is the interval of C_1 plateau offset to C_2 plateau onset. Table 6 shows that C_1 plateau duration clearly depends on the place of articulation in that plateau durations for /p/ are longer than those for /k/. This is in agreement with previous findings: Maddieson (1997) reports – based on observations by e.g. Byrd (1993) – that stop closures universally appear to be longer in bilabials than in velars. C_1 plateau duration is also affected by the manner of articulation of C_2. For /kC/ clusters, it is longer when C_2 is the alveolar nasal than when it is the lateral or the fricative. For /pC/ clusters, C_1 plateau duration is considerably longer when C_2 is the fricative than when it is the lateral.

As discussed above, C_1 plateau duration is consistently lengthened in phrase-initial position. Again, the duration of C_1 plateau plays a crucial role in target latency in all five clusters as target latency is lengthened as a function of prosodic position in parallel to C_1 plateau duration. This parallel variation is not found for plateau lag. In this regard, only /pC/ clusters emerge as sensitive to positional effects but never /kC/ clusters. Stress induced effects on target latency are consistent such that clusters in stressed syllables tend to have longer latencies than clusters in unstressed syllables. Interestingly, there is one interaction of C_1 place, stress, and position which fails to reach significance but the tendency is clearly visible in Fig. 6: At least in phrase initial position /pC/ clusters tend to have longer C_1 plateaus in stressed than in unstressed syllables. This indicates that a prosodically conditioned increase in target latency in /pC/ clusters is due to a parallel increase of C_1 plateau duration. In /kC/ clusters, however, it is due to a parallel increase in plateau lag.

Plateau lag seems to reflect segmental properties which are governed by principles (e.g. bio-mechanical linkages, recoverability constraints) which in turn stipulate how closely two oral articulations may follow each other and which are still not well understood. Target latency on the other hand takes these segmental properties into account as well as prosodic and C_1 specific factors. Effects on target latency can be viewed as the summed effects on C_1 plateau duration and plateau lag. In itself or on its own, target latency does not capture the segmental properties of a cluster. The point we are trying to make here, is that measures such as plateau lag are suitable for studies which are concerned with coarticulation, bio-mechanical linkage or perceptual constraints. Latency measures such as target latency, on the other hand, should serve a different purpose. They can be regarded as parameters reflecting speech planning, i.e. they might better serve to capture the initiation of consecutive gestures, e.g. in terms of *phase windows* (Byrd, 1996; Saltzman & Byrd, 2000), although a measure based on gestural onsets and not on target attainment might be more appropriate here.

The comparison of /pC/ and /kC/ clusters indicates that articulator independence plays a crucial role in inter-gestural timing. Contrary to the scenario which predicted articulator independence as in /pC/ clusters to allow for lesser plateau lags the results show the smallest values of plateau lag in /kl/ clusters in spite of the relative dependency of tongue back and tongue tip. While plateau lag in /ks/ and /kn/ clusters may be greater than in /pC/ clusters, overlap remains unaffected by positional variation in all /kC/ while in /pC/ it varies as a function of position (lesser in initial position). This finding is compatible with *phase windows* (Byrd, 1996) according to which temporal coordination is primarily determined by the type of gestures involved (/VC/, /CC/, /CV/) and by syllable position (e.g. /C#C/ vs. /#CC/). These factors constrain the totally available range of phasing relations (0–360°) to more limited *phase windows* which can be further narrowed by *influencers* such as prosodic factors and segment identity and their interactions. In this framework, our data indicate that independence of articulators (as in /pC/ clusters) seems to allow for more variation (and wider windows) while strong dependence (as in /kC/ clusters) appears to constrain the coordination possibilities to narrow windows. Prosodic position as an

Table 8
Utterances for cluster /kl/. Target words highlighted in bold.

Stress on first syllable	
Utterance initial	*Thomas studiert in Fulda.* ***Claudia*** *geht noch zur Schule*
	'Thomas goes to college in Fulda. Claudia ist still in school'
Phrase initial	*Olga sagt immer*, ***Claudia*** *sei zu jung*
	'Olga always says that Claudia is too young'
List	*Thomas, Peter*, ***Claudia*** *und Elke fahren in den Süden*
	'Thomas, Peter, Claudia and Elke are driving south'
Word initial	*Gestern war* ***Claudia*** *noch frisch und gesund*
	'Yesterday, Claudia was still OK'
Stress on second syllable	
Utterance initial	*Die Arbeit war super.* ***Klausur*** *und mündliche Prüfung waren nicht so toll*
	'The thesis was great. Written and oral exams were not as good'
Phrase initial	*Tine sagt immer*, ***Klausur*** *schreiben macht Spaß*
	'Tine always says it's fun to write exams'
List	*Hausarbeit, Wetter*, ***Klausur*** *und Erkältung machen schlechte Laune*
	'Housework, weather, written exams and a cold cause sulkiness'
Word initial	*Morgen muss sie weder* ***Klausur*** *noch Examen schreiben*
	'Tomorrow she neither has to write a test nor an exam'

Table 9
Utterances for cluster /kn/. Target words highlighted in bold.

Stress on first syllable	
Utterance initial	*Peter ist Fussballtrainer.* ***Kneipe*** *und Stadion sind sein Leben*
	'Peter is a football coach. Pub and stadium are his life'
Phrase initial	*Thomas sagt immer*, ***Kneipe*** *oder Café machen zu viel Arbeit*
	'Thomas always says a pub or a coffee shop are too much work'
List	*Restaurant, Bar*, ***Kneipe*** *und Disco wollen sie heute noch besuchen*
	'The plan to visit a restaurant, a bar, a pub and a disco today'
Word initial	*Sie arbeitet in einer* ***Kneipe*** *als Kellnerin*
	'She works in a pub as a waitress'
Stress on second syllable	
Utterance initial	*Walter trinkt gerne Vodka.* ***Kneipier*** *ist sein Traumberuf*
	'Walter likes Vodka. He dreams of being a pub owner'
Phrase initial	*Peter sagt immer*, ***Kneipier*** *ist ein schöner Beruf*
	'Peter always says that pub owner is a nice job'
List	*Koch, Kellner*, ***Kneipier*** *oder Barkeeper würde er gern werden*
	'He would like to be cook , waiter, pub owner or barkeeper'
Word initial	*Er wollte immer* ***Kneipier*** *werden*
	'He always wanted to be a pub owner or a bar keeper'

influencer on the size of the phase window (see also Cho, 2004) appears to be ineffective in clusters with narrow phase windows such that /kC/ clusters show very little plateau lag variation due to prosodic position.

5. Conclusion

The main conclusions of this study can be drawn from the way prosodic variation affects intra-gestural properties on the one hand and inter-gestural properties on the other. Inter-gestural properties such as plateau lag and target latency analyzed under prosodic variation allow for the conclusion that the gestural coordination in initial consonant clusters is stipulated nonuniformly by the segmental make-up. It is obvious that both perceptual and production-based factors crucially contribute to accounting for the patterns observed here. Segmental make-up also appears to determine the amount of coordination variation due to prosodic variation. /kC/ clusters are less variable in this regard than /pC/ clusters. It is argued that higher variability in /pC/ clusters is due to higher inter-articulator independence.

Table 10
Utterances for cluster /ks/. Target words highlighted in bold.

Stress on first syllable	
Utterance initial	*Volker studiert in Jena.* ***Xaver*** *geht noch zur Schule* 'Volker goes to college in Jena. Xaver is still in school'
Phrase initial	*Walter sagt immer,* ***Xaver*** *sei zu jung* 'Walter always says that Xaver is still too young'
List	*Inge, Walter,* ***Xaver*** *und Elke fahren in den Süden* 'Inge, Walter, Xaver and Elke are driving south'
Word initial	*Am Montag war* ***Xaver*** *noch gesund* 'On monday Xaver was still well'

Table 11
Utterances for cluster /pl/. Target words highlighted in bold.

Stress on first syllable	
Utterance initial	*Olgas Ring ist aus Silber.* ***Platin*** *war ihr zu teuer* 'Olgas Ring is made of silver. She found platinum too expensive'
Phrase initial	*Olga sagt immer,* ***Platin*** *sei zu teuer* 'Olga always says that platinum is too expensive'
List	*Gold, Silber,* ***Platin*** *und Kupfer könnte er gut gebrauchen* 'Gold, silver, platinum and copper would come in handy for him'
Word initial	*Es muss nicht immer* ***Platin*** *oder Gold sein* 'It does not always have to be platinum or gold'
Stress on second syllable	
Utterance initial	*Anna ist voller Eifer.* ***Plakat*** *oder Poster gestaltet sie gern* 'Anna is very eager. She likes to design placards and posters'
Phrase initial	*Volker sagt immer,* ***Plakat*** *und Poster sei dasselbe* 'Volker always says that placards and posters are the same'
List	*Starschnitt, Poster,* ***Plakat*** *und Autogramm hat sie schon* 'She already has a standup, a poster, a placard and an autograph'
Word initial	*Anna will ein lila* ***Plakat*** *an die Wand hängen* 'Anna wants to hang up a purple placard, not a yellow one'

Table 12
Utterances for cluster /ps/. Target words highlighted in bold.

Stress on first syllable	
Utterance initial	*Elke singt gerne Lieder.* ***Psalmen*** *singt sie auch* 'Elke likes to sing songs. She also sings psalms'
Phrase initial	*Thomas sagt immer,* ***Psalmen*** *seinen altmodisch* 'Thomas always says that psalms are old-fashioned'
List	*Sprüche, Lieder,* ***Psalmen*** *und Verse kann sie auswendig* 'She knows quotations, songs, psalms and verses by heart'
Word initial	*David hat viele der* ***Psalmen*** *und Lieder verfasst* 'David has composed many of the psalms and songs'
Stress on second syllable	
Utterance initial	*David war König von Juda.* ***Psalmist*** *war er auch* 'David was the king of Juda. He also was a psalmist'
Phrase initial	*Peter sagt immer,* ***Psalmist*** *wäre er gern* 'Peter always says that he would like to be a psalmist'
List	*Hirte, Kämpfer,* ***Psalmist*** *und Knig ist David gewesen* 'David was a shepard, a fighter, a psalmist, and a king'
Word initial	*Im Alter ist er* ***Psalmist*** *und Sänger gewesen* 'In old age he was a psalmist and a singer'

Concerning intra-gestural properties, C_1 plateaus are consistently lengthened phrase-initially while C_2 plateaus are longer in stressed syllables than in unstressed syllables. The impact of both position and stress is therefore strongest on the immediately adjacent consonant. The results provide sound evidence for the graded nature of both prosodic factors provided the property in question (C_1 or C_2 plateau) be in the range of the prosodic effect. This is especially the case for /kl/ whose summed plateaus and overlap yield the shortest total duration observed among all clusters investigated in this study. Its close coordination may account for the fact that in /kl/ both consonant plateaus are affected by both position and stress. It would be interesting to see whether these results can be replicated for other similarly closely coordinated clusters. This study has put a strong focus on timing in initial consonant clusters and therefore only durational measures were presented. An analysis of spatial effects induced by prosodic and segmental variation would be a welcome addition. Such a study has the potential of shedding light on the differences that are due to manner of articulation to a much larger extent than a purely durational study can.

Acknowledgments

This work was supported by the German Research Council's Grants HO 3271/3-1 and MO 1687/1-1. The second author acknowledges the NIH **NIDCD DC008780 Grant. We thank Susanne Waltl for assistance during data acquisition and Tamara Rathcke for prosodic annotations. We are also very grateful to three anonymous reviewers and Taehong Cho who provided us with very helpful and constructive comments on previous versions of this paper.

Appendix A. Complete speech material

See Tables 8–12.

References

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R* (1st ed.). Cambridge: Cambridge University Press.
Bates, D., & Maechler, M. (2009). lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-32.
Beckman, M., & Pierrehumbert, J. (1986). Intonational structure in Japanese and English. *Phonology Yearbook*, *3*, 255–309.
Bombien, L. (2011). Segmental and prosodic aspects in the production of consonant clusters—On the goodness of clusters. Ph.D. thesis, LMU Müchen: Fakultät für Sprach- und Literaturwissenschaften. Available at ⟨http://edoc.ub.uni-muenchen.de/12840/1/Bombien_Lasse.pdf⟩.
Bombien, L., & Hoole, P. (2013). Articulatory overlap as a function of voicing in French and German consonant clusters. *The Journal of the Acoustical Society of America*, *134*, 539–550, http://dx.doi.org/10.1121/1.4807510.
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*, *38*, 388–403.
Bombien, L., Mooshammer, C., Hoole, P., Rathcke, T., & Kühnert, B. (2007). Articulatory strengthening in initial German /kl/ clusters under prosodic variation. In *Proceedings of the 16th International Conference of Phonetic Sciences* (pp. 457–460). Saarbrücken.
Browman, C., & Goldstein, L. (1992). Articulatory phonology: *An overview*. *Phonetica*, *49*, 155–180.
Byrd, D. (1993). 54,000 American stops. *UCLA Working Papers in Phonetics*, *83*, 97–116.
Byrd, D. (1996). A phase window framework for articulatory timing. *Phonology*, *13*, 139–169.
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. In: C. Fougeron, B. Kühnert, M. d'Imperio, & N. Vallé (Eds.), *Papers in Laboratory Phonology 10: Variation, Detail and Representation* (pp. 31–60). Berlin/New York: Mouton de Gruyter.
Byrd, D., Kaun, A., Narayanan, S., & Saltzman, E. (2000). Phrasal signatures in articulation. In: M. Broe, & J. Pierrehumbert (Eds.), *Papers in Laboratory Phonology 5: Acquisition and the Lexicon* (pp. 70–87). Cambridge: Cambridge University Press.
Byrd, D., Krivokapić, J., & Lee, S. (2006). How far, How far: *how long: On the temporal scope of prosodic boundary effects*. *Journal of the Acoustical Society of America*, *120*, 1589–1599.
Byrd, D., & Saltzman, E. (2003). The elastic phrase: *Modeling the dynamics of boundary-adjacent lengthening*. *Journal of Phonetics*, *31*, 149–180.
Byrd, D., & Tan, C. C. (1996). Saying consonant clusters quickly. *Journal of Phonetics*, *24*, 263–282.
Byrd, D., Tobin, S., Bresch, E., & Narayanan, S. (2009). Timing effects of syllable structure and stress on nasals: *A real-time MRI examination*. *Journal of Phonetics*, *37*, 97–110.
Cambier-Langeveld, T. (2000). Temporal marking of accent and boundaries. Ph.D. thesis, University of Amsterdam.
Cambier-Langeveld, T., & Turk, A. E. (1999). A cross-linguistic study of accentual lengthening: *Dutch vs. English*. *Journal of Phonetics*, *27*, 255–280.
Chitoran, I., Goldstein, L. M., & Byrd, D. (2002). Gestural overlap and recoverability: Articulatory evidence from Georgian. In N. Warner & C. Gussenhoven (Eds.), Laboratory Phonology 7 (Vol. 7, pp. 419–447).
Cho, T. (2004). Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics*, *32*, 141–176.
Cho, T., & Keating, P. (2009). Effects of initial position versus prominence in English. *Journal of Phonetics*, *37*, 466–485.
Cho, T., & McQueen, J. M. (2005). Prosodic influences on consonant production in Dutch: *Effects of prosodic boundaries, phrasal accent and lexical stress*. *Journal of Phonetics*, *33*, 121–157.
Fougeron, C. (1998). Variations articulatoires en début de constituants prosodiques de différents niveaux en francais. Thèse de doctorat Université Paris III.
Fougeron, C. (2001). Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics*, *29*, 109–135.
Fougeron, C., & Keating, P. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America*, *101*, 3728–3740.
Grice, M., & Ladd, R. (2000). On the place of phrase accents in intonational phonology. *Phonology*, *17*, 143–185.
Hardcastle, W., & Roach, P. (1979). An instrumental investigation of coarticulation in stop consonant sequences. In: H. Hollien, & P. Hollien (Eds.), *Current Issues in the Phonetic Sciences (Current Issues in Linguistic Theory, 9)* (pp. 531–540). Amsterdam: John Benjamins.
Hardcastle, W. J., Gibbon, F. E., & Jones, W. (1991). Visual display of tongue–palate contact: *Electropalatography in the assessment and remediation of speech disorders*. *British Journal of Disorders of Communication*, *26*, 41–74.
Hoole, P., Bombien, L., Kühnert, B., & Mooshammer, C. (2009). Intrinsic and prosodic effects on articulatory coordination in initial consonant clusters. In: G. Fant, H. Fujisaki, & J. Shen (Eds.), *Frontiers in Phonetics and Speech Science* (pp. 275–286). Beijing: Commercial Press.
Hoole, P., & Zierdt, A. (2010). Five-dimensional articulography. In: B. Maassen, & P. van Lieshout (Eds.), *Speech motor control: new developments in basic and applied research* (pp. 331–350). Oxford: Oxford University Press.
de Jong, K., Beckman, M. E., & Edwards, J. (1993). The interplay between prosodic structure and coarticulation. *Language and Speech*, *36*, 197–212.
Jongman, A. (1989). Duration of frication noise required for identification of English fricatives. *Journal of the Acoustical Society of America*, *85*, 1718–1725.
Keating, P., Cho, T., Fougeron, C., & Hsu, C. (2003). Domain-initial articulatory strengthening in four languages. In *Laboratory Phonology 6* (pp. 143–161). Cambridge: Cambridge University Press.
Keating, P. A. (1990). The window model of coarticulation: *Articulatory evidence*. In: J. Kingston, & M. E. Beckman (Eds.), *Papers in Laboratory Phonology 1: Between the Grammar and Physics of Speech* (pp. 470–541). Cambridge: Cambridge University Press.
Kochetov, A., Pouplier, M., & Son, M. (2007). Cross-language differences in overlap and assimilation patterns in Korean and Russian. In J. Trouvain, & W. Barry (Eds.), *Proceedings of the 16th International Congress of the Phonetic Sciences* (pp. 1361–1364). Saarbrücken, Germany.
Krakow, R. A. (1993). Nonsegmental influences on velum movement patterns: *Syllables, sentences, stress, and speaking rate*. In: M. K. Huffmann, & R. A. Krakow (Eds.), *Nasals, nasalization, and the velum (Phonetics and Phonolog*, Vol. 5 (pp. 87–116). San Diego: Academic Press, Inc.
Kroos, C. (1996). Kinematische Analyse der Gespanntheitsopposition im Standarddeutschen. Master's thesis Institut für Phonetik und sprachliche Kommunikation der Ludwig-Maximilians-Universitat München.

Kuzla, C., Cho, T., & Ernestus, M. (2007). Prosodic strengthening of German fricatives in duration and assimilatory devoicing. *Journal of Phonetics*, *35*, 301–320.

Kühnert, B., Hoole, P., Mooshammer, C., & Bombien, L. (2008). Influences of manner and voicing on articulatory coordination in German and French initial consonant clusters. *Journal of the Acoustical Society of America*, *123*, 3740.

Ladd, D. R. (2008). *Intonational Phonology. Cambridge Studies in linguistics*. Cambridge University Press.

Maddieson, I. (1997). Phonetic universals. In: W. J. Hardcastle, & J. Laver (Eds.), *Handbook of phonetic sciences* (pp. 619–639). Oxford: Blackwell.

Mattingly, I. G. (1981). Phonetic representation and speech synthesis by rule. In: T. Myers, J. Laver, & J. Anderson (Eds.), *The cognitive representation of speech* (pp. 415–420). Amsterdam: North Holland.

Mooshammer, C., & Geng, C. (2008). Acoustic and articulatory manifestations of vowel reduction in German. *Journal of the International Phonetic Association*, *38*, 117–136.

Oliveira, L., Yanagawa, M., Goldstein, L., & Chitoran, I. (2004). Towards standard measures of articulatory timing. *Journal of the Acoustical Society of America*, *116*, 2643–2644.

Peters, B. (2006). Form und Funktion prosodischer Grenzen im Gespräch (Form and function of prosodic boundaries in conversation). Ph.D. dissertation, Christian-Albrechts-Universität Kiel.

Pierrehumbert, J., & Talkin, D. (1992). Lenition of /h/ and glottal stop. In: G. J. Docherty, & D. R. Ladd (Eds.), *Papers in Laboratory Phonology II: Gesture, segment, prosody* (pp. 90–127). Cambridge: Cambridge University Press.

R Development Core Team. (2009). *R: A language and environment for statistical computing*. Austria: R Foundation for Statistical Computing Vienna ISBN: 3-900051-07-0.

Recasens, D., Pallarés, M. D., & Fontdevila, J. (1997). A model of lingual coarticulation based on articulatory constraints. *Journal of the Acoustical Society of America*, *102*, 544–561.

Reubold, U., Harrington, J., & Kleber, F. (2010). Vocal aging effects on f0 and the first formant: *A longitudinal analysis in adult speakers. Speech Communication*, *52*, 638–651.

Saltzman, E., & Byrd, D. (2000). Task-dynamics of gestural timing: *Phase windows and multifrequency rhythms. Human Movement Science*, *19*, 499–526.

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

Stone, M., Faber, A., Raphael, L. J., & Shawker, T. H. (1992). Cross-sectional tongue shape and linguopalatal contact patterns in [s], [ʃ] and [l]. *Journal of Phonetics*, *20*, 253–270.

Turk, A. E., & White, L. (1999). Structural influences on accentual lengthening in English. *Journal of Phonetics*, *27*, 171–206.

Zsiga, E. C. (2000). Phonetic alignment constraints: *Consonant overlap and palatalization in English and Russian*. *Journal of Phonetics*, *28*, 69–102.