
Volume 39
Issue 3
July 2011
ISSN 0095-4470
ELSEVIER
Journal of
Phonetics
Special Issue: Speech Reduction
Guest Editors:
Mirjam Ernestus and Natasha Warner
Editor
Kenneth de Jong
Associate Editors
Ocke-Schwen Bohn
Taehong Cho
Martine Grice
Jennifer Hay

Journal of Phonetics 39 (2011) 413–428

ELSEVIER

Contents lists available at ScienceDirect

Journal of Phonetics

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

Articulatory factors influencing regressive place assimilation across word boundaries in German

Marion Jaeger *, Philip Hoole

Institute of Phonetics and Speech Processing, Schellingstr. 3, D-80799 Munich, Germany

ARTICLE INFO

Article history:
Received 4 January 2009
Received in revised form
24 February 2011
Accepted 3 March 2011
Available online 16 April 2011

ABSTRACT

This EMA study examined articulatory properties of nasal–stop and stop–stop sequences across word boundaries in German. The immediate aim was to investigate whether the greater tendency of nasals as compared to stops to exhibit regressive place assimilation results from differences in tongue-tip reduction due, in turn, to acoustic–perceptual properties of nasals. In parallel with this aim, the relevance of word frequency for assimilatory processes was investigated. Analysis of data from four speakers showed greater tongue-tip reduction in high-frequency words with a nasal, indicating a combination of factors that causes tongue-tip reduction.

A further route to capturing assimilatory processes is the amount of overlap between C_1 and C_2. Analysis of this was complicated precisely by the fact that particularly for the high-frequency nasals, movement reduction was sometimes so strong that no kinematic analysis could be performed. A tentative conclusion was nonetheless that overlap tended to be greater in nasal–stop than stop–stop sequences.

The discussion points out that high-frequency words in German mostly end in a nasal and concludes that it is word frequency and speakers' knowledge of acoustic–perceptual properties of nasality that allows them to simplify articulation more in nasals than stops. This account presupposes that the nasality itself remains robustly present even when the lingual gesture of the nasal consonant is strongly reduced (or overlapped by the following consonant). Velum movement data available for some of the speakers and inspection of the acoustics indicated that this was indeed the case.

1. Introduction

This study investigated the articulatory patterns that underlie regressive place assimilation in consonant clusters across word boundaries in German. In such assimilations, a word-final consonant C_1 assimilates the place of articulation of a following word-initial consonant C_2. In current phonological accounts, articulatory patterns of place assimilation are often considered to fall into two classes: categorical and gradient. In categorical assimilations, C_1 is completely reduced, i.e., deleted, during prearticulatory processing, and the duration of C_2 is increased in time so as to occupy the slot formerly held by C_1 (Jun, 2004, p. 69). In gradient assimilations, however, C_1 must be partially reduced during prearticulatory processing, and the duration of C_2 is concomitantly extended, so that acoustic cues of C_1 can no longer be perceived. Gradient assimilations coincide with perceptual assimilations, in which a categorical change in the perception of a sound occurs as a result of spatio-temporal overlap of two consecutive gestures (Browman & Goldstein, 1992). In addition, a complete or partial reduction of the tongue-tip movement may or may not occur during gestural overlap (Barry, 1991; Byrd, 1996; Hardcastle, 1995; Nolan, 1992; Wright & Kerswill, 1989). A general decrease of speaking effort over the course of a unit has been considered to cause such gestural reductions (Browman & Goldstein, 1995).

A general decrease of speaking effort is most likely in words with a high word frequency or great lexical predictability. Several acoustic studies (Bybee & Scheibman, 1999; Gregory, Raymond, Bell, Fosler-Lussier, & Jurafsky, 1999; Jurafsky, Bell, Gregory, & Raymond, 2001) showed that there is more reduction in duration in words and phrases that are more frequent and hold greater predictability in spontaneous speech. The greater temporal reduction in words and phrases that are more frequent might be conceived as either due to faster articulation, articulatory reduction, sound deletion or increased C_1C_2 overlap. Previous instrumental investigations (mainly based on electropalatography) that explicitly considered frequency effects are few, and have given a mixed picture of the strength and nature of frequency effects (see e.g., Bergmann, accepted for publication; Mücke, Grice, & Kirst, 2008; Stephenson, 2003). Thus examination with further experimental techniques (in our case electromagnetic articulography) appears to be called for.

* Corresponding author. Tel.: +49 89 3248 9804; fax: +49 89 2180 5790.
E-mail address: jaeger@phonetik.uni-muenchen.de (M. Jaeger).

0095-4470/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.wocn.2011.03.002

Another feature of regressive place assimilation – apart from articulatory reduction – is that alveolar nasals show a greater inclination to assimilate than alveolar stops (for German see Jaeger, 2003; Kohler, 1990). In recent accounts (Jun, 2004; Steriade, 2001), this manner of articulation asymmetry in place assimilation is accounted for by the weaker acoustic cues in nasals than stops. According to Steriade (2001) it is above all the known pattern of perceptual confusion (Hura, Lindblom, & Diehl, 1992) that accounts for the greater inclination of nasals versus stops to assimilate more often. Speakers deliberately use their knowledge that it is perceptually more tolerable to modify the place of articulation of nasals than stops in place assimilation (Steriade, 2001, p. 233).[1] Steriade's discussion, however, leaves open what the precise nature of the articulatory modifications will actually be when speakers exploit this freedom. Will there be a categorical change in the place of articulation? Will the original movement be made more economical in physiological terms by a reduction in movement amplitude? Or could this freedom also take the form of increased overlap between adjacent articulations to exploit the potential of the vocal tract for efficient parallel transmission of segmental information?

Moreover, the weaker acoustic cues in nasals than stops led to a simple and fundamental motivation for the present study. Based on acoustic and perceptual (transcription) data, it is very difficult to compare articulatory reduction patterns across stops and nasals. For example, word-final alveolars may show similar amounts of weakening for nasals and stops, but because of the reduced acoustic saliency of the nasals, a given amount of movement amplitude reduction (or overlap with the following sound) may be more likely to lead to perception of the assimilated form in nasals than in stops. Thus, while it is highly plausible that the acoustics of nasals give speakers the potential for articulatory simplification, rather little is currently known about whether, and how, they actually exploit this freedom. In order to elaborate on the articulatory detail of place assimilation under different phonetic conditions, we directly examined the speech movements in assimilatory contexts by means of the electromagnetic articulograph (EMA, Carstens AG500).

Until now, articulatory studies investigating assimilations across word boundaries confined examination of place of articulation to either stop–stop sequences (Byrd, 1996) or nasal–stop sequences (Ellis and Hardcastle, 2002). In contrast, our study explicitly compared the movement patterns of both kinds of sequences, and combined this analysis with systematic variation of word frequency. Two further factors were also included in the design of the investigation: vowel context (palatal versus non-palatal) and place of articulation of C_2 (labial versus velar).

The motivation for the inclusion of the latter two factors was as follows: results of Kühnert and Hoole (2004) indicated that vowel context (here meaning the vowel preceding the consonant sequence) could have an influence on tongue-tip reductions and/or on patterns of C_1C_2 overlap. However, this potential influence has not yet been extensively investigated. In tongue-tip#tongue-back sequences across word boundaries (indicated by #) in German they had found that tongue-tip excursions decreased more often in consonants following palatal than non-palatal vowels. They attributed this to the strongly constraining influence of palatal vowels on the whole tongue shape, making it difficult for the speaker to anticipate the alveolar articulation in the preceding vowel (vowel context as a factor in our design will henceforward be referred to as V1_pal).

Place of articulation of C_2 (C_2_POA) as an influencing factor was probed because Zsiga (1994) had found that the influence of C_2 velars on the vowel transitions into C_1 alveolars was stronger than that of C_2 labials. However, it was not clear that this need be due to more pronounced C_1C_2 overlap for the C_2 velars. Rather, it was suggested that the link between tongue-back and tongue-tip could lead to a greater influence on the tongue-tip in the velar compared to the labial case. Whether this could take the form of greater reduction of the tongue-tip gesture in the case of velar C_2 is impossible to say on the basis of Zsiga's acoustic data. Since the influence of the place of articulation of the trigger consonant in assimilatory processes does not seem to have been extensively looked at in the meantime, we felt that new articulatory data would be useful. Accordingly the following analyses were carried out.

(1) *Spatial analyses*: Parametric tests to determine how manner of articulation of C_1, word frequency of w_1, vowel context, and place of articulation of C_2 might influence tongue-tip excursion in C_1C_2 sequences across word boundaries.
(2) *Timing analyses*: Overlap of C_1C_2 under different conditions. The initial question here was whether weaker acoustic cues for C_1 and high word frequency lead to a more overlapped mode of speech. Furthermore, we probed C_1C_2 overlap in combination with spatial patterns in the data. While Browman and Goldstein on the one hand have emphasized mutual sliding of gestures as a key mechanism leading to perceived assimilations, Jun on the other hand has argued that this on its own is not sufficient, but that reduction of gestural magnitude must occur as well to give an assimilated percept: "[…] the correct mechanism appears to be gestural reduction of the target consonant, with concomitant temporal extension of the trigger consonant" (2004, p. 70). The latter formulation suggests a close link between changes in gestural magnitude and overlap. As C_1 reduces and C_2 extends, then C_2 will presumably overlap this residual C_1 more and more. But there is a further source of uncertainty here because the assumption of temporal extension of C_2 as a key mechanism may not be justified. Even in cases where C_1 magnitude has reduced to zero (i.e., where C_2 extension may be expected to be clearest), quite variable results have been reported (Ellis & Hardcastle, 2002; Kochetov & Pouplier, 2008). In short, there is much that is still unclear about the relative importance of the various mechanisms such as reduction of gestural magnitude of C_1, increase in overlap of C_1 and C_2, and temporal extension of C_2, and in particular about the strength of their interrelationships.

Rounding up this introductory section, the point of departure of the present study was the unknown nature of the articulatory modifications that underlie the manner of articulation asymmetry in German. Accordingly, we focused on the articulatory processes under systematically varied phonetic and lexical conditions by means of electromagnetic articulography in nasal–stop and stop–stop sequences across word boundaries. Because of the inescapable fact that in German words with a high word frequency more often end in a nasal than a stop (Jaeger, 2003), our basic predictions were as follows: high-frequency words ending in an alveolar nasal will exhibit greater tongue-tip reduction than high-frequency words with an alveolar stop or low-frequency words with an alveolar nasal or stop. In addition, a palatal vowel context or a velar C_2 context will make tongue-tip reductions even more likely under each condition.

2. Materials and methods

2.1. Speaker and speech material

We report the results of a reading task for four speakers, three females and one male (S1, S2, S3, and S4). The speech material

[1] For a different view regarding the perceptibility of place of articulation in nasals versus stops, see Winters (2001).

Table 1
Target items with estimated frequencies (per billion) and respective control items.[a]

C_1_MOA	V_1_pal	w_1_freq	Target C_1 alveolar_C_2 velar	Target C_1 alveolar_C_2 labial
nasal	non-palatal	high	dann kann (1164184, 10650) 'then (it) can'	dann passt (1164184, 4) 'then (it) fits'
		low	Zahn kann (7637, 12) 'tooth can'	Zahn passt (7637, 0) 'tooth fits'
	palatal	high	zehn Kappen (810447, 0) 'ten caps'	zehn Palmen (810447, 2) 'ten palms'
		low	Delfin kann (3787, 0) 'dolphin can'	Delfin passt (3787, 0) 'dolphin fits'
oral	non-palatal	high	statt kommende (447743, 20) 'instead of next'	statt Padau (447743, 0) 'instead of Padua'
		low	Blatt kann (25631, 35) 'page can'	Blatt passt (25631, 2) 'page matches'
	palatal	high	sieht kommende (272328, 1) 'next (week) looks'	sieht Padua (272328, 0) 'Padua views'
		low	Lied kann (27643, 70) 'song can'	Lied passt (27643, 4) 'song matches'

	V_1_pal	w_1_freq	Control C_1 alveolar_C_2 alveolar	Control C_1 velar_C_2 velar	Control C_1 labial_C_2 labial
nasal	non-palatal	high	dann nach	lang kann	kam mal
		low	Zahn nach		
	palatal	high	zehn nach	Ding kann	wem macht
		low	Delfin nach		
oral	non-palatal	high	statt das	Tag kann	hab bald
		low	Blatt da		
	palatal	high	sieht das	Weg kann	Jeep passt
		low	Lied da		

[a] Column 1: C_1 manner of articulation (C_1_MOA), column 2: vowel place of articulation (V_1_pal), column 3: word frequency (w_1_freq), on the top: target items with counts per billion (left number) and co-occurrence frequency (right number) in brackets.

analyzed for this study consisted of short sentences containing $V_1C_1\#C_2V_2$ sequences (# indicating a word boundary). The stimuli were selected for a balanced factorial design. V_1 was a non-palatal vowel /a/ or a tense palatal vowel /i:, e:/ preceding an alveolar stop /t/ or an alveolar nasal /n/ followed by a labial stop /p/ or a velar stop /k/. Table 1 also shows the corresponding control items where place of articulation was kept the same for C_1 and C_2. Word pairs were further divided into groups with high (> 250,000 per billion) and low (< 50,000) word frequencies of the first word. We used the corpora of contemporary written German of the Institute of German Language (IDS- Institut für Deutsche Sprache/ IDS-Korpus) to estimate both the frequency of each test word relative to one billion occurrences as well as the co-occurrence frequency of each word pair.[2] There was one word pair, namely "dann kann", in the group of words with a high word frequency that also featured a significant co-occurrence frequency (a measure similar to joint bigram probability, cf. Gregory et al. (1999) and Jurafsky et al. (2001).

All items were blockwise randomized. Each speaker repeated each item ten times in a carrier phrase that either did (S1 and S2) or did not (S3 and S4) change with each repetition.[3] The sentences were prompted on a computer screen, and the speakers were instructed to produce them at a self-determined but rapid speaking rate.

2.2. Recording and data analysis

Articulatory movements were monitored by means of a 3-D electromagnetic transduction device (Carstens AG500). Data was acquired by three sensors attached to the tongue (tip, mid, and back), one sensor attached to the upper- and lower-lip, the lower jaw, and the velum, respectively. The velum sensor was attached to the soft palate either directly or indirectly. With the indirect method (used with S1, S2, and S4) the velum sensor was mounted on a plastic strip. One end of the strip was glued to the hard palate while the other end (the one with the sensor) was placed slightly posterior of the intersection between the hard and soft palate, where movements of the velum could be registered. The plastic strip was held in place by the moisture of the mucosa. In contrast, with the direct method (used with S3), the velum sensor was glued directly to the velum slightly behind the intersection of the hard and soft palate. To permit the use of the direct method, the speaker's oral cavity had to be rather spacious and could therefore only be used with the male speaker. Additional sensors placed on the nasion, the upper incisors, and behind the left and right ears served as reference coils to permit compensation for head movements. The acoustic signal was recorded synchronously with the movement signals. A detailed discussion of data acquisition, normalization and preparation procedures is outlined in Hoole and Zierdt (2010).

Our focus here was on data from the lower-lip (ll), the tongue-tip (tt), the tongue-back (tb) and the velum transducer.

[2] The corpora of written German of the IDS with a total of 4 billion words (latest update August 2010) currently form world-wide the biggest electronically accessible collection of German contemporary texts. The 109 corpora of written German are searchable via COSMAS II—the Corpus Search, Management and Analysis System of the IDS. We used this application to determine both word frequency and co-occurrence frequency of neighboring words. The co-occurrence frequency of a word pair was used to determine the significance of the lexical cohesion of the two words. A co-occurrence frequency of 100 or more is treated as significant.

[3] Constantly changing carrier phrases were employed to diversify the reading task. However, the inconsistency of the carrier phrases intensified reading and concentration demands upon participants. Increasing slips of the tongue and repetitions were the inevitable consequence, which in turn extended our already long recording sessions (two hours on average, not counting preparation time). Thus, after a few recordings in which the carrier phrase changed with each prompt, we decided to go back to the usual setting with only one carrier phrase for each target item. In this process we also replaced some words in our speech material, i.e., the target word "Delfin" ('dolphin') was replaced by "Wien" ('Vienna') and the non-target words "Palmen" ('palms'), "Kappen" ('caps') and "Padua" were replaced by "Pasten" ('paste'), "Kasten" ('box') and "Passau". These changes were applied with S3 and S4.

The articulatory data was extracted from the $VC_1\#C_2V$ signal using Matlab routines to identify several kinematic landmarks in the displacement and velocity signals of the EMA recordings: on- and offset of C_1 and C_2 nucleus (constriction) phase, C_1 closing gesture of tt, tb and ll as well as C_1 opening gesture of the velum. On- and offsets were defined as points in time at which the tangential velocity exceeded or fell below a threshold value, i.e., 20% of its maximal speed above minimal velocity. Fig. 1 illustrates the parameters derived from the above mentioned landmarks.

To quantify the temporal overlap between C_1 and C_2, we measured the interval between the end of the nucleus of C_1 and the moment in time of the gesture onset of C_2 (equals onset overlap) (Chitoran, Goldstein, & Byrd, 2002; Kühnert & Hoole, 2004). The index is given as a percentage relative to the overall nucleus time, defined as the interval between C_1 nucleus onset and C_2 nucleus offset. A negative percentage of overlap indicates no overlap (signifying a time lag, with the lag increasing as values become more negative), whereas a positive percentage identifies overlap between C_1 and C_2 (with more overlap, more positive values).

As presented in more detail below, there were a fair number of target items in which reduction of tt movement was so strong that kinematically defined landmarks could not be reliably determined in the measurement signal. Therefore it was not possible to measure movement displacement as the difference in position between movement onset and the point of maximum constriction. However, it did not seem justified to simply regard displacement as exactly equal to the corresponding controls in such cases.

On this account, in order to derive a measure of tt behavior that is defined in all items where an alveolar is lexically present, we extracted the position of the tt at a point 25% into the acoustically defined closure phase of the C_1C_2 sequence. The choice of this time point was based on the idea that it would correspond roughly to the midpoint of C_1, given that the complete consonantal sequence consists of two consonants. The well-foundedness of this choice was cross-checked by determining the relative position in the C_1C_2 sequence where maximum constriction for C_1 occurred in items with well-defined tt movement patterns. For two speakers (S1 and S4) the result was indeed very close to 25%. For the other two speakers (S2 and S3) the empirically determined relative time of maximum constriction was somewhat later, i.e., close to 35%. For S2 the later time point gave extremely similar position data to the 25% time point (actually not surprising because movement velocity will be low in the vicinity of maximum constriction) and for S3 the 35% time point showed somewhat stronger movement reductions in the assimilation contexts. Thus for further analysis and display the

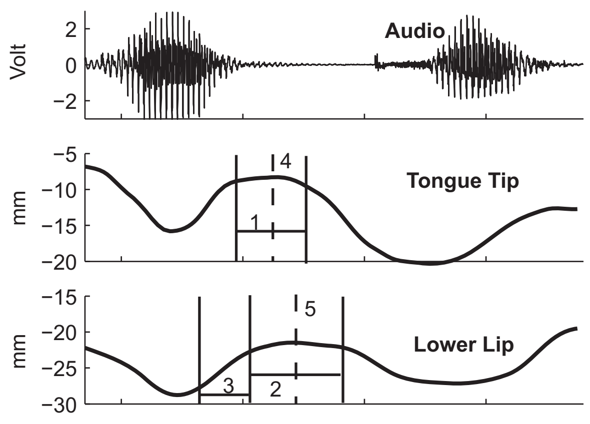

Fig. 1. Kinematic parameter extraction (illustration only) in one production of "Blatt past" produced by one speaker. Top: audio; middle: tongue-tip movement; bottom: lower-lip movement. 1: C_1 nucleus; 2: C_2 nucleus; 3: C_2 closing gesture; 4 and 5: constriction maxima.

use of a single time point for all speakers appeared an acceptable compromise that is, if anything, conservative with regard to the magnitude of the gestural reductions occurring in the target items. An alternative approach to extracting position data for all target items would have been to extract the maximum vertical position of the tt wherever it occurred in the C_1C_2 sequence. However, in cases with strong reduction of tt activity for the alveolar itself, this position could be strongly affected by articulations for the C_2 consonant (Zsiga, 1994), or even the preceding vowel in contexts with palatal vowel (Kühnert & Hoole, 2004). Thus an approach based on a specifically defined time point appeared preferable.

To describe the displacement of the velum gesture we applied a similar procedure (for more details see Section 3.4). Finally, to illustrate some of our non-statistical findings, the vertical positions of the articulators in question or the articulatory trajectories of the articulators in question were plotted. The statistical analyses applied to the data included tests of homogeneity of variance (Levene's Test) as well as parametric tests of variance (ANOVA).[4] Influencing factors (with their abbreviations in the ANOVA tables) were vowel context (V_1_pal), word frequency of w_1 (w_1_freq), manner of articulation of C_1 (C_1_MOA) and place of articulation of C_2 (C_2_POA). Dependent variables were vertical tt position and C_1C_2 overlap.

3. Results

We start the presentation of the results by considering the tt positions extracted at the acoustically defined fixed relative location in the C_1C_2 segment. This gives an immediate appreciation of the frequency and strength of reduction effects on the basis of the complete data set (Section 3.1). Following that, we look in Section 3.2 at patterns of overlap between the tt (C_1) and the C_2 articulator. The latter analyses have to leave out items where there was insufficient tt movement for kinematic analysis. Section 3.3 builds on the previous two sections by examining how closely reduction of tt movement and increase in C_1C_2 overlap are linked. The results section then concludes with consideration of velum movement (Section 3.4) and C_2 duration (Section 3.5).

3.1. Tongue-tip position

3.1.1. Frequency of partial and complete reductions

Using the control items, two thresholds were defined based on the vertical position of the tt (thresholds were defined separately for the two vowel contexts).

(1) Two standard deviations *below* the mean of the tt position in purely alveolar C_1C_2 control sequences.
Data from the target items located higher than this threshold were regarded as ***unreduced***.
(2) Two standard deviations *above* the mean of the tt position in purely labial and purely velar C_1C_2 control sequences.

Data from the target items located lower than this threshold were regarded as ***completely reduced***.

Data falling below threshold 1 but above threshold 2 were regarded as ***partially reduced*** (for a similar procedure see Kochetov and Pouplier (2008)).

[4] Variance was not always homogeneous, and we also applied non-parametric tests. Since the non-parametric results did not contradict the parametric ones, we do not report these results.

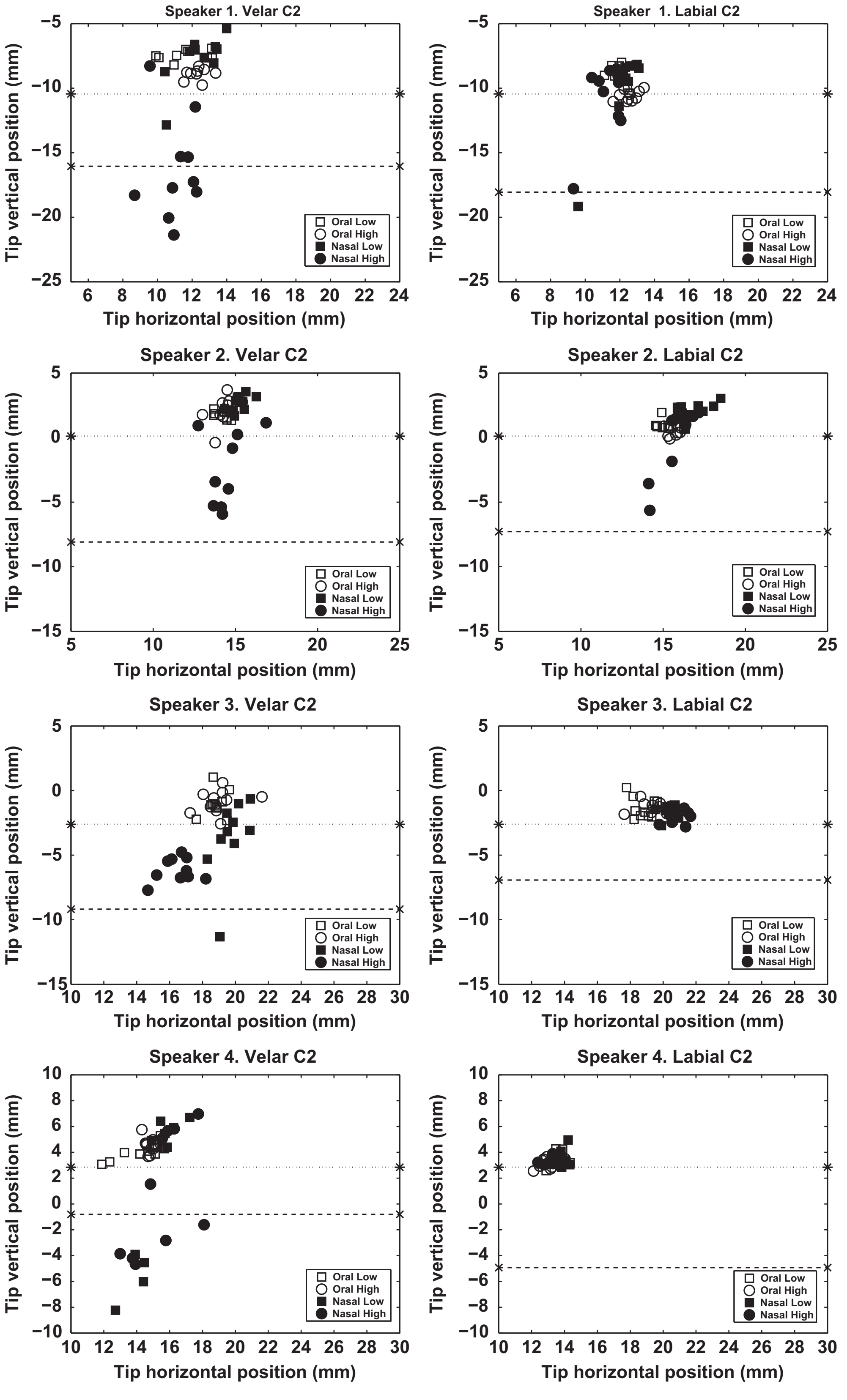

Fig. 2. Plot of vertical tongue-tip position of target items (non-palatal vowel context) at a point 25% into the acoustically defined C_1C_2 closure phase by speaker over C_2 place of articulation (left column: velar, right column: labial), C_1 manner of articulation (filled symbols: nasal, unfilled symbols: oral) and w_1 frequency (circle: high, quad: low). Reference point: upper incisors. Horizontal and vertical coordinates correspond to anterior–posterior and closed-open location within the oral cavity. Upper lines represent two standard deviations below the mean of the tongue-tip in purely alveolar C_1C_2 control items. Data from the target items located higher than this threshold were regarded as unreduced. The lower lines represent two standard deviations above the mean of the tongue-tip position in purely velar C_1C_2, resp., purely labial C_1C_2 control items. Data from the target items located lower than this threshold were regarded as completely reduced. Data falling between these two lines were regarded as partially reduced.

Fig. 2 shows scatter plots of the data from the non-palatal vowel context subdivided by speaker and C_2 place of articulation. The appropriate thresholds are shown as horizontal lines.

Before presenting the results of a series of ANOVAs that consider in detail the effect of all the independent variables with respect to tongue position, we first provide a succinct overview to give the reader some orientation when negotiating the more detailed results below. Accordingly, Table 2 summarizes the frequencies of the reduction categories defined by means of the thresholds, subdivided by the two independent variables of primary interest, namely manner of articulation of C_1 and word frequency. The counts are pooled over speakers, vowel context and place of articulation of C_2. The table shows, from left to right, the counts for complete reductions, partial reductions, and any kind of reduction (sum of first two columns). Separate figures are given for each combination of oral and nasal C_1 with high and low word frequency.

There was a grand total of 121 items in which reductions occurred. The total number of target items recorded was about 640 (16 target sequence types $\times$ 4 speakers $\times$ 10 repetitions), so reductions occurred in about 19% of all cases. Regardless of which column is considered, two points emerged very clearly: more reductions for nasal than oral, and more reductions for high than low frequency. Another point to notice is that complete reductions were almost totally absent for oral items, whereas the number of complete and partial reductions was fairly evenly balanced for the nasal items. Furthermore, there was one item in the corpus that made a particularly large contribution to the preponderance of reductions for high-frequency nasal C_1, namely the sequence "dann kann". This corresponds to the filled circles in the left column of Fig. 2. It will be seen that all four speakers had a large number of reductions for this category. It is worth noting here that this is the item for which the corpus analysis gave the highest co-occurrence frequency of the word pair.

Table 2
Sums of C_1 reductions over C_2 place of articulation, vowel height, and speaker.[a]

C_1_MOA	w_1_freq	Reductions		
		Complete	Partial	All
oral	low	0	11	11
	high	1	21	22
nasal	low	11	13	24
	high	28	36	64

[a] Columns 1 and 2: total of either complete or partial reductions; column 3: total of all reductions, broken down by C_1 manner of articulation (nasal vs. oral C_1) and word frequency (high vs. low).

3.1.2. Detailed analysis by means of ANOVA

For a more detailed quantitative analysis, we examined differences in sample means of the vertical tt positions (extracted at the 25% time point) for each speaker with respect to each of the independent variables and their interactions. As a preliminary step a four-way ANOVA (V_1_pal $\times$ w_1_freq $\times$ C_1_MOA $\times$ C_2_POA) was carried out.

The main effect of vowel was significant for three speakers (S1, S2, and S3), with a lower tt position occurring in non-palatal versus palatal vowel context. Taken by itself this does not have to mean that, contrary to the original hypothesis, more assimilated tokens occurred in non-palatal contexts because the threshold below which items would be categorized as completely reduced was also generally lower in the non-palatal case. In fact the palatal vowel context showed slightly more complete reductions than the non-palatal context (23 vs. 17), but fewer partial reductions (33 vs. 48). Still, this does not offer obvious support for the hypothesis of clearer assimilatory behavior in the palatal context.

Since patterns of variability clearly varied between the two vowel contexts, and word frequency and manner of articulation interacted with vowel context in all speakers, we decided for the main analyses to examine vowel contexts separately. The level of significance was raised to 1% to reduce the chances of a Type I error.

Non-palatal vowel context: As shown in Table 3, the vertical tt position was influenced by word frequency, manner of

Table 3
Mean vertical tongue-tip positions in the non-palatal vowel context.

Factors	Mean vertical tongue-tip position (difference between means in mm)[a]	
	Speaker S1, N=39	Speaker S2, N=39
C_1_MOA	nasal < oral (2.3)	n.s.
w_1_freq	high < low (3.1)	high < low (1.8)
C_2_POA	n.s.	n.s.
w_1_freq $\times$ C_1_MOA	nasal low–nasal high = 4.6	nasal low–nasal high = 3.5
	oral low–oral high = 1.8	oral low–oral high = 0.1
C_2_POA $\times$ w_1_freq	velar low–velar high = 5.0	n.s.
	labial low–labial high = 1.2	n.s.
C_1_MOA $\times$ C_2_POA	nasal labial–nasal velar = 1.6	nasal labial–nasal velar = 0.7
	oral velar–oral labial = 1.5	oral velar–oral labial = 1.3
	Speaker S3, N=40	Speaker S4, N=40
C_1_MOA	nasal < oral (2.3)	nasal < oral (1.6)
w_1_freq	n.s.	n.s.
C_2_POA	velar < labial (1.3)	n.s.
w_1_freq $\times$ C_1_MOA	nasal low–nasal high = 1.4	n.s
	oral high–oral low = 0.1	n.s.
C_2_POA $\times$ w_1_freq	n.s.	n.s.
	n.s.	n.s.
C_1_MOA $\times$ C_2_POA	nasal labial–nasal velar = 3.1	nasal labial–nasal velar = 2.5
	oral velar–oral labial = 0.3	oral velar–oral labial = 1.0

[a] Level of significance: $p < 0.01$. Abbreviations: <: lower than, >: higher than. n.s. = non-significant. For significant interactions, the difference between the means of the relevant pairs is given in mm, always in the order of higher tongue-tip position minus lower tongue-tip position.

articulation and place of articulation in the following way: tt position was lower in high than low-frequency words (S1, S2; S3 at $p<0.02$), and lower in nasal than oral C_1 items (S1, S3, and S4). Only one speaker (S3) showed a significant effect of C_2_POA (velar lower than labial; but see additional note on interactions with C_1_MOA below). In addition, a w_1_freq $\times$ C_1_MOA interaction was found for three speakers (S1, S2, and S3) indicating that the C_1_MOA effect was salient only in high-frequency words, i.e., lower vertical tt positions in nasal high-frequency words as compared to stop high-frequency words. C_1_MOA $\times$ C_2_POA interactions were present in all speakers. Thus, in nasal items the tt position was lower in velar C_2 items, whereas in stop items, the tt position was lower in labial C_2 items. The former bias is most readily explained by the item "dann kann", whereas the latter bias might be due to higher jaw positions in labial versus velar items. To make the pattern underlying each significant interaction evident, we provide in Table 3 the relevant pair-wise differences for all significant two-way interactions rather than performing additional post-hoc analyses.

Looking at speakers individually, we found the following: For S1, a w_1_freq $\times$ C_1_MOA $\times$ C_2_POA interaction in combination with a main effect of word frequency, a main effect of manner of articulation but no main effect of place of articulation suggested that "dann kann" exhibited the lowest tt position. For S2, interactions of w_1_freq $\times$ C_1_MOA and C_1_MOA $\times$ C_2_POA in combination with a main word frequency effect indicated that for this speaker, high word frequency and to a lesser degree articulatory factors such as manner and place influenced the vertical tt position. For S3, interactions of w_1_freq $\times$ C_1_MOA and C_1_MOA $\times$ C_2_POA were obtained. In contrast to S2, main effects of C_1_MOA and C_2_POA suggested that for S3, articulatory parameters such as manner and place yielded the most influence. Finally, for S4, manner of articulation in a velar context seemed to matter for vertical tt displacement. Because of an additional C_1_MOA $\times$ C_2_POA interaction but no word frequency effect (see Fig. 3), it seems that for this speaker manner of articulation in a velar context was, in fact, the only underlying cause for tt reduction.

In short, the results presented in this section showed an intricate pattern of interaction effects. It is important to emphasize that this actually reflects a very consistent feature in the data. Strong reduction effects were concentrated on just one of the eight possible combinations of the levels of the three factors (nasal C_1, velar C_2, and high frequency), with – in a very speaker-specific manner – at most one further combination also strongly affected. In other words, there were no cases where all four combinations involving nasal C_1, or all four involving high-frequency words, or all four involving velar C_2 reduced in parallel.

Palatal vowel context: In S1, the vertical tt position was not influenced by any of the independent factors. The observed lower tt position differences in labial versus velar C_2 items in S2 were small and barely exceeded an assumed measurement error of 0.5 mm (Hoole, Zierdt, & Geng, 2003; Zierdt, Hoole, & Tillmann, 1999). For S3, significant differences in tt position were obtained for all independent factors. However, as in S2, these differences, with the exception of a lower tt position in velar versus labial C_2 items, barely exceeded the assumed measurement error. Only when we looked at interactions did tt position differences become meaningful. Items with a velar C_2_POA yielded lower vertical tt positions. The lowest vertical tt positions were seen in items with a high word frequency, a nasal C_1_MOA and a velar C_2_POA. For speaker S4, palatal items with a high w_1_freq, a nasal C_1_MOA or a velar C_2_POA showed lower tt positions than palatal items with a low w_1_freq, an oral C_1_MOA or a labial C_2_POA. The differences were on average 1 mm. Because w_1_freq and C_2_POA did not interact and according to the differences in means in velar and labial contexts (see Table 4), these findings indicate that the word pair "zehn kann" ('ten can') caused the lower tt position.

In summary: The tt positions of all items clearly showed that tt position was affected by word frequency and manner of articulation of C_1. Statistical tests confirmed and differentiated this observation. In the non-palatal vowel context, the vertical tt position was on average lower in high-frequency words with a nasal than a oral C_1. Despite this similarity each speaker stressed influencing factors differently. For S1, co-occurrence frequency rather than word frequency might have played a decisive role in tt reduction because this was the speaker in which reduction was concentrated most exclusively on the "dann kann" item. For S2, word frequency showed the greatest influence. For S3, a lowering of the tt position mostly occurred in velar C_2 context. S4 also showed some place of articulation effect. For her, tt reduction in the velar C_2 context occurred in both frequency conditions, albeit only when C_1 was nasal. With regard to the palatal vowel context,

Table 4
Mean vertical tongue-tip positions in the palatal vowel context.

Factors	Mean vertical tongue-tip position (difference between means in mm)[a]	
	Speaker S1, $N=39$	Speaker S2, $N=40$
C_1_MOA	n.s.	n.s.
w_1_freq	n.s.	n.s.
C_2_POA	n.s.	velar < labial (0.7)
w_1_freq $\times$ C_1_MOA	n.s	n.s.
C_2_POA $\times$ w_1_freq	n.s	n.s.
C_1_MOA $\times$ C_2_POA	n.s	n.s.
	Speaker S3, $N=39$	Speaker S4, $N=40$
C_1_MOA	nasal < oral (0.7)	nasal < oral (1.0)
w_1_freq	high < low (0.7)	high < low (0.8)
C_2_POA	velar < labial (1.0)	velar < labial (0.9)
w_1_freq $\times$ C_1_MOA	nasal low–nasal high = 1.3 oral low–oral high = 0.4	nasal low–nasal high = 1.6 oral low–oral high = 0.0
C_2_POA $\times$ w_1_freq	velar low–velar high = 1.8 labial high–labial low = 0.1	n.s.
C_1_MOA $\times$ C_2_POA	nasal labial–nasal velar = 2.1 oral velar–oral labial = 0.2	nasal labial–nasal velar = 2.3 oral velar–oral labial = 0.5

[a] Level of significance: $p<0.01$. Abbreviations: <: lower than, >: higher than, n.s.: non-significant. For significant interactions, the difference between the means of the relevant pairs is given in mm, always in the order of higher tongue-tip position minus lower tongue-tip position.

statistically significant differences were so small that they cast doubt as to whether they were functionally meaningful in two speakers (S1 and S2). For S3, the influence of the velar stop on tt reduction was most salient in high frequency and nasal C_1 contexts, and for S4 in nasal C_1 context only.

3.2. Consonantal overlap

Before reporting on how vowel context, word frequency and manner of articulation influenced C_1C_2 overlap, we first discuss some of the difficulties encountered in measuring C_1C_2 overlap under the assumptions made using a gestural model of regressive place assimilation (Browman & Goldstein, 1992).

In a gestural account of regressive place assimilation, it is assumed that an increase in the spatial-temporal overlap of two consecutive gestures will eventually produce the effect of one of the gestures being hidden and perceived as deleted. However, several studies of regressive place assimilation (Barry, 1991; Byrd, 1996; Hardcastle, 1995; Nolan, 1992; Wright & Kerswill, 1989) showed that there are many instances of temporally reduced C_1C_2 sequences in which a partial or complete reduction of the tt gesture of the word-final C_1 concomitantly occurred. Now, when calculating any measure of C_1C_2 overlap, it has to be assumed that a tt displacement and, correspondingly, a nucleus phase in the overt articulation can be detected and measured. However, C_1C_2 sequences that exhibit complete articulatory reduction can obviously not be accounted for that way. Moreover, since it is not implausible that overlap increases as gestural magnitude decreases, it would clearly be misleading to regard the unanalyzable items as random missing data. Thus, it is necessary to bear in mind that the results presented in this section could be biased by their omission. In Section 3.2.3 we hence suggest a supplementary analysis that can be applied to all items with nasal C_1 regardless of the amount of tt reduction and which gives a rough estimate of how serious any bias in the present section is likely to be.

A complete tabulation of the numbers of non-measurable tt excursions for each speaker and word pair is given in Appendix A.

Essentially, in the non-palatal vowel context the vast majority of non-measurable items were those with a high word frequency and a nasal C_1 (items in Fig. 2 located below the lower threshold) The palatal vowel context (not shown in the figure) actually had a slightly higher total of unanalyzable cases than the non-palatal one – not surprisingly as the overall range of movement from vowel to alveolar consonant is much more restricted in the palatal context (Perkell & Cohen, 1989) – but a clear effect of phonetic context was not present, though most cases occurred with high-frequency words.

We present the results for tt–tb and tt–ll sequences separately, because it is not necessarily meaningful to combine overlap results that are based on different articulators. Results are summarized in Tables 5 and 6, respectively (scatter plots of tt position versus C_1C_2 overlap can be found in Section 3.3). Note that in these tables the number of analyzed cases (*N*) is given for each level of each factor. So in cases where this is not equal (for example if *N* is lower for high-frequency than low-frequency) this indicates the presence of non-measurable items.

3.2.1. tt–tb onset overlap

A three-way ANOVA (V_1_pal × w_1_freq × C_1_MOA) obtained a main effect of vowel context for all four speakers, suggesting greater tt–tb overlap in non-palatal vowel contexts. Similarly to the previous section, vowel context interacted with manner of articulation and/or word frequency in three of four speakers, so we present the detailed results based on separate tests of the non-palatal and palatal contexts. Level of significance was set to 1%.

In the *non-palatal* vowel context, a main effect of C_1_MOA was obtained for S2 and S4, showing greater tt–tb onset overlap in nasal than oral C_1 items. For S1, a main effect of w_1_freq and a w_1_freq × C_1_MOA interaction were seen, indicating that the greater tt–tb onset overlap in high- versus low-frequency words was only salient in words with a nasal C_1. For S3, no effects were obtained.

In the *palatal* vowel context, no main effects or interactions were obtained for any of the speakers.

3.2.2. tt–ll onset overlap

Just as for tt–tb overlap, a preliminary three-way ANOVA (V_1_pal × w_1_freq × C_1_MOA) obtained a main effect of V_1_pal for all speakers (more overlap in non-palatal context). And again vowel context interacted with word frequency or manner of

Table 5
C_1C_2 onset overlap by vowel context for items in which overlap could be calculated, velar C_2.

V_1_pal	Factors	tt_tb_onset_overlap (differences between means in %)[a]			
		Speaker S1		Speaker S2	
		N	Onset	*N*	Onset
non-palatal	w_1_freq	11 20	high > low (31)	15 20	n.s.
	C_1_MOA	12 19	n.s.	15 20	nasal > oral (26)
	w_1_freq × C_1_MOA		nasal high–nasal low=64		n.s.
			oral low–oral high=2		
palatal		Non-significant		Non-significant	
		Speaker S3		Speaker S4	
		N	Onset	*N*	Onset
non-palatal	w_1_freq	10 19	n.s.	18 19	n.s.
	C_1_MOA	12 17	n.s.	17 20	nasal > oral (38)
	w_1_freq × C_1_MOA	n.s.		n.s.	
palatal		Non-significant		Non-significant	

[a] Level of significance: $p < 0.01$. Abbreviations: $<$: smaller than, $>$: greater than, n.s.: non-significant. Order of number of items (*N*) (e.g., 11 20) corresponds to high, low and nasal–oral, respectively. For significant interactions, the difference between the means of the relevant pairs is given in percent, always in the order of greater onset overlap minus smaller onset overlap.

Table 6
C_1C_2 onset overlap by vowel context, labial C_2.

V1_pal	Factors	tt_ll_onset_overlap (difference between means in %)[a]			
		Speaker S1		Speaker S2	
		N	Onset	*N*	Onset
non-palatal	w_1_freq	19 20	n.s.	17 19	n.s.
	C_1_MOA	19 20	n.s.	18 18	nasal > oral (19)
	w_1_freq × C_1_MOA		n.s.		n.s.
		Speaker S3		Speaker S4	
		N	Onset	*N*	Onset
	w_1_freq	20 20	n.s.	20 20	high < low (9)
	C_1_MOA	20 20	nasal > oral (23)	20 20	nasal > oral (11)
	w_1_freq × C_1_MOA		n.s.		nasal low–nasal high = 16 oral low–oral high = 2
		Speaker S1		Speaker S2	
		N	Onset	*N*	Onset
palatal	w_1_freq	18 18	high > low (27)	20 20	n.s.
	C_1_MOA	20 16	nasal > oral (14)	20 20	nasal > oral (13)
	w_1_freq × C_1_MOA		n.s.		nasal high–nasal low = 35 oral low–oral high = 34
		Speaker S3		Speaker S4	
		N	Onset	*N*	Onset
	w_1_freq	16 20	high > low (17)	13 19	n.s.
	C_1_MOA	20 16	n.s.	12 20	n.s
	w_1_freq × C_1_MOA		nasal high–nasal low = 36 oral low–oral high = 3		nasal high–nasal low = 29 oral low–oral high = 21

[a] Level of significance: $p < 0.01$. Abbreviations: < : smaller than, > : greater than, n.s.: non-significant. Order of number of items (*N*) (e.g., 20 16) corresponds to high, low and nasal–oral, respectively. For significant interactions, the difference between the means of the relevant pairs is given in percent, always in the order of greater onset overlap minus smaller onset overlap.

articulation, this time for all speakers. Accordingly, detailed results are presented separately for the two vowel contexts.

In the *non-palatal* vowel context, a main effect of C_1_MOA was obtained for S2, S3 and S4, revealing greater tt–ll onset overlap in nasal than oral C_1 items. S1 showed a tendency towards greater tt–ll onset overlap in nasal versus oral C_1 items. For S4, a main word frequency effect was obtained, indicating, surprisingly, greater tt–ll onset overlap in low than high-frequency words. However, an additional w_1_freq × C_1_MOA interaction indicated that this word frequency effect was only salient in the nasal C_1 context.

In the *palatal* vowel context, a main effect of w_1_freq for S1 and S3 was obtained that showed greater tt–ll onset overlap in high- versus low-frequency words. An additional w_1_freq × C_1_MOA interaction for these speakers indicated that the word frequency effect was more salient in nasal than oral C_1 items. For S2, a main effect of C_1_MOA was seen, suggesting greater tt–ll onset overlap in nasal than oral C_1 items. Finally, for S4 a w_1_freq × C_1_MOA interaction was obtained but no main effects.

In summary: Overall, tt–tb and tt–ll onset overlap were greater in non-palatal than palatal vowel contexts. However, vowel context interacted with word frequency, manner of articulation or both in all speakers suggesting that manner of articulation and/or word frequency effects were not independent of vowel context. Looking at each vowel context separately, results were as follows. In the non-palatal vowel context, tt–tb onset overlap as well as tt–ll onset overlap were mainly influenced by manner of articulation of C_1 (more overlap in nasals than stops). In the palatal vowel context none of the factors seemed to influence tt–tb onset overlap in particular. The situation was a little bit more intricate for tt–ll onset overlap. Depending on the speaker, either manner of articulation (nasal) or word frequency (high), or their combination, led sporadically to greater tt–ll onset overlaps.

On the face of it, it seems that manner of articulation is what matters for C_1C_2 overlap. In the next section we consider whether this result may be biased by the non-analyzable items. Regarding word frequency, one reason why significant effects were found less consistently than for manner of articulation may be that the word frequency differences were simply not distinct enough – except in the specific case of "dann kann", which in turn had so many strongly reduced items that overlap was often not measurable.

3.2.3. *Confirming the reliability of C_1C_2 overlap results*

The previous section found some support for the hypothesis that greater overlap of C_1 and C_2 will be found when C_1 is nasal rather than oral. However, the results could be biased because the items with very strong reduction of movement amplitude, and thus not analyzable kinematically, were mostly nasal C_1. The bias could actually be conservative with respect to that hypothesis because it is quite plausible that strongly reduced items would, if measurable, be found to have high overlap. In search of independent evidence that the present result is a conservative one, we adopted the following procedure: Given that the velum opening movement for nasal C_1 remains robustly present even when the tt gesture reduces strongly (see Sections 3.4 and 3.5), and further assuming that the velum and tt gestures maintain the same timing relationship when C_1C_2 overlap increases, then the stronger the overlap between C_1 and C_2 the later will the soft palate raise to generate the oral occlusion for C_2. In cases where C_2 is a voiceless plosive this will be reflected in the acoustic signal by the duration of the non-nasal occlusion phase. This leads to a prediction that can be tested for all items, regardless of how strongly the tt movement reduces[5]: If nasal items with strongly reduced tongue movement are underlyingly characterized by high rather than low overlap, then shorter duration of the non-nasal occlusion phase for C_2 should accompany a lower position of the tt for C_1.[6]

To operationalize this we looked for positive correlations between these two variables for target word pairs with C_1 nasal, C_2 voiceless plosive, and where the standard deviation of tt vertical position exceeded 1 mm (word pairs characterized by a more restricted range of tt positions would not be informative).

S4 had four relevant word pairs, all with positive correlations, of which three were significant at $p<0.05$. S2 had two cases, both positive, of which one was significant. S3 had three cases, all showing weak negative correlations, but none were significant. S1 had five cases, all weak (two positive and three negative) and not significant. Since the only statistically significant correlations were positive ones, we interpret this result as an indication that any items with strongly reduced tt excursion not included in the main overlap analysis above were at least unlikely to exhibit particularly ***low*** overlap. Thus, if the C_1C_2 overlap results reported in Sections 3.2.1 and 3.2.2 were biased, then it was ***not*** in the direction of finding higher overlap for nasals where this was not actually the case.

3.3. Tongue-tip reduction and degree of C_1C_2 overlap

We have seen in the previous sections that both spatial reduction and increased temporal overlap were particularly common for nasal C_1. The present section examines the strength of the link between these two processes. Fig. 3 shows scatter plots of the vertical position of the tt versus the overlap of the tt with the C_2 articulator for all speakers and for both C_2 categories. For sake of brevity only the non-palatal vowel context was plotted because the palatal vowel contexts were much more restricted both in the range of tt positions and in overlap values, and added nothing to the range of patterns observable for the non-palatal vowels.[7]

The expectation would be for negative correlations if tt reduction and increase of overlap are linked at the level of individual tokens. It is quite clear that a close link of this kind is not present. It is probably only for the tt–tb sequences of S4 (bottom left panel) that any resemblance to this hypothetical pattern was found. Even if we concede that the strength of the association might be underestimated here because the unanalyzable items with strong tt reduction and potentially high overlap could not be taken into consideration, there still remained several very clear patterns where the number of unanalyzed items was insufficient to affect the picture. Consider in particular the tt–ll data for S2, S3, and S4 (bottom three panels on right). There were clearly numerous C_1 nasal items with relatively high overlap values but no trace of spatial reduction. Conversely, for the tt–tb sequences of S3 there was clear spatial reduction for the nasals, but no increase in overlap.

In short, while nasal C_1 showed globally more reduction and more overlap than oral C_1, this connection did not apply at an item-by-item level. In other words, tt reduction and degree of overlap appear to represent independent components of connected speech processes. We come back to the implications of this finding in the final discussion.

3.4. Tongue-tip and velum excursion in nasal C_1 items

This section examines to what extent reduction of the tt closing gesture is paralleled by reduction in the velum opening gesture. In other words, is there any evidence in nasal C_1 items that the velum stays closer to its position in oral control items precisely in those cases where little tt raising occurs? Currently, there is very little information available about whether anatomically (fairly) independent articulators such as tongue and velum that are tightly coordinated in unreduced forms (Krakow, 1999) can become functionally decoupled when targeted by reduction processes. In the present case, one explanation for the nasal–oral asymmetry in assimilation processes would seem to require such a decoupling. Speakers only receive the freedom to reduce the tt gesture if they simultaneously ensure that nasality remains robustly present acoustically, by this means ensuring that the weakening of place of articulation information does not become too salient for the listener. Accordingly, this scenario would predict only a weak link between reduction of tt movement and reduction of velum movement.

We were able to obtain data on velum movement for two of the four speakers (S1 and S3). For S1 data was, however, only valid for the first six blocks of repetitions.[8]

Velum data was extracted for analysis as follows. First, the time point of the lowest velum position in the C_1C_2 segment of each utterance was determined, and the anterior–posterior and vertical coordinates of the velum sensor were extracted at this time point. Second, because velum movement shows highly correlated vertical and anterior–posterior components, a principal component analysis was carried out of the two-dimensional positions over all items of each speaker, and the measure of velum movement used for further analysis was the projection of the 2D-coordinates onto the first principal component. This reflects the overall range of velum movement better than would either the anterior–posterior or vertical dimension used on its own. We then examined the relationship between this measure of velum position and the vertical tt position at the 25% time point within the C_1C_2 segment.

Clearly, the only situations worth considering in detail are those where the tt showed a range of reduced and unreduced realizations. For both speakers this applied in particular to the

[5] In addition, this approach does not require measurement of velum movement itself, which as will be seen below was not successful for all speakers.

[6] We readily concede that the basic assumption on which this section is based (a stable timing relationship between velum and tt gestures) is at best plausible rather than conclusively demonstrable. However, it does have the advantage of once again being conservative with respect to the hypothesis that nasal items attract both reduced tongue-tip movement and higher overlap. Expressed in task-dynamic terms, as the gestures for nasal C_1 and oral C_2 increase their overlap, then there will be increasing competition between the two sounds for control of the velum articulator. This means that velar raising may occur even though gestural activation of C_1 is still present, and means in turn that using shortening of the non-nasal occlusion phase of C_2 as an indication of C_1C_2 overlap may tend to underestimate the extent of this overlap.

[7] The arrangement of this figure exactly parallels that of Fig. 2, but because now only those data points are plotted where tt movement was sufficient to allow analysis of overlap, the scaling of the *y*-axis has been changed for better readability.

[8] This speaker had been recorded using the plastic strip method; its attachment to the hard palate became loose in the last part of the experiment. Because of the non-elastic nature of the plastic strip, movements of the sensor are effectively restricted to a single 'orbit' in the sagittal plane. At the end of the sixth repetition this movement trace shifted forward several millimeters (and a consistently higher position for the non-nasal controls was no longer observed).

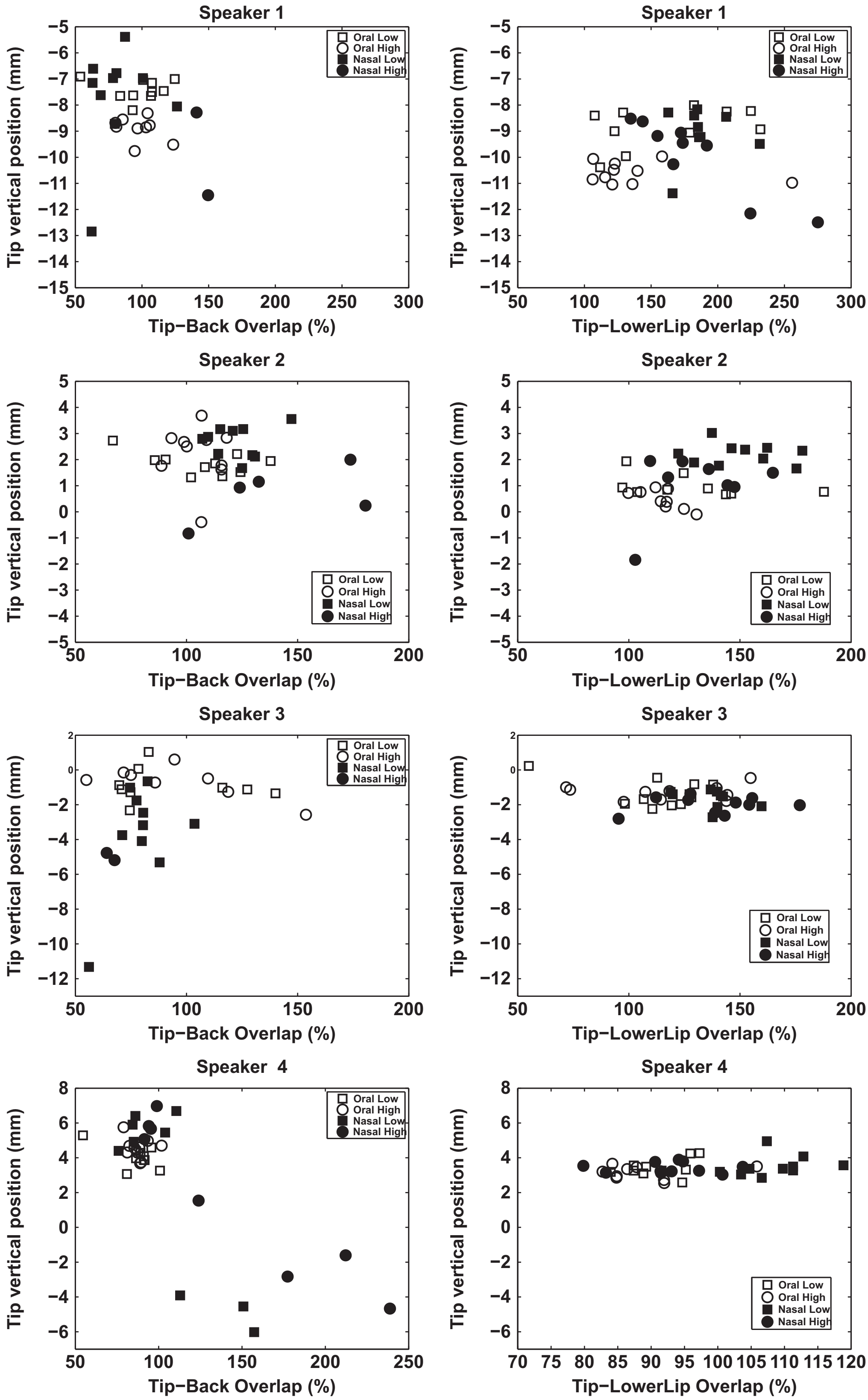

Fig. 3. Vertical tongue-tip position of target items (non-palatal vowel context) at a point 25% into the acoustically defined C_1C_2 closure phase plotted against C_1C_2 onset overlap for each speaker (S1–S4 from top to bottom). Left column: velar C_2, right column: labial C_2. Filled symbols: nasal C_1, unfilled symbols: oral C_1. Circles: high w_1 frequency, quads: low w_1 frequency).

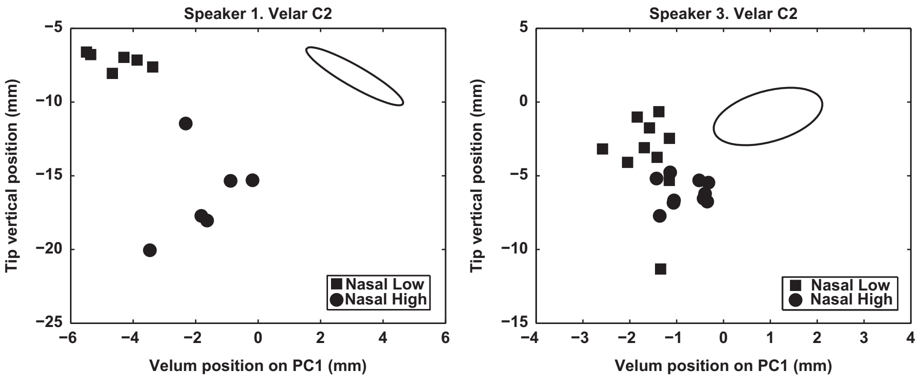

Fig. 4. The vertical tongue-tip position of nasal C_1 target items (non-palatal vowel context; velar C_2) at a point 25% into the acoustically defined C_1C_2 closure phase plotted against the lowest velum position in the C_1C_2 segment (defined as the first principal component ('PC1') of the vertical and anterior–posterior coordinates at the lowest velum position), divided by w_1 frequency. S1 (left) and S3 (right). Tongue-tip position vs. velum position of oral control items is shown as 2-sigma ellipses.

context with non-palatal vowels and velar C_2. The corresponding scatter plots are shown in Fig. 4. The nasal C_1 tokens are shown as individual data points. The range for the corresponding oral C_1 items is shown for reference using a two-sigma ellipse.

For S1, pooling over the high- and low-frequency nasal items there was a significant negative correlation between tongue and velum position ($r=-0.72$, $p<0.01$). Velum lowering was less strong (i.e., velum position was closer to the oral items) when the tt showed more reduction. Thus reduction did seem to proceed here to some extent in parallel. The situation was muddied by the fact that the relationship just within the high-frequency items (filled circles) was reversed. Note that the high-frequency item with the lowest tt position showed the lowest and thus most unambiguously nasal position of the velum.

For S3, the overall relationship was again negative. Notice here the group of high-frequency items whose velum position approaches the range of the oral items. However, the correlation coefficient just missed significance ($r=-0.43$, $p=0.061$).

In addition to the contexts just discussed, which showed the most continuous distribution of tt values for both speakers, one further context should be mentioned for each of the speakers: For S1 in the context with non-palatal vowels and labial C_2, there were two isolated items with strongly reduced tt positions (visible in the top right panel of Fig. 4). These showed absolutely no evidence of a parallel reduction in velum position ($r=0.16$, $p=0.63$). For S3, the context with palatal vowels and velar C_2 showed fairly clear reduction of the tt in the high-frequency items, but any parallelism with velum position was even weaker than in the corresponding non-palatal context shown in the figure, and nowhere near reaching significance ($r=-0.11$, $p=0.65$).

On balance then, while cases were found where velum opening for nasals was somewhat weaker when the tt gesture was reduced, there was no compelling evidence for a close coupling of the magnitude of reduction. Certainly there were no cases where elimination of the tt gesture could be accompanied by elimination of the velum gesture.

3.5. *Temporal expansion of C_2*

Finally, we address the question of whether reduction of tt movement for C_1 is associated with temporal expansion of the gesture for C_2. In the absence of observable tt movement C_2 durations might then be indistinguishable from those of control items with identical place of articulation for C_1 and C_2.

This question in particular allows the accounts of Articulatory Phonology on the one hand, and Jun (2004) on the other hand to be contrasted. As outlined in the introduction, Articulatory Phonology emphasizes the role of increasing overlap in leading to perceptual assimilations, and, while acknowledging the possibility of gestural reduction of $C_{1,}$ certainly does not predict expansion of C_2. Jun, by contrast, has argued that overlap on its own is an insufficient mechanism, contending that the key mechanism is rather "gestural reduction of the target consonant, with concomitant temporal expansion of the trigger consonant" (Jun, 2004, p. 70). In short, if a link of this kind could be shown it would be strong evidence against a gestural account of assimilation (for discussion of the implications of temporal patterns in models of assimilatory processes, see also Nolan, 1992).

Evidence for a link between tt spatial reduction and C_2 temporal expansion was assessed quantitatively using the following procedure: For each target sequence (and for each speaker) the correlation (Pearson's r) was calculated between the tt position at the acoustically defined 25% location and the duration of the C_2 gestural nucleus. As in previous sections we concentrated first on cases in which the standard deviation of tt position exceeded 1 mm. The interesting cases are those in which the correlation was negative (lower tt position accompanied by longer C_2 duration).

Over all speakers there were 19 cases in which the standard deviation criterion was fulfilled. Within these cases two significant negative correlations (at $p<0.05$) and one positive one were found. Thus there is negligible evidence for a robust link between tt reduction and C_2 expansion.

In retrospect, we suspect that it might not be at all straightforward to demonstrate such a correlation. There may be two competing influences on C_2 duration: (1) C_2 may increase as C_1 reduces, (2) but whenever C_1 reduces, it may also be that the phonetic explicitness of the word or word pair is also reduced, which could concomitantly reduce C_2 duration as well. Hence, to provide an additional perspective on C_2 duration we looked at the specific case of "dann kann", since this consistently attracted strong tt reduction, and tested whether its C_2 duration matched that of the corresponding control sequence "lang kann".

The answer to the question could hardly have been less consistent over the four speakers. For two, the duration of the velar constriction phase for "dann kann" was shorter than for the control "lang kann". For one it was the same, and for one it was actually longer. Given that for two of the four speakers the C_2 duration of the target sequence did not approach that of the controls, temporal expansion of C_2 does not seem to be a robust effect. For the other two speakers consideration of further sequences indicated the need for caution in the interpretation. For example, the sequence "Blatt kann" typically showed

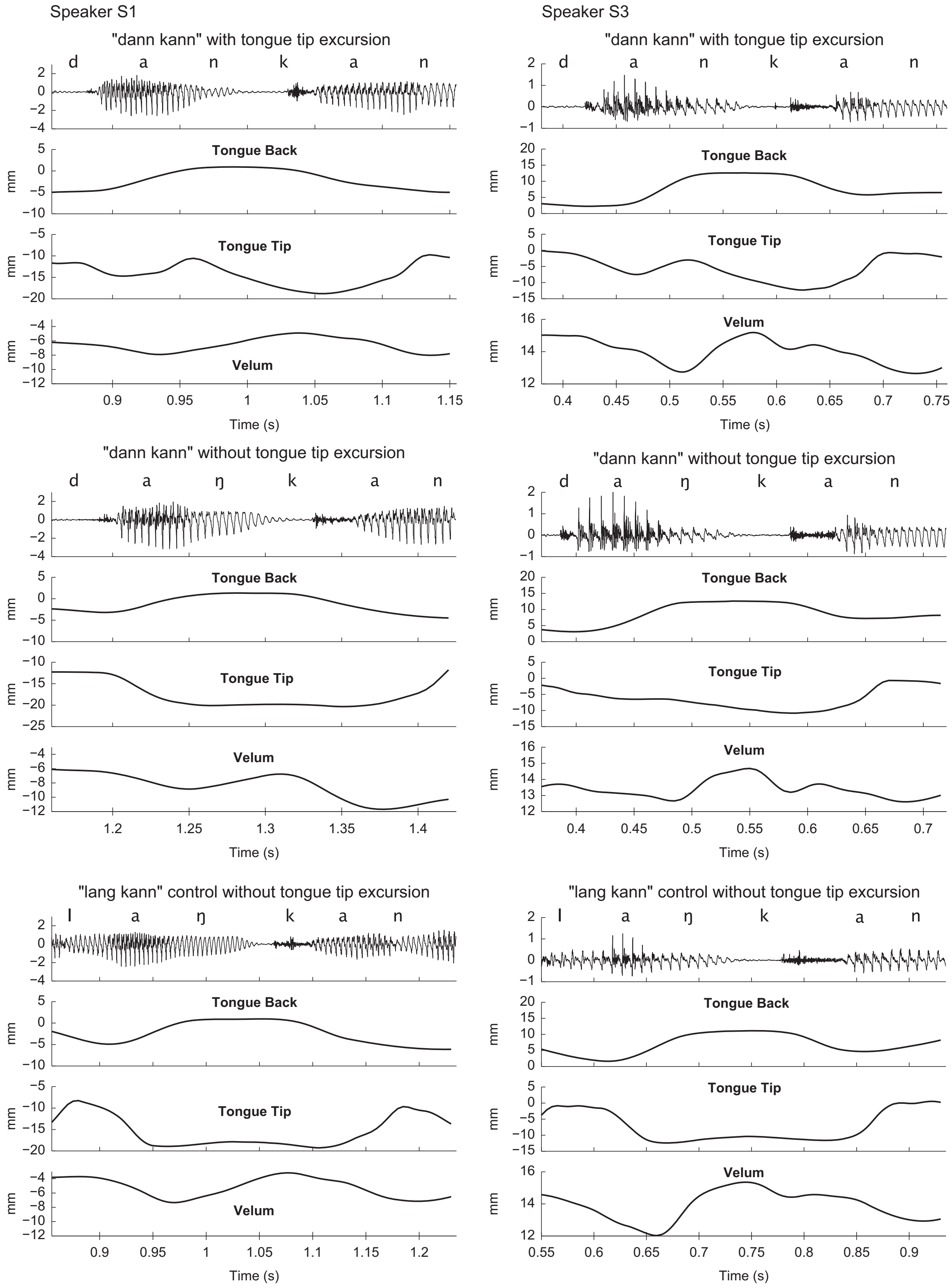

Fig. 5. Movement trajectories in target items with and without tongue-tip excursions. Audio signal (top) and movement trajectories of tongue-back, tongue-tip and velum for two speakers (left column: S1, right column: S3) in selected target and control items with and without visible tongue-tip excursion during the C_1C_2 closure phase. Top row: tt–tb target items with visible tongue-tip excursions, middle row: tt–tb target items without visible tongue-tip excursions, bottom row: tb–tb control item.

negligible tt reduction, so we would assume that temporal extension of C_2 is not expected here. Nonetheless, for only one of the four speakers was tb nucleus duration shorter in "Blatt kann" items than in the most closely matching control sequence "Tag kann". For one speaker it was about the same, and for two speakers (S2 and S3) it was actually longer. These latter two speakers were the ones who had the longest C_2 durations for "dann kann", supporting the interpretation that apparent expansion of C_2 duration in target sequences was not related to C_1 reduction.

In short, it appears there were influences on the durational properties of the geminate controls in our material that were not well controlled for, making durational comparisons with the target sequences difficult. Since we had not anticipated these problems, we hadn't included in our corpus further control items with singleton velar (tb) segment and no potential overlap. This would probably have provided a better basis for assessment of any tendency to the lengthening of C_2 in assimilation contexts.

The results do, however, fit in well with other indications in the literature that it is difficult to find evidence for a close link between reduction of the target consonant and temporal expansion of the trigger consonant in assimilation contexts. For example, Kochetov and Pouplier (2008) found in three Korean speakers that reduced tt–tb sequences had rather similar durations to the purely tb control sequences, but this certainly did not apply to the tt–ll sequences. They had a strong tendency to remain shorter than the controls. Similarly, in the investigation of n#k assimilations in English of Ellis and Hardcastle (2002) four of the eight speakers consistently showed spatial patterns for /n/ that were indistinguishable from the velar nasal controls. But of these four only two showed matching tb durations for the assimilation and control contexts. The other two had clearly shorter tb durations in the assimilation context.

We conclude this section with an illustration of movement patterns in cases in which tb constriction durations were actually highly similar for both target and control sequences and where the tt movement in "dann kann" was either invisible, or strongly reduced and completely overlapped by the tb movement (see Fig. 5).

Coupled with the fact that the velum movement was present in all these cases, and timed similarly relative to the tb movement, we clearly have in the first part of the tb constriction phase of the "dann kann" cases a configuration that corresponds phonetically to a velar nasal [ŋ]. However, on the basis of the above discussion it would be premature to assume that it was really temporal expansion of the tb constriction that led to the emergence of the velar nasal.

4. Discussion

The principal contribution of the analyses in this paper has been to give a better understanding of the articulatory substrate of the nasal asymmetry in assimilatory processes. Both stronger spatial reduction of the target as well as stronger overlap of the target and trigger consonant was found for nasal compared to oral targets.

This finding was by no means a foregone conclusion. Since observations of the manner of articulation asymmetry have been mostly based on auditory studies, it was at least conceivable that nasals and orals do not differ in their spatial or temporal properties, but that listeners simply have a stronger tendency to hear assimilated forms in nasal contexts because of less clear cues to place of articulation in nasals. This account, essentially a purely passive perceptual effect on the part of the listener, is clearly not tenable. The account we believe is most consistent with the data makes the same assumption of weaker information about place of articulation in nasals, but links it with an active process of reorganization by the speaker. We thus essentially follow the original proposal of Steriade (2001) that speakers exploit a situation in which they have the freedom to reorganize and simplify articulation without this being unduly salient to the listener (for an account along similar lines see Kohler, 1990). We are, however, now in a position to be more specific about this articulatory freedom.

The most obvious expression of this freedom to simplify articulation is the reduction in the magnitude of the tongue-tip movement. But increased overlap can also be seen in this light because it simplifies the speaker's task to articulate many segments in a given space of time without requiring physiologically costly high accelerations. Moreover, we believe that the absence of correlations between spatial reduction and temporal overlap fits in with this interpretation in terms of "articulatory freedom". Both processes are important components of fluent speech production, but can apparently be employed independently of each other. In other words, it is probably not possible to predict in detail which process will be emphasized by which speaker in which context.

It is also worth pointing out here that perhaps the most basic assumption of an account in which the acoustic properties of nasals lead to articulatory freedom for the speaker is clearly met. Trivial as it might seem, it is obviously crucial that articulatory reduction of the nasal sound should not extend to the velum opening movement itself. The elimination of velum opening would presumably make changes in the tongue-tip articulation acoustically very salient to the listener. Thus any evidence for global weakening of the gestures associated with the nasal segment would require the development of a different account from the one being put forward here. In fact, although some weakening of the velum opening movement seemed occasionally to occur, there was no evidence that the frequently encountered complete reductions of the tongue-tip movement could be accompanied by complete absence of velum opening. No cases were detected in the acoustic record of missing nasal segments, and in Section 3.5 we presented cases where velum movement was indistinguishable from the control context despite complete reduction of the tongue-tip movement.

In addition to reductions in gestural magnitude and increase in gestural overlap a third process often associated with assimilation is temporal extension of the trigger consonant. There was much less evidence for the relevance of this process, compared to the other two. This is of interest in the light of the proposed interpretation, and actually fits in well with it, because it is much less obvious how extension of C_2 could be seen as reflecting the speaker's exploitation of a situation in which simplification of articulation is perceptually tolerable. While some deficiencies in the construction of our control items do not allow final resolution of this issue, the unclear role of the extension of C_2 is supported in the literature. Moreover, the correlation analyses of C_2 duration with tongue-tip position are unaffected by the control items, and did not support the close link predicted by Jun's account (2004).

In concluding this line of argument, we need to concede that one foundation on which the above interpretation rests cannot be conclusively demonstrated here. This refers to the initial assumption that information on place of articulation is less salient in nasal contexts. Since – as also pointed out in the introduction – this contention has not been universally accepted, it remains important to substantiate it further, even if the general pattern of results indicates its plausibility. This was beyond the scope of the present investigation because, so we believe, completely different experimental techniques are called for. In particular, it could be interesting to use articulatory synthesis to generate stimuli for perception experiments in which degree of reduction of C_1 and degree of overlap of C_1 and C_2 are systematically manipulated.

Table A1
Number of trials with non-measurable tongue-tip excursions.

V_1_pal	w_1_freq	C_1_MOA	C_2_POA	$V_1C_1\#C_2V_2$	Word pair	*Speaker*			
						S1	S2	S3	S4
non-palatal	low	nasal	velar	an#ka	Zahn kann				1
	high	nasal	velar	an#ka	dann kann	8	5	8	
	high	nasal	labial	an#pa	dann passt	1	2		
	high	oral	velar	at#ko	statt kommende			2	
palatal	high	nasal	velar	en#ka	zehn Kappen	1		5	7
	high	oral	labial	it#pa	sieht Padua	3		4	7
	high	oral	velar	it#ko	sieht kommende	3		1	1
	low	oral	labial	it#pa	Lied passt	2	1		1

The main prediction would be that listeners become unsure in their categorization of place of articulation at a lesser degree of reduction for nasals than for oral consonants.

Concerning the remaining independent variables the present results did not confirm the hypothesis derived from the earlier study of Kühnert and Hoole (2004) that more evidence of assimilation would be found in the palatal vowel context.

With regard to the influence of place of articulation of C_2 the main point of interest was whether conflicting requirements on the tongue for C_1 and C_2 (in the velar C_2 context) could lead to stronger reduction of the tongue-tip gesture than in the non-conflicting context (C_2 labial) (Zsiga, 1994). It is difficult to give a firm answer to this question. In the non-palatal vowel context there was certainly overall more reduction for velar than labial C_2. But as repeatedly emphasized, this was mainly due to the high-frequency nasal item "dann kann". So there is not much evidence for a generalization of the effects to low frequency and oral items. In the palatal vowel context the effects were even more mixed. Only two speakers showed a robust effect of more reduction for velar C_2, and one speaker showed a small but significant effect in the opposite direction. In short, on the basis of the present results it would be hazardous to conclude that there is a general effect of velar place of articulation per se beyond the effect probably due to the special frequency properties of the "dann kann" item, to which we now turn as the final element in the discussion.

We have chosen to consider frequency effects last of all because this probably represents the area that would be most interesting to follow up in future work. We based the discussion above on the contention that it is speakers' knowledge of acoustic–perceptual properties that allows them to simplify articulation in nasal–stop sequences (Honorof, 2003; Steriade, 2001). However, the inclination to make use of this knowledge may depend in large part on word frequency and lexical probability. In fact, the influence of nasality and frequency may show mutual reinforcement, because it can be plausibly assumed that it forms part of speakers' knowledge of the sound structure of German that there is a particular concentration of high-frequency words that end in a nasal (Jaeger, 2003). However, this is precisely where further investigation would be required. There were pervasive effects of frequency throughout our results and although they were not concentrated exclusively on the item with high co-occurrence frequency, it is tempting to suspect that it is co-occurrence rather than word frequency that is the crucial factor (see Bybee, 2003, p. 157), who assumes that words that are often used together become processing units and changes within these units arise as the result of automating production). But tempting as it may be, it is equally clear that it would be premature on the basis of the present results to assume that this is actually the case. To test this more rigorously it would be necessary to look at word pairs with a greater range of word and co-occurrence frequencies. This was beyond the scope of the present experiment because of the complexity introduced by our original aim of including vowel and C_2 context in the balanced factorial design, and so must remain as probably the most interesting avenue to explore in future work.

Acknowledgments

Work supported by German Research Foundation grant HO 3272/1 to Philip Hoole. The authors are grateful to Hélène Loevenbruck and one anonymous reviewer for very helpful comments on an earlier version.

Appendix A

See Table A1.

References

Barry, M. C. (1991). Temporal modeling of gestures in articulatory assimilation. In *Proceedings of the 12th international congress of phonetic sciences* (pp. 14–17). Aix-en-Provence, France.

Bergmann, P. Articulatory reduction and assimilation in n#g sequences in complex words in German. In P. Hoole, L. Bombien, M. Pouplier, C. Mooshammer, & B. Kühnert (Eds.), *Clusters and structural complexity*. Berlin: de Gruyter; accepted for publication.

Browman, C. P., & Goldstein, L. M. (1992). Articulatory phonology: An overview. *Phonetica*, *49*, 155–180.

Browman, C. P., & Goldstein, L. M. (1995). Gestural syllable position effects in American English. In F. Bell-Berti, & L. J. Raphael (Eds.), *Producing speech: Contemporary issues* (pp. 19–32). New York: American Institute of Physics.

Bybee, J. (2003). *Phonology and language use*. Cambridge: Cambridge University Press.

Bybee, J., & Scheibman, J. (1999). The effect of usage on degree of constituency: The reduction of *don't* in American English. *Linguistics*, *37*, 575–596.

Byrd, D. (1996). Influences on articulatory timing in consonant sequences. *Journal of Phonetics*, *24*(2), 209–244.

Chitoran, I., Goldstein, L., & Byrd, D. (2002). Gestural overlap and recoverability: Articulatory evidence from Georgian. In C. Gussenhoven., T. Rietveld, & N. Warner (Eds.), *Papers in laboratory phonology*, Vol. 7 (pp. 1–39). Cambridge: Cambridge University Press.

Ellis, L., & Hardcastle, W. J. (2002). Categorical and gradient properties of assimilation in alveolar to velar sequences: Evidence from EPG and EMA data. *Journal of Phonetics*, *30*, 373–396.

Gregory, M. L., Raymond, W. D., Bell, A., Fosler-Lussier, E., & Jurafsky, D. (1999). The effect of collocational strength and contextual predictability in lexical production. *Proceedings of the Chicago Linguistic Society*, *35*, 151–166.

Hardcastle, W. J. (1995). Assimilation of alveolar stops and nasals in connected speech. In J. W. Lewis (Ed.), *Studies in general and English phonetics* (pp. 49–67). New York: Routledge.

Honorof, D. (2003). Articulatory evidence for nasal de-occulusivization in Castilian. In *Proceedings of the 15th international congress of phonetic sciences* (pp. 1759–1762). Barcelona, Spain.

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. 331–349). Oxford: Oxford University Press.

Hoole, P., Zierdt, A., Geng, C. (2003). Beyond 2D in articulatory data acquisition and analysis. In *Proceedings of the 15th international congress of phonetic sciences* (pp. 265–268). Barcelona, Spain.

Hura, S. L., Lindblom, B., & Diehl, R. L. (1992). On the role of perception in shaping phonological assimilation rules. *Language and Speech*, *35*, 59–72.

Institut für Deutsche Sprache/IDS-Korpus 〈http://www.ids-mannheim.de〉.

Jaeger, M. (2003). Perception of lost contrast. In *Proceedings of the 15th international conference of phonetic sciences* (pp. 1735–1738). Barcelona, Spain.

Jun, J. (2004). Place assimilation. In B. Hayes, R. Kirchner, & D. Steriade (Eds.), *Phonetically based phonology* (pp. 58–86). Cambridge: Cambridge University Press.

Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In J. Bybee, & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 254–299). Amsterdam: John Benjamins Publishing Company.

Kochetov, A., & Pouplier, M. (2008). Phonetic variability and grammatical knowledge: An articulatory study of Korean place assimilation. *Phonology*, *25*, 399–431.

Kohler, K. (1990). Segmental reduction in connected speech in German: Phonological facts and phonetic explanations. In W. Hardcastle, & A. Marchal (Eds.), *Speech production and speech modeling* (pp. 69–92). Dordrecht: Kluwer.

Krakow, R. A. (1999). Physiological organization of syllables: A review. *Journal of Phonetics*, *27*, 23–54.

Kühnert, B., & Hoole, P. (2004). Speaker-specific kinematic properties of alveolar reductions in English and German. *Clinicial Linguistics & Phonetics*, *18*, 559–575.

Mücke, D., Grice, M., & Kirst, R. (2008). Prosodic and lexical effects on German place assimilation. In *Proceedings of the eighth international seminar on speech production* (pp. 225–228). Strasbourg, 2008.

Nolan, F. (1992). The descriptive role of segments: Evidence from assimilation. In D. Ladd., & G. Docherty (Eds.), *Papers of laboratory phonology, II: Gesture, segment, prosody* (pp. 261–280). Cambridge: Cambridge University Press.

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

Stephenson, L. (2003). An EPG study of repetition and lexical frequency effects in alveolar to velar assimilation. In *Proceedings of the 15th international conference of phonetic sciences* (pp. 1891–1894). Barcelona, Spain.

Steriade, D. (2001). Directional asymmetries in place assimilation. In E. Hume, & Johnson (Eds.), *The role of speech perception in phonology* (pp. 219–250). London: Academic Press.

Winters, S. J. (2001). VCCV perception: Putting place in place. *Ohio State University Working Papers in Linguistics*, *55*, 70–87.

Wright, C., & Kerswill, P. (1989). Electropalatography in the analysis of connected speech processes. *Clinical Linguistics & Phonetics*, *3*, 49–57.

Zierdt, A., Hoole, P., & Tillmann, H. G. (1999). Development of a system for three-dimensional fleshpoint measurement of speech movements. In *Proceedings of the 14th international congress of phonetics sciences*. San Francisco.

Zsiga, E. C. (1994). Acoustic evidence for gestural overlap in consonant sequences. *Journal of Phonetics*, *22*, 121–144.