An EPG study of initial /kl/ clusters in varying prosodic conditions in German

Lasse Bombien[1*], Christine Mooshammer[2], Phil Hoole[1], Barbara Kühnert[3], Jennifer Schneeberg[2]

[1]Ludwig-Maximilians-Universität München
Institut für Phonetik und sprachliche Kommunikation
Schellingstraße 3 – D-80799 München – Germany

[2]Christian-Albrechts-Universität Kiel
Institut für Phonetik und digitale Sprachverarbeitung
D-24098 Kiel – Germany

[3]Université Paris 3 – Sorbonne Nouvelle
Institut du Monde Anglophone
5, rue de l'Ecole de Médicine – F-75006 Paris – France

lasse@phonetik.uni-muenchen.de, timo@ipds.uni-kiel.de

hoole@phonetik.uni-muenchen.de,barbara.kuhnert@uni-paris3.fr

jennifer.schneeberg@web.de

Abstract. *Articulatory strengthening as a function of prosodic variation has so far been largely investigated for singleton consonants and vowels only. The aim of this study is to analyze effects of prosodic variation on the word intial cluster /kl/ by varying lexical word stress, accent and prosodic boundary levels. Therefore, three subjects were recorded by means of Electropalatography. Derived temporal parameters provide evidence corroborating the assumption that /k/ - which was closer to the boundary - was more affected by boundary strength than the more distal segment /l/, while /l/ tended to be stronger influenced by lexical stress than /k/. At lower prosodic boundaries, the consonants overlapped to a greater degree as compared to higher prosodic boundaries.*

1. Introduction

Prosodic structure is not only marked by durational and intonational means but also by so-called articulatory strengthening. This study aims at a better understanding of the interplay between the segmental tier and higher levels in the prosodic hierarchy by analyzing articulatory principles governing the internal organization of initial consonant clusters under varying prosodic conditions. Specifically the question is which parts or properties of the initial cluster /kl/ are affected, the first consonant, the second or the

*Supported by DFG grants HO 3271/3-1 and MO 1687/1-1. We also thank our colleagues from ZAS in Berlin, especially Jörg Dreyer and Susanne Fuchs, and Christine Baran for help with labelling.

internal organization of the cluster, by varying lexical word stress, sentence accent and prosodic boundary strength. This will be addressed here by means of electropalatographic recordings. In the last two decades articulatory strengthening as a means for prosodic phrase marking or prominence enhancement has been studied for a number of languages, segments and hierarchy levels (see e.g. Keating et al. 2003; Cho and McQueen 2005; Beckman and Edwards 1994). Articulatory strengthening involves several temporal and spatial supralaryngeal enhancement strategies, such as longer and larger gestures and less coarticulation across boundaries and at higher prosodic levels. However, the empirical basis is limited to single segments and gestures whereas cluster have received very little attention. Based on results mainly on singleton consonants from the literature our hypotheses on how the prosodic hierarchy affects the realization of the constituents of word initial clusters are sketched in Figure 1. Theoretically, two possibilities are plausible for the gestural coordination within the initial clusters: no change in overlap or less overlap. No changes in overlap could be expected if the temporal coordination of clusters is entirely determined by language-specific constraints (Gafos, 2002) or by perceptual demands (Browman and Goldstein, 2000; Wright, 2004)[1]. A reduction of overlap at higher prosodic levels would be predicted by the H&H theory (Lindblom, 1990), as an enhancement of the distinctiveness of adjacent segments, and by the π-gestures approach (Byrd and Saltzman, 2003), as a local slowing down of the clock. In this case the composite gestures of a cluster would be only loosely coordinated.

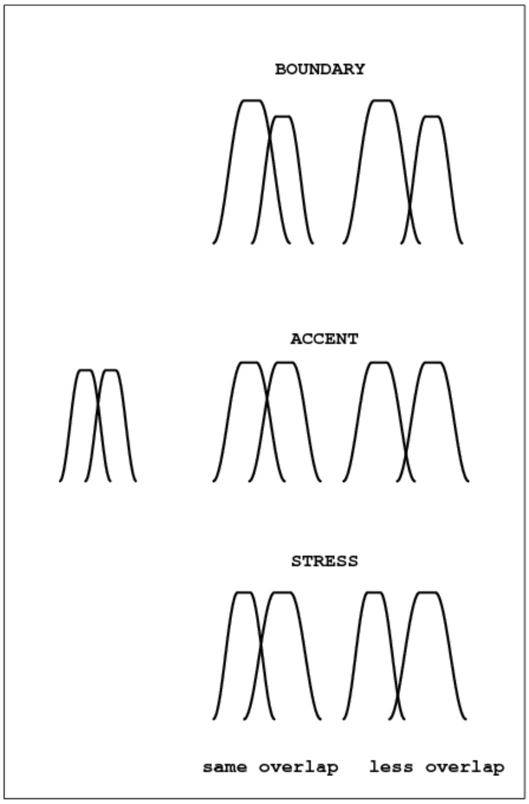

Figure 1. Schematic representations of prosodic strengthening effects on word initial consonant clusters.

For the different kinds of prosodic variations we have the following predictions: Regarding prosodic phrase marking, results from the literature indicate that the onset of a postboundary CV syllable is more extensively lengthened than the following nucleus. Spatial effects are restricted to segments adjacent to the boundary. To our knowledge complex syllable structures were only addressed by Fougeron (1998) for French and by Byrd and Choi (2006) for American English. Both studies agree regarding the domain of strengthening, mainly the element closest to the boundary, and the domain of lengthening, showing that the segment further apart from the boundary is lengthened to a lesser degree. However, concerning the internal organization of the clusters, the two studies

[1]These authors do not explicitely exclude that the internal organization of clusters is sensitive to prosodic variation

contradict each other with no boundary effect for gestural overlap measured by means of EPG contact patterns in French and less overlap measured by means of EMMA in American English, even though both studies focused on /#kl/ clusters. These two possibilities, same or less overlap, are shown on top in Figure 1 with a hypothesized lengthening of the first element of the cluster adjacent to higher prosodic boundaries.

Since the domain of sentence accent is the whole word (see Turk and White 1999), both elements of the cluster should be equally affected. Furthermore results from the literature (cf. Beckman and Edwards 1994; Mooshammer and Harrington 2005) indicate that accent will be more consistently distinguished by temporal changes as compared to spatial (see middle row of Figure 1). We speculate that for stress the strongest effect will be on the vowel (cf. Mooshammer and Fuchs 2002) and that the strengthening effect will decrease with the distance from the vowel. Therefore the second consonant in a cluster will be more strongly affected as compared to the first (see Fig. 1, bottom). In a very recent study Keating and Cho (2005) found that stress effects on consonants are restricted to the temporal domain and more extensive for stress as compared to accent.

2. Method

In order to test these hypotheses, three native speakers of German (two females (*SF, MV*), one male (*DP*)) were recorded by means of EPG. *SF* originates from Saxonia but has been living in Berlin for 10 years, and *DP* was born in Berlin. *MV* is of northern German origin. None of them reported any speech or hearing disorders. Word initial /kl/-clusters were investigated in two words: *Claudia* [ˈklaʊdi̯a] with stress on the first syllable and *Klausur* [klaʊˈzuɐ] (‘written test’) with stress on the second syllable. For both words, utterances were designed for four prosodic contexts similar to the approach of Cho and McQueen (2005): *utterance initial (U)*, *phrase initial (P)*, *list item (L*, in non-final position of the list), *word (W)*, see Table 1. In all utterances, the word preceeding the target word was chosen to end on an open vowel (either [a] or [ɐ]), to ensure a minimum linguopalatal contact before the cluster onset. We also recorded words with initial /sk/ and /kn/ clusters and a fifth prosodic condition, namely deaccentuation, which will not be included in this study. The speakers were recorded in Berlin (speakers *SF, DP*) and in Kiel (speaker *MV*) using the *Reading EPG system* (Hardcastle et al., 1989). In Berlin, audio data were recorded on digital audio tape (DAT, sampling frequency 48 kHz) and later synchronized with the EPG data (100 Hz). The audio data were downsampled to 16 kHz. A different setup was used in Kiel, yielding one EPG file (200 Hz) and one audio file (20 kHz) per utterance. All utterances were repeated ten times in randomized order. Acoustical labelling was performed using the *Munich Automatic Segmentation System MAUS* (Schiel, 1999). All label files were then converted to be usable for the *EMU speech database system* (Cassidy and Harrington, 1996). A custom *EMU* module was prepared to calculate two EPG indices: The *anteriority index* (AI) indicating the amount of anterior contact (rows 1 - 5) and the *dorsopalatal index* (DI) as a measure of dorsopalatal contact (rows 6 - 8). DI was refined using speaker-dependent region profiles as proposed by Byrd et al. (1995). Both indices were calculated by summing all contacts defined for the respective region and deviding the result by the number of contacts to yield a value between 0 and 1. The module then automatically determined a number of events defined for the articulatory gestures: the point of maximum contact (*max*) as well as onset (*on*) and offset

Syl.	Pos.	Utterance
stressed (s)	U	Thomas studiert in Fulda. **Claudia** geht noch zur Schule.
		Thomas goes to college in Fulda. Claudia is still in school.
	P	Olga sagt immer, **Claudia** sei zu jung.
		Olga always says that Claudia is too young.
	L	Thomas, Peter, **Claudia** und Elke fahren in den Süden.
		Thomas, Peter, Claudia and Elke are travelling to the south.
	W	Gestern war **Claudia** noch gesund.
		Yesterday, Claudia was still well.
unstressed (u)	U	Die Arbeit war super. **Klausur** und mündliche Prüfung waren nicht so toll.
		The thesis was very good. Written and oral examinations were not so good.
	P	Tine sagt immer, **Klausur** schreiben macht Spaß.
		Tine always says that it is fun to write a test.
	L	Hausarbeit, Wetter, **Klausur** und Erkältung machen schlechte Laune.
		Chores, weather, exams and a cold cause sulkiness.
	W	Morgen muss sie wieder **Klausur** schreiben.
		Tomorrow, she has to write a test again.

Table 1. Material: Target words containing the cluster are highlighted. The utterances are sorted by stress of the syllable containig the cluster and by position of the target word within the utterance(U - utterance initial, P - phrase initial, L - list item, W - word.

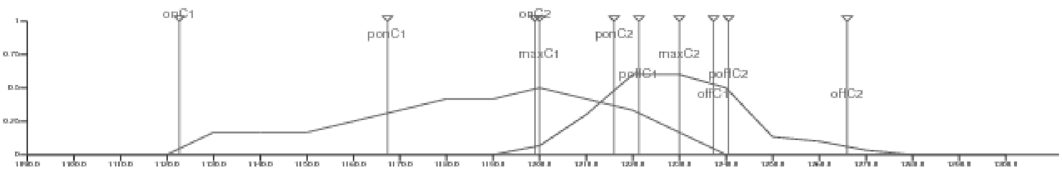

Figure 2. Dorsopalatal index (C1) and anteriority index (C2) over time [ms]. For event labels see text.

(*off*) at the 10 % threshold of the amplitude difference between maximal amplitude and lowest amplitude before/after *max*. Furthermore, a plateau of high contact was defined as the interval between the 70 % thresholds (*pon, poff*), as we consider this plateau to be less arbitrary than *max* itself. The duration of this plateau corresponds roughly to the constriction duration. Figure 2 illustrates this procedure. For statistics and analysis, the *R*-Software[2] in combination with the *EMU/R*-package was used. Four durational measures were extracted for individual speakers including all valid data. *Pause* represents the interval between the burst og the /kl/-cluster onset and the offset of the preceeding vowel. Depending on the prosodic boundary associated with the cluster onset, this interval might include a real pause. Two measures of strengthening are *plateau /k/* and *plateau /l/* defined as the interval between the events *pon* and *poff* for the respective gesture. *Sequence overlap* is the percentage of the cluster sequence in which contacts for both regions are registered (100 * (*offC1* - *onC2*) / (*offC2* - *onC1*)), see Byrd et al. (1995), higher values

[2]The R Project for Statistical Computing (http://www.r-project.org)

Sp.	Effect	Df	pause	plateau /k/	plateau /l/	seq. over.
	bd	3	55.13***	15.16***	9.1**	6.43***
SF	**stress**	1	0.18	0.3	40.6***	2.85
	bd:stress	3	2.24	0.59	3.47*	0.06
		64	U>PLW, P>W	U>LW, PL>W	u: UPL>W, s>u	UL<W
	bd	3	130.67***	9.03**	3.38*	12.77***
DP	**stress**	1	1.63	0.25	1.57	6.2*
	bd:stress	3	1.63	0.56	2.07	1.39
		54	U>PLW, PL>W	UPL>W	u: P>W	UPL<W
	bd	3	40.53***	2.05	1.15	6.27***
MV	**stress**	1	7.49**	0.03	0.06	0.69
	bd:stress	3	2.77*	0.1	0.28	0.6
		68	U>PLW, U: u>s			UL<W

Table 2. Boundary (bd) and stress effects on four durational paramters. *pause*, *plateau /k/*, *plateau /l/*, *sequence overlap*. Statistics are based on ANOVAs. Results of pairwise *t*-tests with *Bonferroni* adjustments are also given with > indicating greater duration / more overlap (*: $p<0.05$; **: $p<0.01$; ***: $p<0.001$).

indicating more overlap. For each individual speaker, ANOVAs were carried out using these measures as dependent variables. Two independent variables were defined: *Boundary* denotes the utterance types *U, P, L, W*. The two levels of *stress* correspond to clusters in stressed syllables *(s)* as in *Claudia* or in unstressed syllables *(u)* as in *Klausur*. Due to the small amount of *U* tokens for speaker DP, this condition was excluded from statistic calculations.

3. Results

3.1. Boundary effects

In order to assess whether the syntactic structure of our test corpus was realized by our speakers as predicted, we included the interval from the offset of the preceding vowel to the burst of /k/ in our statistical analysis (see Table 2, *pause*). This duration decreased significantly in accordance with our postulated boundary levels except for phrase boundaries and list, which were not distinguished by any of the speakers. In the utterance initial condition, speakers SF and DP often swallowed during pauses. Since this caused EPG patterns with almost total contacts prohibiting the detection of /k/ onset, these items were excluded from the analysis (SF: 6 items; DP: 16 items). For the plateau durations, we predicted that the /k/ should be lengthened more extensively at higher boundary levels than /l/. Two speakers realize /k/ with significantly longer dorsal plateaus in utterance initial position than in word initial position. The same speakers also shortened /l/ significantly in word initial position as compared to higher boundaries but only in unstressed syllables. All speakers showed significantly more overlap in word initial position than at higher boundary levels, indicated by higher values in *sequence overlap*, see Figure 3.

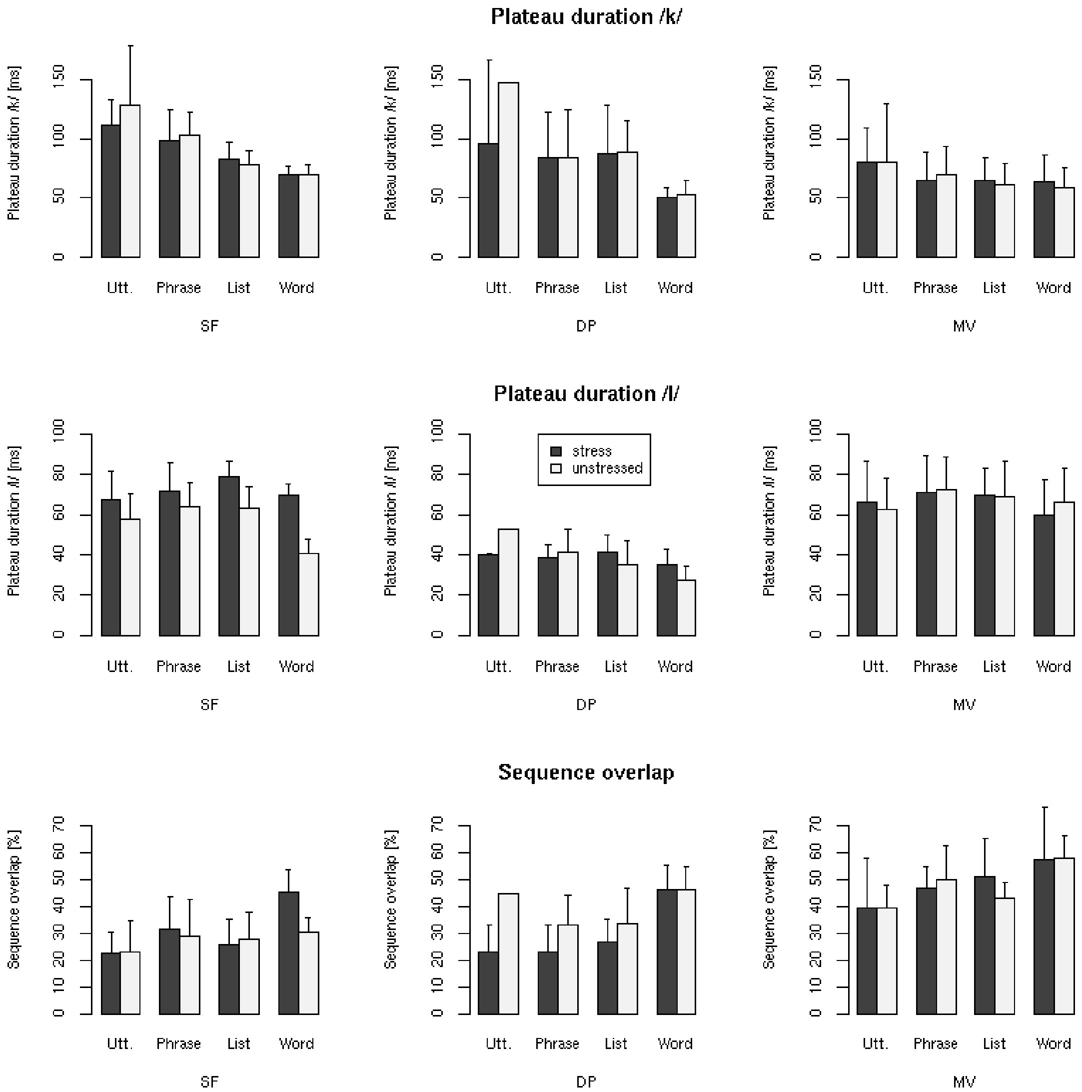

Figure 3. Barplots of the measures *plateau /k/, plateau /l/* and *sequence overlap* seperately for all speakers.

3.2. Stress effects

None of the speakers distinguished stressed from unstressed clusters by the duration of the plateau for /k/. The following lateral, however, was lenghtened significantly in stressed syllables word initially for the speakers DP and in all conditions for speaker SF. The internal coordination of the clusters was only very inconsistently affected by stress.

4. Conclusion

As was predicted, the segment closer to the boundary was stronger influenced by boundary strength than the more distal segment. This is in accordance with the pi-gestures approach proposed by Byrd and Saltzman (2003). However, this was significant for only two of three speakers (SF and DP). Speaker MV showed no lengthening effects at all.

More consistently, larger overlap between /k/ and /l/ was found word initially as compared to the higher boundary levels, confirming Byrd and Choi (2006) but contrary to Fougeron (1998) and Hardcastle (1985). Therefore, the internal organization of clusters is sensitive to prosodic variation and not entirely determined by language-specific grammar or perceptual demands. However, the overlap measure applied here is dependent on the duration of the entire /k/ gesture. The /k/ gestures, in turn, depend on the duration of pauses which were frequently realized at higher boundary levels, because full contact for /k/ can be present throughout the pauses. This poses a question which cannot be answered satisfactorily here: Are the effects on /k/ found in this study due to articulatory strengthening or due to the presence of pauses? As mentioned above, /k/ onset cannot be separated from the pause acoustically. We hope that analyzing of fricative-stop clusters, which are included in our corpus, will solve this problem.

Finally, our study is limited to one cluster type only. As was found by Chitoran et al. (2002) and Kühnert and Hoole (2006), the internal organization of clusters is also determined by simple constraints on the execution of the motor system (e.g. less overlap in /pl/ than in /kl/). Therefore, in the near future we will also analyze EPG patterns of words beginning with /sk/ and /kn/ clusters of more speakers.

References

Beckman, M. and Edwards, J. Articulatory evidence for differentiating stress categories. In Keating, P., editor, *Papers in Laboratory Phonology 3: Phonlogical Structure und Phonetic form*, pages 7–33. University Press, Cambridge, 1994.

Browman, C. and Goldstein, L. Competing constraints on intergestural coordination and selforganization of phonological structure. *Les Cahiers de l'ICP*, 5:25–34, 2000.

Byrd, D. and Choi, S. At the juncture of prosody, phonology, and phonetics?the interaction of phrasal and syllable structure in shaping the timing of consonant gestures. In *Proceedings of the 10th Conference on Laboratory Phonology*, Paris, June 2006.

Byrd, D., Flemming, E., Mueller, C. A., and Tan, C. C. Using regions and indices in EPG data reduction. *Journal of Speech and Hearing Research*, 38:821–827, August 1995.

Byrd, D. and Saltzman, E. The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics*, 31(2):149–180, 2003.

Cassidy, S. and Harrington, J. Emu: an enhanced hierarchical speech data management system. In McCormack, P. and Russels, A., editors, *Proceedings of the Sixth Australian International Conference on Speech Science and Technology*, pages 361–366, Adelaide, 1996. ASSTA.

Chitoran, I., Goldstein, L., and Byrd, D. Gestural overlap and recoverability: Articultory evidence from Georgian. In Gussenhoven, C. and Warner, N., editors, *Laboratory Phonology 7*. Mouton de Gruyter, Berlin/New York, 2002.

Cho, T. and McQueen, J. M. Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics*, 33(2):121–157, 2005.

Fougeron, C. *Variations articulatoires en début de constituants prododiques de différents niveaux en français*. Thèse de doctorat, Universit Paris III, 1998.

Gafos, A. A grammar of gestural coordination. *Natural Language and Linguistic Theory*, 20(2):169–337, 2002.

Hardcastle, W. J. Some phonetic and syntactic constraints on lingual coarticulation during /kl/ sequences. *Speech Communication*, 4:247–263, 1985.

Hardcastle, W. J., Jones, W., Knight, C., Trudgeon, A., and Calder, G. New developments in electropalatography: A state-of-the-art report. *Clinical Linguistics and Phonetics*, 3 (1):1–38, 1989.

Keating, P., Cho, T., Fougeron, C., and Hsu, C. Domain-initial articulatory strengthening in four languages. In Local, J., Ogden, R., and Temple, R., editors, *Papers in Laboratory Phonology 6: Phonetic interpretation*, pages 143–161. Cambridge University Press, Cambridge, 2003.

Keating, P. and Cho, T. Influence of prosodic factors on segment articulations and acoustics in English. Poster presented at October ASA meeting in Minneapolis, 2005.

Kühnert, B. and Hoole, P. Cohésion temporelle dans les groupes C1/l/ initiaux en français. In *Actes des XXVIes Journées d'Etude sur la Parole*, pages 545–548., 2006.

Lindblom, B. Explainig phonetic variation: A sketch of the H&H theory. In Hardcastle, W. J. and Marchal, A., editors, *Speech Production and Speech Modelling*, pages 403–439. Kluwer, Dordrecht, 1990.

Mooshammer, C. and Harrington, J. Linguistic prominence and loudness: a systematic comparison between lexical word stress, sentence accent and vocal effort, 2005.

Mooshammer, C. and Fuchs, S. Stress distinction in German: Simulating kinematic parameters of tongue tip gestures. *Journal of Phonetics*, 30:337–355, 2002.

Schiel, F. Automatic phonetic transcription of non-prompted speech. In *Proceedings of the 14. International Congress of Phonetic Sciences*, pages 607–610, San Francisco, August 1999.

Turk, A. and White, L. Structural influences on accentual lenghtening in English. *Journal of Phonetics*, 27:171–206, 1999.

Wright, R. A review of perceptual cues and cue robustness. In Hayes, B., Kirchner, R., and Steriade, D., editors, *Phonetically Based Phonology*, pages 34–57. Cambridge University Press, 2004.