CLINICAL LINGUISTICS & PHONETICS, 2004, VOL. 18, NO. 6–8, 559–575

Speaker-specific kinematic properties of alveolar reductions in English and German

BARBARA KÜHNERT†‡ and PHIL HOOLE§

† Institut du Monde Anglophone and ‡ Laboratoire de Phonétique et Phonologie (CNRS – UMR 7018), Université Paris III – Sorbonne Nouvelle, 5, rue de l'Ecole de Médécine, F-75006 Paris, France
§ Institut für Phonetik und Sprachliche Kommunikation, Ludwig-Maximilians-Universität München, Schellingstr. 3, D-80799 München, Germany

(*Received 11 February 2004; accepted 6 April 2004*)

Abstract

A simultaneous EPG/EMA study of tongue gestures of five speakers was conducted to investigate the kinematic events accompanying alveolar stop reductions in the context of a velar plosive /k/ and in the context of a laryngeal fricative /h/ in two languages, English and German. No systematic language differences could be detected. Alveolar productions before a following /h/ showed only a marginal weakening of the formation of complete occlusion, while alveolar productions before a following /k/ showed a wide range of reductions, including instances of a complete deletion of the alveolar gesture. The extension of movement reduction varied between and within subjects. Importantly, while speakers were consistent with themselves, they employed different articulatory patterns with respect to the timing relationship between movement initiation, overall movement duration, peak velocity as well as closure duration. An attempt is made to relate the observed movement patterns to the dynamic factors of the speech mechanism.

Keywords: Electropalatography, electromagnetic articulography, experimental phonetics, weakening, assimilation.

Introduction

The analysis of the (co-)articulation of consonant sequences has recently gained new importance, primarily due to several publications by Byrd and colleagues (Byrd, 1996a,b; Byrd and Tan, 1996). Adopting the framework of Articulatory Phonology, in which linguistically relevant vocal tract configurations, or *gestures*, are the basic units of the phonetic and phonological specification (Browman and Goldstein, 1986, 1992), Byrd and colleagues found that inter-gestural coordination in consonant sequences is subject to a variety of influences, such as place and manner of articulation, as well as syllable structure. Specifically, electropalatographic data showed that in utterances such as '*type bad gab again*' the consonants

Clinical Linguistics & Phonetics
ISSN 0269-9206 print/ISSN 1464-5076 online
http://www.tandf.co.uk/journals
DOI: 10.1080/02699200420002268853

were nearly completely overlapped, with contact of [g] often starting synchronously with that for [d]. Increase in speaking rate had minimal influence on the temporal extent of consonant overlap, however contact duration shortened and maximum contact decreased.

Word-final alveolars are also the consonants most likely to be assimilated to a following consonant (Paradis and Prunet, 1991). Assimilations in connected speech, in turn, have been considered as a test case of speech production patterns that are either best explained at the level of central control, or, in the spirit of Articulatory Phonology, as the result of temporal co-occurrence in the articulation of subsequent sounds. Thus, much of the discussion in the literature concerns the dichotomy between a discrete change of 'segments' implied by a phonological account of assimilation and the phonetic continuum of articulatory reductions that has been observed in experimental studies of assimilated utterances (Barry, 1992; Ellis and Hardcastle, 2002).

Without repeating the arguments here, it should be noted that within the apparatus afforded by any phonological approach the process of assimilation is essentially formulated as a replacement of one phonological unit by another. As a corollary, in an assimilated production such as '*the kid comes*', in which the specification of the alveolar is reassigned to the following velar, the resulting consonant sequence should be comparable in duration (and in other characteristics) to the consonant sequence in an utterance such as '*the wig comes*'. In fact, some durational lengthening of assimilated consonants has been reported, but the evidence is not unequivocal (Barry, 1991; Nolan, Holst and Kühnert, 1996). Furthermore, the finding of what has been called *residual gestures* (Nolan, 1992), i.e. the execution of weakened coronal movements without complete closure, poses a problem (although not insolvable, Hayes, 1992; Scobbie, 1995) for any essentially binary oriented approach.

Such residual gestures can be accounted for more elegantly within the framework of Articulatory Phonology. As the abstract gestural units are coordinated with each other in a highly complex fashion, concurrently activated gestures can overlap as well as compete for articulators simultaneously, resulting in spatial and temporal reductions of executed speech movements (Browman and Goldstein, 1990, 1992). However, the model does not foresee the deletion of gestural units and is therefore unable to account for cases of what appears to be a complete loss of an alveolar articulation (Holst and Nolan, 1995; Ellis and Hardcastle, 2002).

There is yet another proposal which explains the observed gradual reductions of alveolar plosives from a more phonological perspective. Hayes (1992) suggests that two distinct processes have to be differentiated, an assimilation process, which is phonological and discrete, and a syllable-final weakening process, which is phonetic and gradual, and is able to generate residual articulations. Since syllable-final weakening of coronals should occur regardless of the following context, this would imply that in an utterance such as '*the kid hunts*', in which the coronal consonant is followed by a neutral laryngeal fricative, the alveolar should show a similar amount of reduction as in the utterance '*the kid comes*'.

The kinematic properties of alveolar-velar sequences are also relevant from a general point of view of speech motor control. The control of oral closure is a recurrent topic in recent studies (Mooshammer, Hoole and Kühnert, 1995; Löfqvist and Gracco, 2002), which seek to understand the commonly observed forward sliding movement of the tongue during occlusion, in particular during the occlusion of velar plosives. Although it has been shown that the forward movement is only slightly affected by variations in speech rate during VCV productions (Munhall, Ostry and Flanagan, 1991), it is not clear how robust these patterns are during the production of connected speech and how they are influenced by a temporally close formation of a second oral closure in the vocal tract.

The purpose of the present study is to evaluate the kinematic properties of alveolar stop productions in the context of a following velar plosive /k/ and a following laryngeal fricative /h/ in two languages: English and German. Recordings of tongue movements were made to explore four questions: (a) how do the characteristics of alveolar reductions differ in the two consonantal contexts; (b) which movement patterns during closure phase can be observed for overlapping alveolar and velar plosives; (c) how do residual alveolar gestures differ from canonical closure productions; and (d) how do completely assimilated alveolars differ from corresponding velar gestures.[1]

Method

Subjects and material

Three German (GER1, GER2, GER3) and two British male speakers (ENG1, ENG2) participated in the experiment. The material consisted of utterances containing strings of /-V1C1#C2V2-/, with # indicating a word boundary. The consonant clusters were either /d#k, g#k, d#h/ for English, and /t#k, k#k, t#h/ for German, i.e. the consonant sequences include potential alveolar-velar assimilation sites, velar control contexts, and contexts in which the alveolar is followed by a gesture commonly not known for motivating alveolar reductions. Voiced alveolar stops were chosen for English, since here voiceless alveolar stops are often reinforced or replaced by glottal stops. In German, however, consonants exhibit word-final voicing neutralization and only voiceless plosives will occur.

For both languages lax vowels were used, with one exception (English: V1 = /ʌ, e, ɪ, ɒ, ʊ/ and V2 = /ʌ, ɑː, ʊ/; German: V1 = /a, e, ɪ, ɔ, ʊ/ and V2 = /a, ɑː, ʊ/). All vowels were combined with all consonant combinations. Examples of test utterances are:

Consonants	English	German
/d#k/, /t#k/	The rod can't crack.	Dein Spott kam überraschend.
/g#k/, /k#k/	The dog can't bite.	Der Schock kam nach dem Unfall.
/d#h/, /t#h/	It is an odd harem.	Sie will Kompott haben.

The subjects produced the speech material once at their normal self-determined rate, once according to the instruction 'as fast as possible'. Each subject repeated

[1]In the following 'movement' and 'gesture' are used interchangeably to refer to the actual observable movements in the data. The term 'underlying gesture' is used to refer to an underlying abstract gestural representation in the sense of Articulatory Phonology. 'Alveolar reduction' is used to refer to any alveolar stop production without complete closure. More specifically, 'assimilated alveolar' refers to an alveolar reduction before a following velar plosive, and 'instable alveolar' refers to an alveolar reduction before a laryngeal fricative /h/.

the list of sentences twice, resulting in a total of about 180 utterances, or 27–30 repetitions for each consonant sequence and speaking rate.

Recording procedure

Articulatory movements were monitored by the simultaneous use of electropalatography (Reading EPG2-system) and electromagnetic articulography (Articulograph AG100; Carstens Medizinelektronik). Movement signals were sampled at 200 Hz; the audio signal was recorded on DAT tape with synchronization pulses generated by the EMA computer (for further details, see Hoole, 1996).

Receiver coils were placed on three positions of the tongue, referred to as tongue tip, mid and back. The rearmost coil was placed about 0.5 cm behind the place where the tongue touched the rear end of the artificial palate during the production of a velar occlusion. The front coil was fixed approximately 1 cm posterior to the tongue tip, and the third (middle) coil was placed in-between. A further transducer on the lower incisors registered jaw movement.

EMA data were checked for reliability (e.g. possible rotational misalignment of sensors), which led to the last 12 utterances of GER2 being discarded. In a further step, movement data were corrected for head movement, low-pass filtered, and transformed to a coordinate system in which the origin corresponds to the position of a reference coil attached to the upper incisors. The vertical axis of the coordinate system was defined by the principal component of jaw movement in speakers' /aC1#C2ɑː/ productions. Tangential velocity signals were calculated for tongue-tip and tongue-back.

Measurements

Several kinematic parameters were identified for each consonant sequence using measurements made in the EPG displays and the position and velocity signals of the EMA representations. An illustration of parameter extraction is given in figure 1.

- *T_Closure* = duration of alveolar closure; interval between the first and last EPG frame indicating a complete closure in the anterior region of the palate.
- *K_Closure* = duration of velar closure: (i) in the context of V1 = /ɪ, e/, interval between the first and last EPG frame with complete closure in the posterior region of the palate; (ii) in the context of V1 = /ʌ/a, ɒ/ɔ, ʊ/, interval between the points in time at which vertical displacement of the back coil exceeded 80%, and at which it fell short of 90% of its maximal amplitude.
- *T_Duration* / *K_Duration* = duration of alveolar and velar raising movement; interval between the moment at which the tongue tip or tongue back speed signal exceeded 5 % of its maximal value after the preceding vowel, and the moment at which it fell short of 5 % of its maximum after velocity peak.
- *T_Velocity* / *K_Velocity* = peak velocity of tongue tip and tongue back receiver during alveolar or velar raising movement.
- *T_Displacement* / *K_Displacement* = maximal alveolar and velar displacement; difference of tongue height between minimal and maximal vertical displacement of tongue tip or tongue back transducer during alveolar or velar raising movement.

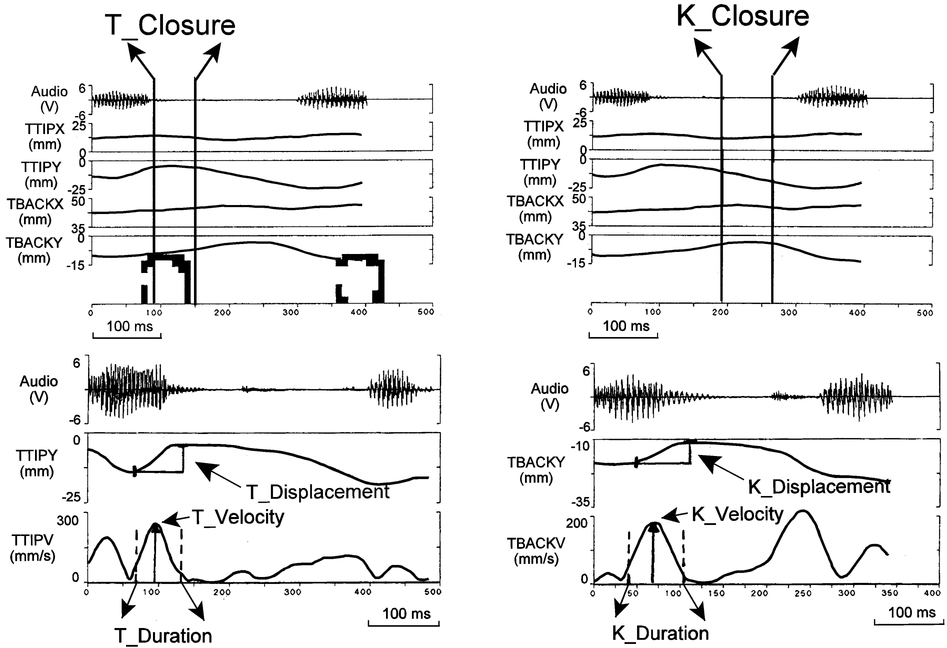

Figure 1. *Illustration of the determination of alveolar and velar closure duration (top panels) and kinematic parameters of tongue-blade and tongue-dorsum gestures (bottom panels). The signals labelled TTIPV and TBACKV are the tangential velocity of the tongue-tip and tongue-back sensors, respectively. The top left panel shows schematic EPG patterns (with full closure at the front of the palate) at the time locations indicated by the arrowed cursors.*

Results

Qualitative data description

Initially, every token of an alveolar stop has been identified as a reduction in which there was a gap of one or more EPG electrodes in horizontal tongue-palate contact in the alveolar region. The frequency of these reductions is listed in table 1 (A). In the /k/ as well as the /h/ context incomplete alveolar closure in the normal speaking condition only occurs sporadically. In fast speech, alveolar reductions are more frequent when preceding the velar plosive.

Figure 2 shows a summary of the tongue coil positions for all speakers at the beginning of the acoustically defined closure for canonically produced alveolars (T), assimilated alveolars in the context of /k/ (A), instable alveolars in the context of /h/ (I), and the velar control sequences (K).

Looking at alveolar reductions in the /k/ context (A), it can be seen that subjects differ in the extent of *spatial* movement reduction. Speaker GER2 produces assimilated utterances with a still prominent raising of the tongue tip, with clearly more displacement than in the velar context. For GER1 and ENG2, the tongue tip receiver indicates that assimilations are most commonly produced with a residual gesture, which sometimes might no longer be evident. The strongest alveolar movement reduction can be observed for ENG1 and GER3, whose positional values of assimilated and velar control sequences closely coincide.

Table 1. *Frequencies of alveolar reductions*

(A)
Frequency of alveolar instability in /h/-context and alveolar assimilation in /k/-context for the speech-rate conditions normal and fast. Number of tokens (max = 30).

Context	GER1	GER2	GER3	ENG1	ENG2
/h/-Normal	1	1	3	3	1
/h/-Fast	14	4	7	16	6
/k/-Normal	1	3	4	0	0
/k/-Fast	21	9	28	23	10

(B)
Breakdown of alveolar assimilation in /k/-context at the fast speech rate.
/t/_ra = residual alveolar; /t/_na = zero alveolar.

Reduction category	GER1	GER2	GER3	ENG1	ENG2
/t/_ra	16	9	4	7	6
/t/_na	5	0	24	16	4

A comparison of tongue back positions between the categories T, A and K reveals that *temporal* overlap between the alveolar and velar gesture increases during the production of assimilated alveolar-velar sequences. Although the subjects differ in the extent, all show a distinctively higher position of the tongue dorsum receiver at the beginning of acoustic closure for category A than for category T.

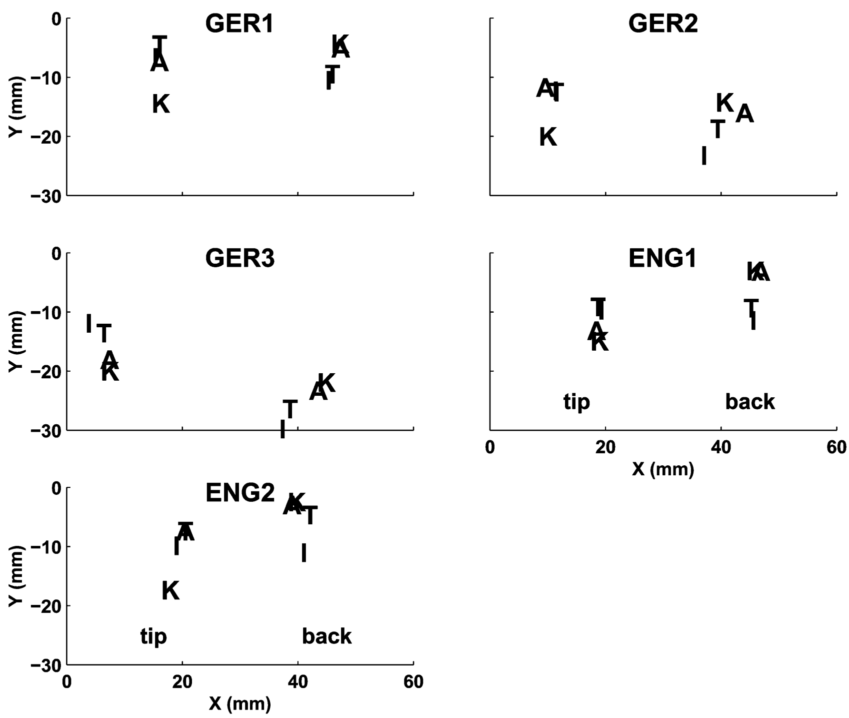

Figure 2. *Average position of tongue-tip and tongue-back sensors at the start of the closure phase for full alveolars (T), reduced alveolars (A), instable alveolars (I) and velars (K).*

Finally, not only are alveolar reductions before /h/ less frequent than before /k/, they are also less prominent. For all subjects, the positions of tongue tip coil show only marginal differences between the categories I and T, while the positions of the tongue dorsum coil are de facto identical. Alternatively, the positions of the tongue tip coil in assimilated tokens (A) are further reduced as compared to instable articulations (I), whereas the positions of the tongue dorsum receiver are situated much higher in anticipation of the following velar.

For the following analysis, alveolar reductions in the /k/ context were divided into two further categories on the basis of an inspection of the EMA representations: residual alveolars, defined as sequences in which a coronal movement was still present, and zero alveolars, defined as sequences in which the raising movement of the tongue tip did not exceed the one of the corresponding velar control token. The distribution of residual and zero alveolars is listed in table 1, part B.

Quantitative data analysis

Analyses of variance for unequal measures were conducted (SAS Institute Inc., 1988) for each speaker separately to assess the influence of consonant sequence on several kinematic events. Unequal measures were given by the fact that the subjects differed in the number of instable, assimilated and zero alveolars they produced. The factor *consonant sequence* is coded as follows:

$TK_{\_normal}$ / $TK_{\_fast}$	alveolar-velar sequence in normal or fast speech rate with full alveolar closure.
$TK_{\_residual}$	alveolar-velar sequence in fast speech rate with residual tongue tip raising movement.
$TK_{\_zero}$	alveolar-velar sequence in fast speech rate with *no* visually detectable tongue tip raising.
$TH_{\_normal}$ / $TH_{\_fast}$	alveolar-laryngeal fricative sequence in normal or fast speech rate with full alveolar closure.
$TH_{\_instable}$	alveolar-laryngeal fricative sequence in fast speech rate without full closure.
$KK_{\_normal}$ / $KK_{\_fast}$	velar-velar sequence in normal or fast speech rate.

Since *consonant_sequence* covers a complex group of categories, references to the statistical results below will be to selected post-hoc comparisons. Note that not every consonant category is available for every speaker: the category $TK_{\_fast}$ has been omitted from the analysis of GER3, since it only contained two tokens. $TK_{\_zero}$ is absent for GER2 since no tokens occurred, and has also been omitted for GER1 and ENG2 since the few tokens that occurred were confined to environments where the preceding vowel was /ɪ/ and would thus have given a biased view of the results.

Kinematic properties of alveolar and velar closure

Figure 3 shows the mean durations of alveolar and velar closure phase (T_Closure, K_Closure). The results of T_Closure reveal a prominent difference between German and English speakers. For German speakers, T_Closure only seems to be influenced by speech rate. For English speakers, duration of T_Closure is not only

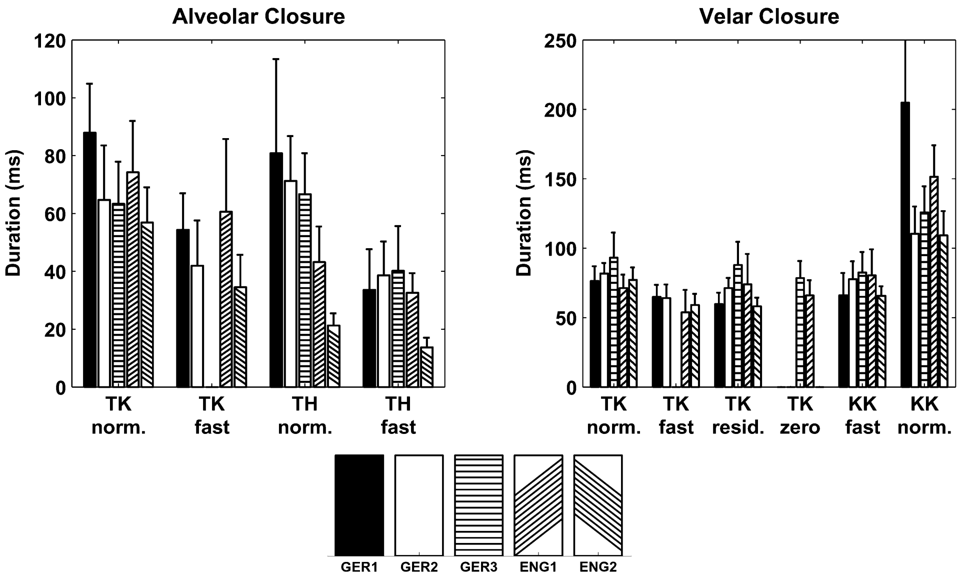

Figure 3. *Overview of alveolar and velar closure duration for all subjects and sound categories (error bars indicate one standard deviation).*

influenced by rate, but also by the following consonant, being far longer before /k/ than before /h/.

The language-specific difference of alveolar closure duration might be motivated by two reasons. Firstly, in theory the study investigates voiced alveolar plosives in English. In practice, many of the alveolar stops were partially or completely devoiced due to the following voiceless [k]. They were however never devoiced in the context of /h/. Thus, the durational difference is the result of a difference in voicing since voiced stops are known to have a shorter closure phase than voiceless stops (Crystal and House, 1988). Secondly, the alveolar closure duration could partially reflect a different articulatory strategy of alveolar productions triggered by the velar context.

Considering the results of K_Closure, $KK_{\_normal}$ shows, unsurprisingly, the longest closure period. Furthermore, the reduction of K_Closure of single velar stops in $TK_{\_normal/fast}$ is considerably smaller than the reduction of T_Closure presented above. Recall that the most important parameter in the discussion of (phonological) re-implementation vs. (phonetic) gestural reduction is the duration of velar closure phase in fast speech. There is some evidence that the duration of K_Closure increases in assimilated alveolar-velar sequences. For three of the subjects (GER2, ENG1, ENG2), K_Closure is significantly longer for $KK_{\_fast}$ than for the single velars in $TK_{\_fast}$, i.e. the velar clusters show some kind of durational blending of two gestures. Importantly, K_Closure of $TK_{\_residual}$ does not differ significantly from either single velars or velar clusters, lying in-between and thus reflecting a partial increase in closure period. However, some inconsistency can be found in the data of GER1, for whom all three categories $TK_{\_fast,}TK_{\_residual}$ and $KK_{\_fast}$ do *not* show any significant durational differences.

In order to test the timing relation between T_Closure and K_Closure, the percentage of closure overlap has been examined for $TK_{\_normal}$ and $TK_{\_fast}$. For every token the interval between the offset of alveolar and the onset of velar closure was calculated and considered in relation to overall closure duration of the token.

Table 2 shows the mean values of closure overlap. A negative value indicates that velar closure starts after alveolar closure offset. The range of possible overlap has GER1 and ENG2 at opposite extremes. In $TK_{\_normal}$ speaker GER1 produces the two closure periods clearly sequentially, in $TK_{\_fast}$ with minimal overlap. Likewise, GER3 shows no overlap in $TK_{\_normal}$. GER2 and ENG1 exhibit minimal overlap between the two closure periods in the normal condition, which slightly increases with an increase in rate. Finally, ENG2 shows at both speech rates a high degree of simultaneous closure production. Interestingly, ENG2 shows at the same time a strong correlation between the amount of closure overlap and the duration of alveolar closure: the longer T_Closure, the earlier the onset of K_Closure. It would thus seem that the simultaneous production of the second stop lengthens the production of closure period of the first.

In order to evaluate lingual movement during alveolar and velar stop closure, the horizontal distance covered by the tongue tip and tongue back sensor between the on- and offset of alveolar and velar closure, respectively, has been measured and is shown in figure 4. Since the magnitude of tongue movement during stop closure periods is known to be context dependent (Mooshammer *et al*., 1995), statistical analyses were conducted separately in two sets, once with V1 = /ɪ, e/, once with V1 = /ʌ/a, ɔ/ɒ/ and /ʊ/. Negative values indicate that the tongue moved backwards, positive values that it moved forwards.

For all speakers and in all environments the tongue tip *always* moved forward during T_Closure, and stop production was *always* released with a movement to the front. In the neutral /h/ context, all speakers show more tongue tip movement during closure for central/back vowels than for front vowels. Variation of speech rate seemed to have little impact. The amount of lingual tongue tip movement in the /k/ context shows inter-individual differences. Speakers GER1, GER3 and ENG1 display essentially the same movement pattern as in the case of following /h/. For GER2 and ENG2, however, horizontal tongue tip movement increases significantly in the velar context.

Considering horizontal movements of the tongue back sensor during velar closure, the main observation is that they are quite small. In the category $KK_{\_normal}$ in the context of central/back vowels, in which all but one subject (ENG1) show the most extensive movement, the average length of the path is just above 3mm, decreasing only weakly in the fast speech condition. In the context of a preceding

Table 2. *Mean overlap (SD in brackets) of alveolar and velar closure as percentage of total consonantal duration. Negative values indicate velar closure starts after alveolar closure offset. No value determined for GER3/fast rate as only two tokens had alveolar closure*

Subject	Normal rate	Fast rate
GER1	−22.16 (7.69)	1.76 (12.19)
GER2	4.28 (11.18)	6.47 (13.89)
GER3	−13.61 (9.11)	=
ENG1	6.66 (8.44)	16.21 (20.06)
ENG2	36.67 (14.92)	33.30 (17.66)

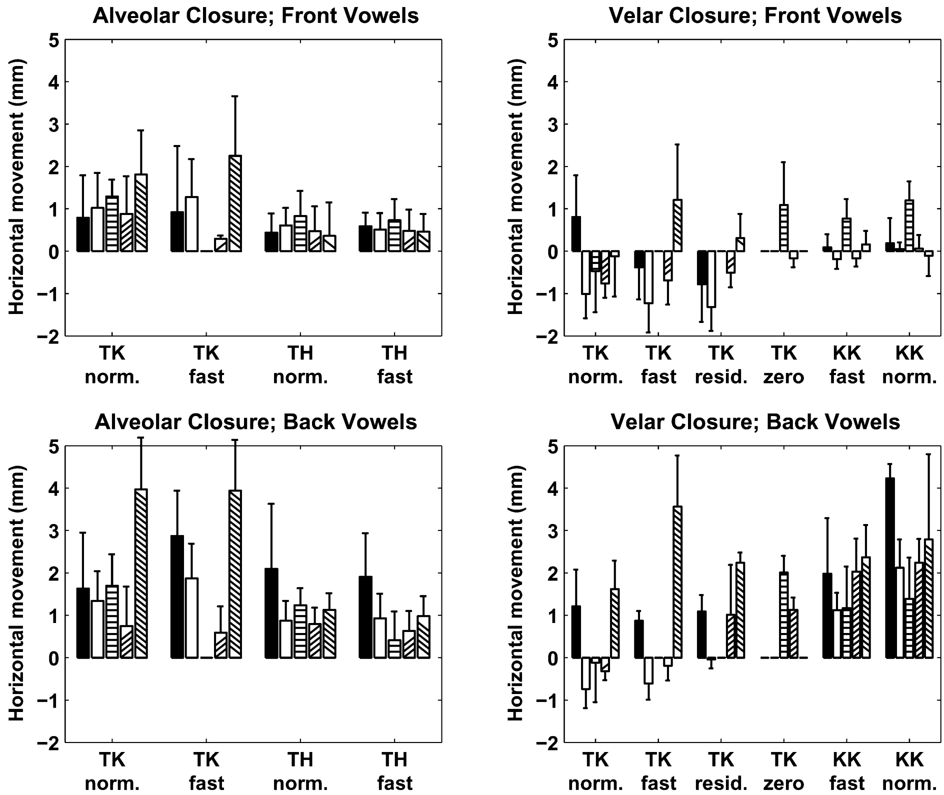

Figure 4. *Overview of amount of horizontal movement during alveolar and velar closure duration for all subjects and sound categories. Alveolar results based on tongue-tip sensor, velar results based on tongue-back sensor. Positive values indicate forward movement of the tongue. Same coding of subjects by hatch pattern as figure 4.*

front vowel, horizontal movement of the tongue dorsum receiver is essentially cancelled out for all speakers in both speaking rates.

Leaving the patterns of speaker ENG2 aside at first, horizontal tongue movement during velar closure is suppressed by the articulation of a preceding alveolar stop. Regardless of V1, the tongue moves either not at all or slightly backward in the categories $TK_{\_normal/fast}$. The results of category $TK_{\_residual}$ suggest that the spatial extent of alveolar movement reduction influences the movement path during velar closure. For speaker GER2, who produced residual alveolars with a still prominent raising gesture, tongue back movement during velar closure is the same as in $TK_{\_fast}$. For the remaining speakers, for whom spatial reduction in $TK_{\_residual}$ was more pronounced, the pattern reverses and the direction and extent of tongue dorsum movement resembles the one found for velar control sequences. $TK_{\_zero}$ and $KK_{\_fast}$ show no significant differences.

Subject ENG2, finally, is the only speaker who displays a *clear* forward movement of the tongue back sensor for $TK_{\_normal/fast}$ in spite of his large overlap of T_Closure and K_Closure. However, the forward movement associated with the velar closure is still less than the very prominent forward movement associated with the alveolar closure production.

Kinematic properties of alveolar and velar raising movement

This section examines the parameters of duration (T_Duration, K_Duration), peak velocity (T_Velocity, K_Velocity), maximal vertical displacement (T_Displacement, K_Displacement) and stiffness (T_Stiffness, K_Stiffness) of the alveolar and velar raising movements in all /C1#C2/ sequences. The results of the kinematic properties are assembled in figure 5.

Apart from subject ENG2, the alveolar raising movement is relatively constant across subjects and consonant categories and amounts on average to 60–80ms. The trend to produce alveolar stops with shorter movement duration in fast speech can be observed for three speakers (GER1, GER3, ENG2). None of the subjects shows a significant variation in T_Duration with regard to $TK_{\_residual}$, $TK_{\_instable}$ and $TK_{\_fast}$.

Looking at the parameter K_Duration, a first thing to note is that movement duration within the category $KK_{\_normal}$ is surprisingly homogeneous across subjects and lies around 90–100ms, and, as such, is consistently longer than T_Duration. Furthermore, there is a tendency for all speakers to reduce K_Duration for $KK_{\_fast}$. The categories $TK_{\_residual}$, $TK_{\_zero}$ and $KK_{\_fast}$ show *no* difference in K_Duration.

A second thing to note is the fundamental difference between speakers with respect to the initiation of the velar gesture following an alveolar plosive. The three German speakers produce velars in the category $TK_{\_normal}$ with a significantly longer K_Duration than velars in all other consonant categories, i.e. the velar raising movement is initiated significantly earlier. Since tongue tip and tongue dorsum are assumed to be relatively independent during alveolar productions (Perkell, 1969), the German speakers seem to apply in this context what is generally referred to as a *look-ahead* strategy (Henke, 1966). The raising of the tongue dorsum for the velar production starts as soon as the articulator is allowed to do so, and only moves a little further upwards shortly before velar closure. Speaker ENG1, by contrast, seems to prefer fixed movement durations before velar occlusion across all contexts. The two different strategies are illustrated in figure 6, which contrasts the sequence /et#kɑː/ produced in normal speech rate by ENG1 and GER3.[2]

Overall, T_Velocity is not influenced by the nature of C2 ($TK_{\_normal}$ vs. $TH_{\_normal}$). An increase in speech rate either results in a similar or slightly decreased velocity. For all speakers, T_Velocity is further reduced during the production of $TK_{\_residual}$ and $TH_{\_instable.}$

K_Velocity in velar control sequences increases with an increase in tempo, indicating that variations in speech rate can have opposite effects on different articulators: while tongue tip raising movements are produced slower, tongue dorsum movements are produced faster in fast speech. K_Velocity of $TK_{\_zero}$ and

[2]Another explanation would be that part of the initial raising of the tongue dorsum receiver is not to be attributed to the velar production *per se*, but is merely a corollary of overall tongue raising for the articulation of the alveolar. However, a reanalysis of the duration of the velar raising movement as the interval between the moment of maximal acceleration and the end of K_Duration (cf. Perkell and Matthies, 1992, who propose that the moment of maximal acceleration more reliably reflects the tongue movement associated with the actual target gesture), did not show a fundamental change in the duration of velar raising movement for any of the speakers.

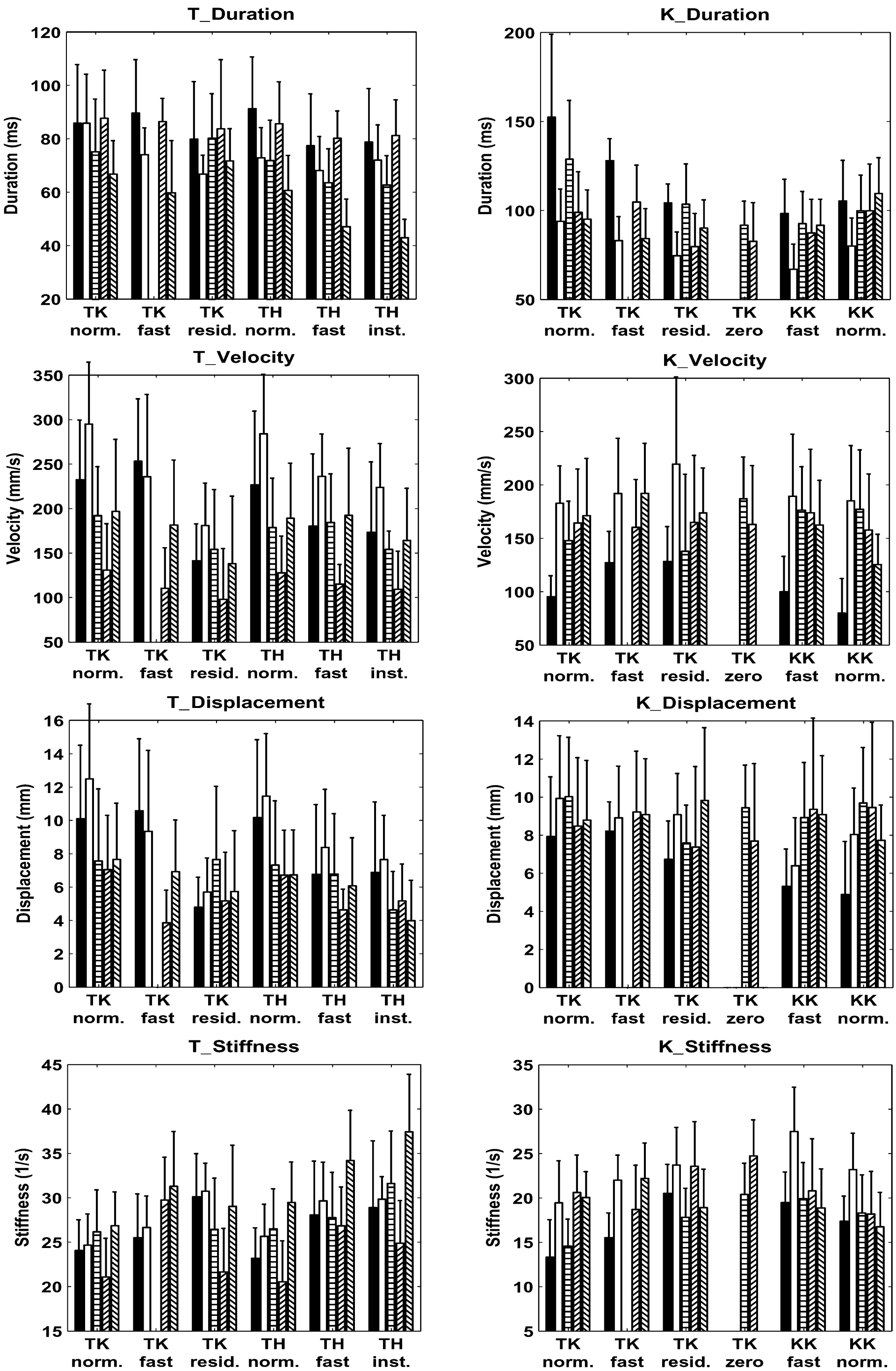

Figure 5. *Kinematic parameters of alveolar and velar tongue-raising movement for all sound categories and subjects. Error-bars indicate 1 standard deviation. Same coding of subjects by hatch pattern as figure 4.*

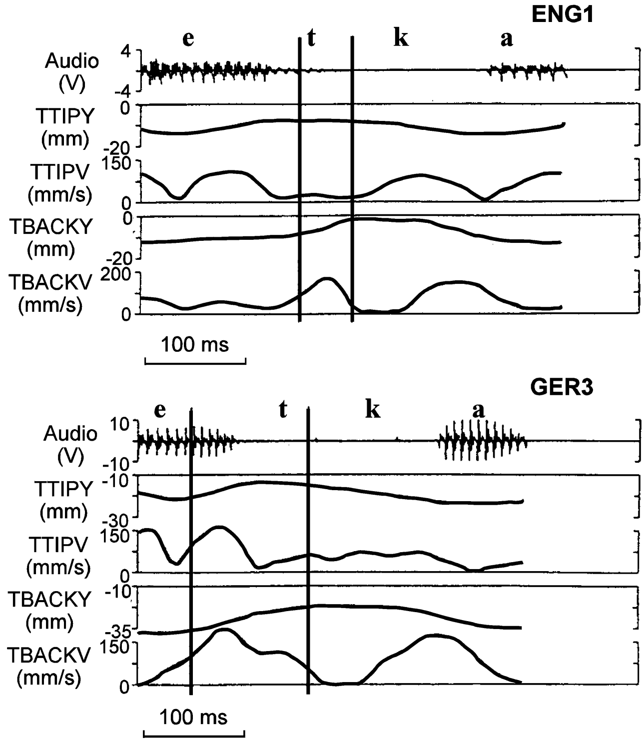

Figure 6. *Example of letkal for subjects ENG1 and GER3. Vertical position and tangential velocity for tongue-tip and tongue-back sensors. The left cursor is located at the point of maximum acceleration of the tongue-dorsum towards velar closure. The right cursor is positioned at the kinematically-defined end of the movement towards velar closure.*

$KK_{\_fast}$ are not different from one another. Considering K_Velocity for the different tokens in the alveolar-velar categories, no consistent pattern across subjects emerges.

T_Displacement slightly decreases for sequences with fully produced alveolars when speech rate increases. No difference in T_Displacement can be found between the categories $TH_{\_instable}$ and $TH_{\_fast,}$ reflecting the observation made earlier that instable alveolar stop productions were often missing one or two contacts only in the front most EPG row. T_Displacement in the production of $TK_{\_residual}$ generally decreases.

The mean value of K_Displacement does not reveal any consistent trend either across different sequences containing a velar plosive or across speakers. Differences between $TK_{\_zero}$ and velar control sequences cannot be found.

The final parameter to be examined is the ratio of peak velocity to displacement (parameter *k*) since previous work (e.g. Kuehn and Moll, 1976; Ostry and Munhall, 1985) has found positive correlations across speakers between articulatory velocity and movement displacement, indicating that it is the ratio of the two parameters that appears to be relevant in implementing changes in rate. The ratio has been considered to be a relative indicator of articulator 'stiffness' (Ostry and Munhall, 1985; Perkell, Zandipour, Matthies and Lane, 2002).

At normal speaking rate, T_Stiffness is not influenced by the following consonant ($TK_{\_normal}$ vs. $TH_{\_normal}$). However, there is a general trend across speakers for the parameter to increase with an increase in rate, and for three out of five speakers to increase further during the production of $TK_{\_residual}$ and $TH_{\_instable}$. In other words, estimated stiffness is higher, although, as seen previously, alveolar peak speed and maximal displacement were reduced in faster speech. The same amount of distance is thus produced with faster velocity.

The same influence of speech rate can be observed for K_Stiffness in a comparison of $KK_{\_normal}$ and $KK_{\_fast}$. Thus, here estimated stiffness increases, although in absolute terms peak velocity increased in the fast speech rate. K_Stiffness during the raising movement in the categories $TK_{\_normal}$, $TK_{\_fast}$ and $TK_{\_residual}$ does not show any regularities.

Discussion

The most consistent outcome across speakers in the present study is that alveolar reductions before a fricative /h/ are qualitatively different from alveolar reductions before a velar stop. They occur less frequently, and, if present, show only a marginal weakening of the occlusion. Essentially, an increase in speech rate results in shorter closure and movement duration, produced with a higher ratio of peak velocity to displacement. Arguably, this data does not rule out Hayes' proposal (1992) that gradual reductions of alveolar stops are the outcome of a syllable final-weakening process, according to the rule '(…) depending on rate and casualness of speech, lessen the degree of a COR segment (…)' (p. 284), but it seems that the rule would need an additional specification, such as 'depending on rate and casualness of speech *and the nature of the following segment*'. This, however, would strike us as an overly concrete phonological specification (cf. Scobbie, 1995) and it seems plausible to suggest that some other factors play a role in alveolar-velar place assimilation.

All speakers produced alveolar-velar assimilations in fast speech and the EMA recordings *showed evidence of a residual alveolar gesture at least some of the time for all speakers*, *and instances of no alveolar gesture for four out of five speakers*. Differences that hinted at a genuine language-specific difference between English and German have not been found. As compared to alveolar stops produced with complete closure, residual alveolars show no consistent difference with respect to the duration of the alveolar raising movement, but they generally show a reduction in displacement and overall velocity, with an equally high or higher value of the parameter 'stiffness'. Tokens in which no alveolar gesture was present were identical to the production of velar control tokens with respect to all kinematic parameters. As to the crucial parameter of velar closure duration in assimilated alveolar-velar utterances, our results provide some evidence of a partial increase in tokens containing a residual gesture. However, the data also suggest that the complete fusion of two velar stops in fast speech could as well result in closure durations identical to an individual stop. This data is therefore a healthy reminder that the interpretation of closure duration in fluent speech still has to proceed cautiously, since there is no indication in the movement trajectories that allows us to differentiate unambiguously between single and double articulations. Consequently, the interpretation of closure duration is open in all directions: either the productions reflect the loss of the alveolar gesture without velar involvement, or they reflect a reassignment to the velar articulation, which, in turn, merges with the following velar.

One factor in our data that seems to influence the occurrence and the degree of assimilations relates to the specific properties of articulator movement during velar occlusion, and the observation that '…the sagittal outline of the tongue for /k/ seems to increase in length compared to its length for other consonants' (Perkell, 1969, p. 58; for further evidence, see Löfqvist and Gracco, 2002). All speakers showed a horizontal forward movement of the tongue body during closure period of the velar-velar sequences following a central/back vowel, hardly any movement following a front vowel. The introduction of a second oral closure in alveolar-velar sequences evoked a variety of movement patterns, depending on the degree of overlap of the two closure phases. In cases where the two closure periods were produced with no or little overlap, i.e. when the tongue front was free to move during velar production, the lingual movement during velar closure period remained essentially the same. However, in cases where the two closure periods were produced with more overlap, the pattern changed: either the tongue tip moved forwards, while tongue back simultaneously moved backwards, or both tongue tip and back moved forwards, but the forward movement of the tongue tip gesture was greater than the simultaneously executed forward movement of the tongue back gesture. The effect of a residual coronal gesture, finally, seemed to depend partially on its degree of reduction, with prominent raising movements still causing the tongue body to move backwards during velar occlusion, and less prominent raising gestures resulting in the typical velar forward movement. In other words, the tongue outline always increases in its overall length. Interestingly, the two speakers with the strongest evidence of an additional tongue lengthening in alveolar-velar sequences are the two speakers who overall showed the least alveolar reductions. It is therefore tempting to speculate that the more frequent and extreme alveolar reductions of the other speakers are partially the manifestation of an alternative strategy to cope with the same articulatory requirements, namely to furnish enough space for velar stop productions.

Another factor that seemed to foster the occurrence of assimilations relates to the preceding vocalic environment. All subjects showed an increased disposition to produce highly assimilated alveolars in the context of V1 = /ɪ/. A potential reason for this context-dependent asymmetry could be related to the articulatory precision required for vowel production. As has been shown previously (Perkell and Cohen, 1989), variability of articulatory configurations is smallest at the location in the vocal tract at which maximal constriction is produced, and the formation of a constriction for a palatal vowel such as /ɪ/ strongly constrains the position of the whole tongue, so that little variability induced by neighbouring consonants is observed (Hoole and Kühnert, 1995). Similar regularities were found in the present data (cf. Kühnert, 1996, p. 305). This implies that when the preceding vowel is /ɪ/ the tongue has little freedom to anticipate the articulation of the following alveolar, and with an increase in speech rate of alveolar-velar articulations, it might become increasingly more difficult for the tongue tip to even execute a residual coronal raising gesture.

A final remark concerns the notion of 'gestural overlap'. Apart from the relative overlap of alveolar and velar occlusion, the movement recordings in this study did not allow us to specify any concrete parameter or interval which could have been consistently applied to the data of all speakers. Moreover, as the speaker-specific kinematic parameters during alveolar-velar sequences at normal rate demonstrate, overlap between alveolar and velar speech movement and overlap between alveolar and velar occlusion can be implemented in very different ways. Such inter-individual differences in movement initiation and duration have been observed in other studies (e.g. Perkell and Matthies, 1992) and make the investigation of phasing relations between gestures on the basis of actual movement data exceedingly difficult. Yet another complication is that it is not clear to what underlying principle the observed overlap of speech movements actually testifies. Often considered as an indicator of increased casualness and rate within the broader framework of Articulatory Phonology, we have seen that the early initiation of the velar movement in alveolar-velar sequences of the German subjects could as well be attributed to an articulatory strategy to begin the movement as soon as the articulator is allowed to do so, and as such would support elements of a quite different model of speech production, more associated with central planning.

In summary, our data does not provide any unambiguous evidence that place assimilations in connected speech are either centrally programmed or the result of some kind of blending process of concurrent demands on the speech articulators. However, a division that assigns any gradual gesture to phonetics, and any apparent deletion to phonology might sometimes artificially cut across the same underlying principle. As we tried to show, there are a number of mechanical and dynamic factors of the involved articulations, which help to understand the observed alveolar reductions, even though speakers do employ a variety of means of implementing them. As we also tried to show, it still remains a delicate matter to distinguish kinematic patterns, which are attributable to a process of central planning, from those which arise from more dynamic factors of speech production.

Acknowledgement

The research was supported by a grant of the Studienstiftung des Deutschen Volkes awarded to the first author while at the University of Munich. The authors are grateful to Hans Tillmann, Francis Nolan, and Fiona Gibbon for their insightful comments and assistance. We also thank our speakers for their generous participation.

References

Barry, M., 1991, Temporal modelling of gestures in articulatory assimilation. *Proceedings of the XIIth International Congress of Phonetic Sciences*, **4**, 14–19.

Barry, M., 1992, Palatalization, assimilation and gestural weakening in connected speech. *Speech Communication*, **11**, 393–400.

Browman, C. and Goldstein, L., 1986, Towards an articulatory phonology. *Phonological Yearbook*, **3**, 219–252.

Browman, C. and Goldstein, L., 1990, Tiers in articulatory phonology with some implications for casual speech. In J. Kingston and M. Beckman (Eds), *Papers in Laboratory Phonology I* (Cambridge: Cambridge University Press), pp. 341–367.

Browman, C. and Goldstein, L., 1992, Articulatory phonology: an overview. *Phonetica*, **94**, 155–180.

Byrd, D. and Tan, C., 1996, Saying consonant clusters quickly. *Journal of Phonetics*, **24**, 263–282.

Byrd, D., 1996a, Influences on articulatory timing in consonant sequences. *Journal of Phonetics*, **24**, 209–244.

Byrd, D., 1996b, A phase window framework of articulatory timing. *Phonology*, **13**, 139–169.

Crystal, T. and House, A., 1988, The duration of American-English stop consonants: an overview. *Journal of Phonetics*, **16**, 285–294.

Ellis, L. and Hardcastle, W., 2002, Categorical and gradient properties in alveolar and velar sequences: evidence from EPG and EMA data. *Journal of Phonetics*, **30**, 373–396.

Hayes, B., 1992, Comments on chapter 10. In D. Ladd and G. Docherty (Eds), *Papers in Laboratory Phonology II* (Cambridge: Cambridge University Press), pp. 280–286.

Henke, W., 1966, Dynamic Articulatory Model of Speech Production Using Computer Simulation. Doctoral Dissertation, MIT, Cambridge, MA.

Holst, T. and Nolan, F., 1995, The influence of syntactic structure on [s] to [ʃ] assimilation. In B. Conell and A. Avanti (Eds), *Papers in Laboratory Phonology IV* (Cambridge: Cambridge University Press), pp. 315–333.

Hoole, P. and Kühnert, B., 1995, Patterns of lingual variability in German vowel production. *Proceedings XIIIth International Conference of Phonetic Sciences, Stockholm*, **2**, 442–445.

Hoole, P., 1996, Issues in the acquisition, processing, reduction and parameterization of articulographic data. *Forschungsberichte des Insituts für Phonetik und Sprachliche Kommunikation der Universität München*, **34**, 158–173.

Kuehn, D. and Moll, K., 1976, A cineradiographic study of VC and CV articulatory velocities. *Journal of Phonetics*, **4**, 303–320.

Kühnert, B., 1993, Some kinematic aspects of alveolar-velar assimilations. *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München*, **31**, 263–272.

Kühnert, B., 1996, Die alveolar-velar Assimilation bei Sprechern des Deutschen und des Englischen – kinematische und perzeptive Grundlagen. *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München*, **34**, 175–335.

Löfqvist, A. and Gracco, V., 2002, Control of oral closure in lingual stop consonant production. *Journal of the Acoustical Society of America*, **111**, 2811–2827.

Mooshammer, C., Hoole, P. and Kühnert, B., 1995, On loops. *Journal of Phonetics*, **23**, 3–21.

Munhall, K., Ostry, D. and Flanagan, J., 1991, Coordinate spaces in speech planning. *Journal of Phonetics*, **19**, 293–307.

Nolan, F., 1992, The descriptive role of segments: evidence from assimilation. In D. Ladd and G. Docherty (Eds), *Papers in Laboratory Phonology II* (Cambridge: Cambridge University Press), pp. 261–280.

Nolan, F., Holst, T. and Kühnert, B., 1996, Modelling [s] to [ʃ] accommodation in English. *Journal of Phonetics*, **24**, 113–137.

Ostry, D. and Munhall, K., 1985, Control of rate and duration of speech movements. *Journal of the Acoustical Society of America*, **77**, 640–648.

Paradis, C. and Prunet, J.-F., 1991, Introduction: asymmetry and visibility in consonant articulation. In C. Paradis and J.-F. Prunet (Eds.), *Phonetics and Phonology: The Special Status of Coronals* (San Diego, CA: Academic Press), pp. 1–28.

Perkell, J. and Cohen, M., 1989, An indirect test of the quantal nature of speech in the production of /i/, /ɑ/ and /u/. *Journal of Phonetics*, **17**, 123–133.

Perkell, J. and Matthies, M., 1992, Temporal measures of anticipatory labial coarticulation for the vowel /u/: within- and cross-subject variability. *Journal of the Acoustical Society of America*, **91**, 2911–2925.

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

Perkell, J., Zandipour, M., Matthies, M. and Lane, H., 2002, Economy of effort in different speaking conditions. I. A preliminary study of intersubject differences and modeling issues. *Journal of the Acoustical Society of America*, **112**, 1627–1641.

SAS Institute Inc, 1998. *SAS/STAT User's Guide*, Release 6.03 Edition (Cary, NC: SAS Institute Inc).

Scobbie, J., 1995, What do we do when phonology is powerful enough to imitate phonetics? Comments on Zsiga. In B. Connell and A. Arvaniti. (Eds), *Papers in Laboratory Phonology IV* (Cambridge: Cambridge University Press), pp. 303–314.