From canonical word forms to reduced variants: A study of assimilation and elision in German.

Bernd J. Kröger
Institut für Phonetik der Universität zu Köln

Abstract

In the framework of Articulatory Phonology (Browman and Goldstein 1992) the variety of discrete segmental changes describing the transition from canonical word forms to reduced variants (i.e. elision and assimilation phenomena) can be accounted for by two continuous and non-discrete gestural alteration processes: increase in overlap and decrease in temporal extent of articulatory gestures.
It can be shown that many segmental phenomena like elisions and assimilations in German can be ascribed to these two basic gestural alteration processes. But some assimilation phenomena (progressive and regressive assimilation of place and regressive assimilation of manner) can be described only by introducing a discrete gestural process: gestural (or articulatory) reorganization.
Further we will show that both continuous gestural processes are strongly related to each other. Increase in overlap can be attributed to reduction of temporal extent of gestures if basic gestural association relations are taken into account. In order to develop a comprehensive theory of reduction, we will illustrate that all continuous and discrete gestural processes can be seen as consequences of minimizing articulatory effort.

1 A brief introduction to Articulatory Phonology

1.1 The gesture as a phonetic and phonological unit

Gestures are the basic units of Articulatory Phonology (Browman and Goldstein 1992). They are units of articulatory activity, realizing linguistically relevant vocal tract constrictions like "labial closure" or "glottal opening". Consequently gestures are phonetic as well as phonological units. On one hand gestures are distinctive units and define discrete phonological categories like [place], e.g. labial vs. apical gestures, [manner], e.g. full-closing gestures (for plosives or nasals) vs. near-closing gestures (for fricatives), or [voice], e.g. occurrence vs. no occurrence of a glottal or velic opening gestures. On the other hand each gesture represents a family of functionally equivalent articulatory movement patterns that are actively controlled with reference to speech relevant goals, i.e. the formation of vocal tract constrictions (Saltzman and Munhall 1989). Consequently in the framework of Articulatory Phonology we have no separation of phonological and phonetic units as occurring in segmental theories (e.g. the separation between phoneme and sound).
There are a lot of reasons, which illustrate the importance of the gesture. Firstly - as illustrated above - the gesture is a phonological as well as a phonetic unit. Consequently there is no need to define a phonetic-phonological interface in this approach. The concept "gesture" can be used both in phonetic *and* phonological investigations. In a quantitative model of speech pro-

duction (chapter 2) a phonological description of a gesture can be transformed into an equivalent phonetic one by specifying values for continuous phonetic parameters like target location, strength of gestural activation, and duration of gestural activation (chapter 2.1). Secondly, articulatory measurements indicate that the spatio-temporal structure of articulatory transitions - i.e. the articulatory shape of gestures - is more stable than the spatio-temporal structure in the region of articulatory targets (i.e. around the maxima and minima of articulatory trajectories) (Fujimura 1981 and 1986). Thirdly, a variety of different discrete segmental changes (e.g. elisions and assimilations) occurring in reduced formes (e.g. words in unstressed positions, at high speech rate, or in casual speech) can be ascribed to few continuous gestural processes: increase in overlap of two gestures and decrease of temporal extent of a gesture (Browman and Goldstein 1989 and 1990, Kröger 1993, and this paper, chapter 3). No discrete gestural change - especially no deletion of gestures - occurs in the case of reduction. This is promising since the degree of reduction can result from varying paralinguistic parameters like speech rate. And a variation of a continuous paralinguistic parameter should not lead to a discrete change of linguistic units. Fourthly, the gesture can be seen as a unit of speech production as well as of speech perception. The motor theory of speech perception (Liberman and Mattingly 1985) defines the gesture as its central unit.

1.2 The gestural organisation of a word

In order to understand the gestural approach it is important to understand how an utterance (or at least a word) is organised in the gestural concept. Figure 1 indicates the phonological specification (i.e. the gestural score) of the German word "Kompaß" (compass) in the framework of Articulatory Phonology. Three types of gestures must be differentiated: tract-shaping gestures (TSG), constriction-forming gestures (CFG), and opening gestures (OPG). Gestures can be phonologically specified by four-letter-symbols: {oldl} or {aldl} are dorsal-labial gestures for the realization of the German lax /ɔ/ or lax /a/, {fcla} or {fcdo} are labial or dorsal full-closing gestures, {ncal} is a alveolar near-closing gesture, and {opgl} or {opve} are velic or glottal opening gestures ("velic" is chosen as a term for the active articulator velum while the term "velar" indicates a (passive) place of articulation in our approach). All gestures can be ordered in different gestural tiers as function of their type. Association lines indicate which gesture is timed or phased with respect to which other gesture.

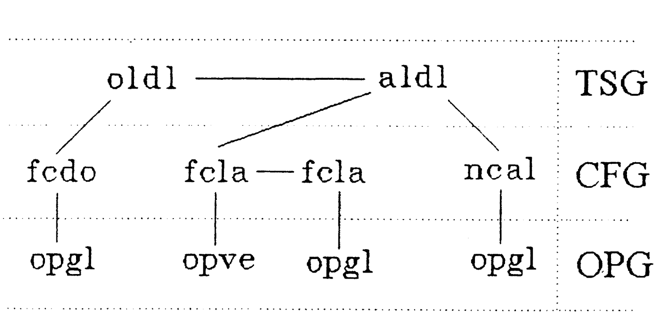


Figure 1 The phonological specification of /kɔmpas/.

2 A gestural speech production model

In the gestural approach any phonological specification can be realized phonetically. A phonetic speech production model has been developed in order to generate the articulation and the acoustic speech signal for a given discrete gestural specification. The first step is the transformation of the phonological gestural specification (four-letter-abbreviations, fig. 1) into a phonetic specification by specifying the values of all (phonetic) parameters for all gestures of an utterance.

2.1 Gestural parameters

Each gesture is defined phonetically (1) by the articulator(s) executing the gestural movement, by target location(s) indicating the gestural target shape(s) or location(s) which is (are) approximated by the gesture-executing articulators(s), (2) by the temporal location and duration of the gestural activation interval, i.e. of the time interval in which the gesture is actively controlling the articulator(s), and (3) by the strength of gestural activation. The gestural parameter "gesture-executing articulator" can be taken directly from the phonological specification of the gesture (chapter 1.2). The parameter "gestural target" is quantitatively defined by specifying values for control parameters like lip protrusion, tongue position, or glottal aperture. These control parameters and its range are model-specific (e.g. Kröger 1993). All other gestural parameters - i.e. the parameter "associated gesture" (e.g. {fcdo} for {oldl} or {opgl} for {fcdo} in "Kompaß", fig. 1) and the three continuous gestural parameters "eigenperiod", "release phase", and "association phase" - specify the strength, temporal location, and temporal extent of the gestural activation interval. Eigenperiod determines the strength of gestural activation; Eigenperiod together with release phase determines the length of gestural activation (Kröger 1993, Kröger et al 1995 and chapter 2.2); Association phase determines the temporal location of the gestural activation interval relative to the location of the associated gesture.

Figure 2 gives the temporal location and the extent of gestural activation intervals for "Kompaß". The gestures are ordered here in five articulatory tiers according to the gesture-performing articulators, i.e. tongue body (TB), tongue tip (TT), lips (LI), velum (VE) and glottis (GL). This figure illustrates two main conditions for an articulator: (1) If gestural activation occurs the articulator is controlled by a gesture. In this case the articulator performs a movement towards the gestural target. (2) If no gestural activation occurs for an articulator this articulator performs a movement towards its inherent neutral position. The neutral position of all articulators defines the production state of a voiced non-nasalized schwa-sound.

Furthermore, this figure shows that gestures overlap in time. Especially different types of gestures overlap considerably in time: Tract forming gestures always overlap with constriction-forming gestures, and constriction-forming gestures always overlap with opening gestures. But also constriction-forming gestures and opening gestures can overlap in time.

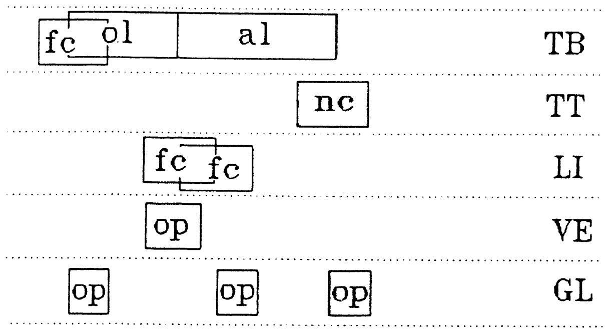


Figure 2 The temporal location and extent of the activation intervals for all gestures of "Kompaß". The abscissa represents time. Each box marks the beginning and ending of a gestural activation interval.

Phonetically the gesture can be seen as a unit of articulatory control. If all gestural parameters are specified, a gesture leads to a defined articulatory movement. Consequently, the specification of all gestures of an utterance leads to an explicit description of its articulation. A vocal tract model can be driven by the gestural specification which generates vocal tract shapes as function of time and subsequently the acoustic speech signal of the utterance. Figure 3 indicates the articulation and the acoustic speech signal for the word "Kompaß" realized in our production model (Kröger 1993). Control parameters and their values are defined with respect to this production model. The control parameters in figure 3 are tongue height (TH), tongue position (TP), tongue tip height (TTH) lip aperture (LA), velic aperture (VA), and glottal aperture (GA).

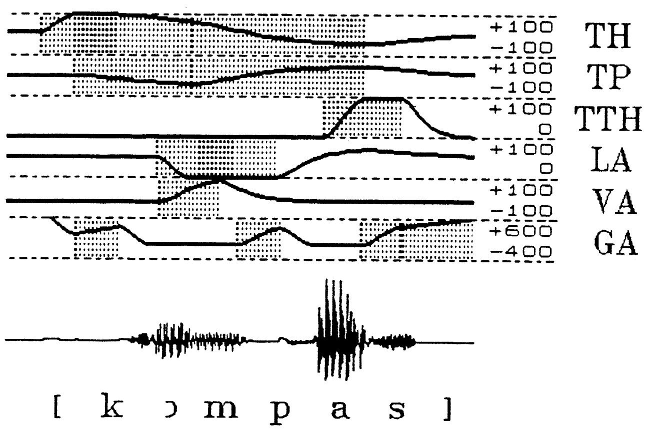


Figure 3 Control parameter time functions (thick lines), gestural activation intervals (shaded areas), and the oscillogram of the synthetic audio signal for "Kompaß". The last glottal opening gesture is followed by a postphonatory opening gesture.

2.2 The dynamic concept for gestures

The dynamics of each gesture determines the spatio-temporal shape of a gesture (i.e. the gestural time function or the gestural movement pattern) and can be generated by a critically damped harmonic oscillator (Saltzman and Munhall 1989, Browman and Goldstein 1990, Kröger 1993, Kröger et al. 1995). In this case the gestural articulator movement is the pattern of an exponential time function asymptotically descending to zero-displacement, i.e. to the gestural target. Examples for gestural movement patterns are shown in figure 4. Here the abscissa indicates relative time values (i.e. phase values, see below) and the ordinate indicates the articulator-target displacement relative to initial displacement. 0% relative articulator-target displacement indicates the target location and 100% indicates initial displacement.

One important parameter of a harmonic oscillator is eigenperiod (i.e. the reciprocal of eigenfrequency) which indicates the level of activation of the oscillator. Low eigenperiod (high eigenfrequency) indicates high activation and vice versa. In the case of critical damping the degree of activation (i.e. eigenperiod) defines the time interval needed for reaching a definite (small) relative articulator-target distance. A gesture with a low eigenperiod value reaches a defined (small) articulator-target distance faster than a gesture with a high eigenperiod value (fig. 4). A relative time scale - the phase scale - can be defined for each gesture if the strength of gestural activation is known: According to the eigenperiod of the gesture the distances on the phase scale are large for low and small for high eigenperiod. In figure 4 the phase scales are indicated for both gestures: above for the dashed lined gesture and below for the solid lined gesture. The figure shows that phase values depend solely on the relative articulator-target-distance: *Phase values indicate the degree to which a gesture is performed.*

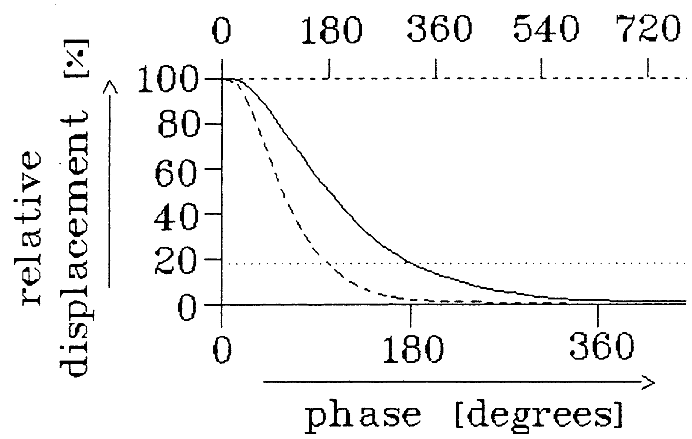


Figure 4 Gestural time functions. The eigenperiod value is lower in the dashed lined than in the solid lined time function. The steady state portion of the gesture occurs below and the transient portion of the gesture above the dotted horizontal line.

As a rule of thumb it can assumed that the gestural target region - i.e. the quasi steady state portion of a gesture - is reached for each gesture at about 180 degrees. For (consonantal) constriction-forming gestures this phase value indicates the beginning of the consonantal obstruction (e.g. full or near closure). Thus the rapid articulator movement towards the gestural target, i.e. the *transient portion* of a gesture, takes place at phase values below 180 degrees whereas the *quasi steady state portion* in which the articulator is near the gestural target (relative articulator-target distance is lower than approximately 20%, see fig. 4) appears at phase values above 180 degrees.

The remaining gestural parameters are defined with respect to the gestural phase scale. The release phase value indicates the final degree of realization of a gesture, i.e. the degree of articulatory undershoot of a gesture. If for example the release phase value of a consonantal gesture is lower than 180 degrees no consonantal closure will be produced. And the longer a gesture is activated above 180 degrees the more quasi steady state portion of the gesture (e.g. the more time portion of a consonantal closure, of a quasi steady (vocalic) vocal tract shape, or of a glottal or velic opening) is produced. Together with eigenperiod the release phase value determines the temporal extent of a gestural activation interval.

The association phase value of a gesture determines the position of the phase scale of this gesture with respect to the time instant defined by the association line (fig. 1). If for instance the association phase value is 0 degrees, the gesture starts at the time instant defined by the association line; if the association phase value is 180 degrees for a (consonantal) constriction-forming gesture, this gesture is timed (or phased) with respect to the beginning of its consonantal obstruction.

2.3 Gestural phasing rules

The association phase value determines the temporal location of a gesture (i.e. the position of its phase scale) with respect to the time instant defined by the association line. In order to determine gestural phasing completely, association rules are added defining which gesture has to be phased with respect to which other gesture (Browman and Goldstein 1990). Consequently, these rules define the time instants of phasing, i.e. the time instants represented by association lines. Four association rules can be established: (1) Each vocalic gesture is phased with respect to the offset of the preceding vocalic gesture (horizontal phasing line in the tract-shaping tier in fig. 1). (2) The first consonantal gesture of a consonant cluster is phased with respect to the onset of the syllable-defining vocalic gesture if the cluster is syllable-initial, and with respect to its offset if the cluster is syllable-final (transversal phasing lines from the tract-shaping to the constriction-forming tier in fig. 1). (3) Non-first consonantal gestures of a consonant cluster are phased with respect to the offset of the preceding consonantal gesture within this cluster (horizontal phasing line in the constriction-forming tier in fig. 1). (4) Opening gestures are phased with respect to the offset of the pertinent consonantal gesture (vertical phasing lines from the constriction-forming tier to the tier of opening gestures in fig. 1).

The association phase values of each gesture, together with these association rules, determine the intergestural constellation completely. The rules given above together with the specification of association phase values lead to four main principles for gestural coordination: (1) Vocalic gestures are in an immediate succesion without gaps (rule 1 and association phase value of zero). They act as a "ground" to consonantal "figures" (Browman 1991). The articulatory movements resulting from these series of vocalic gestures are comparable to the "vocalic base function" in Fujimura's C/D model (Fujimura 1992) or to the "vowel component" in Öhman's model (Öhman 1967). Consequently, consonantal gestures are completely overlapped by vocalic gestures. (2) The consonantal obstruction interval of a consonant (cluster) coincides with the transient portion of a vowel gesture (rule 2 and association phase value of 180 degrees). Thus the vocalic transition portions of the tongue body are hidden by consonantal constrictions. (3) Consonantal obstructions within a consonant cluster are produced without gaps (rule 3 and association phase value of 180 degrees) as such gaps would be perceived as interconsonantal vocalic segments. (4) The temporal extent of the activation interval of an opening gesture coincides with the consonantal obstruction interval (rule 4 and association phase value of 180 degrees). This rule ensures the proper "intrasegmental" timing in the case of voiceless sounds or nasals. In our elaborated quantitative model this rule is differentiated for plosives, fricatives, and nasals according to the articulatory measurements of Löfqvist and Yoshioka (1984):

2.4. Comparison of segmental and gestural approach

Firstly the gestural approach can be interpreted as a non-segmental concept. Gestures are the central units of phonological as well as phonetic description. While segments are serially ordered in segmental approaches gestures are ordered on parallel gestural or articulatory tiers (fig. 1 and fig. 2). Furthermore it is an important feature of this approach that gestures overlap in time. This overlap of gestures in time and the non-serial ordering of gestures is an important feature of the gestural approach. Consequently Articulatory Phonology belongs to non-linear and non-segmental phonologies.
Secondly, it should be mentioned that the concept of "coarticulation" is defined in different ways in segmental theories. The gestural approach allows the replacement of the concept "coarticulation" by the concept of "gestural coproduction". Temporal overlap is clearly defined in this qualitative and quantitative concept and consequently gestural overlap can be concretely measured in this approach.
Thirdly, it is an advantage of the gestural approach that a discrete phonological specification can be transformed into a concrete phonetic realization by means of our gestural production model as is illustrated here for the example "Kompaß" (fig. 1, 2, and 3). Articulator movements and the acoustic signal can be generated from a phonological specification in Articulatory Phonology. This provides us the possibility of perceptual evaluation of gestural specifications and also of gestural processes as has been taken advantage of in this study (see chapter 3).

3 Assimilations and elisions in the gestural framework

3.1 Assimilations and elisions in German and gestural processes

As a consequence of reduction in connected speech a lot of segmental changes - mainly assimilations and elisions - can be found. The main hypothesis of Articulatory Phonology is that these different kinds of *discrete segmental changes* can be ascribed to few simple *continuous gestural alteration processes*, i.e. increase in temporal overlap of gestures and decrease of the temporal extent of a gesture (Browman and Goldstein 1990). It is important to emphasize that these discrete segmental changes can be realized without deletions of any gesture.
In our investigation we verified this hypothesis in the case of assimilations (e.g. assimilation of place, manner, nasality, or voice) and elisions (e.g. elision of [ə]or of [t]) occurring in German (Kohler 1995, p. 205ff).

The cases investigated by using our gestural production model are listed in table 1. It must be emphasised, that in some cases the forms on the left side of table 1 (starting forms) are reduced forms (e.g. in the case of reduction of double consonants in "kommen" [mm]->[m]: the reduction of double consonants is preceded by elision of [ə] and progressive assimilation of nasality. So the starting form is [kɔmm̩]).

Category of segmental change	examples investigated by using the gestural production model
Elision of [ə]	"eben" [bən]->[bn], "reden" [dən]->[dn], "legen" [gən]->[gn], "Adel" [dəl]->[dl]
Elision of [t]	"Glanz" [nts]->[ns], "erhältst" [lts]->[ls], "restlich" [stl]->[sl], "rechtlich" [çtl]->[çl]
Reduction of double consonants	"mitteilen" [tt]->[t], "komm(e)n" [mm]->[m], "wegkommen" [kk]->[k]
Progressive assimilation of place	"eb(e)n" [bn]->[bm], "komm(e)n" [mn]->[mm], "Beamt(e)n" [mtn]->[mpm], "verlog(e)n" [gn]->[gŋ]
Regressive assimilation of place	"mit mein(e)m" [nm]->[mm], "mit jed(e)m" [dm]->[bm], "mit fett(e)m" [tm]->[pm]
Regressive assimilation of manner	"das Schiff" [sʃ]->[ʃʃ], "Eisschrank" [sʃ]->[ʃʃ]
Progressive assimilation of nasality	"umbenennen" [mb]->[mm], "Bundes" [nd]->[nn], "angegeben" [ŋg]->[ŋŋ]
Regressive assimilation of nasality	"Agnes" [gŋ]->[ŋŋ], "eb(e)n" [bm]->[mm], "wird(e)n" [dn]->[nn]
Progressive assimilation of voicelessness	"ratsam"[tz]->[ts], "dasselbe" [sz]->[ss], "das Bad" [sb]->[sb̥], "wegbringen" [kb]->[kb̥]
Sonorization (i.e. assimilation of voice)	"muß ich" [s]->[z], "hat er" [t^h]->[d]
Reduction of degree of opening	"ich habe" [b] -> [β$_T$], "ich lege" [g]->[ɣ$_T$]

Table 2 Categories of segmental changes and concrete examples for each category investigated by using our gestural speech production model.

During the procedure of generation of the segmental changes we firstly generated the unreduced form (starting form) of these words. Secondly we tried to generate the reduced forms by identifying and applying the appertaining gestural alteration process for each word. Thirdly transcriptions of both acoustic forms were analysed to find out whether the segmental change has occurred.

We identified underlying gestural alteration processes in the case of 8 of the given 11 categories of segmental changes, i.e. (1) in the case of elision of [ə] and (2) of [t], (3) in the case of reduction of double consonants, (4) in the case of regressive and (5) progressive assimilation of nasality, (6) in the case of progressive assimilation of voicelessness, (7) in the case of sonorization and (8) in the case of reduction of degree of opening. The appertaining gestural alteration processes are indicated in figure 5 for one example for each category of segmental change. The figure shows the gestural score before (left side) and after (right side) each segmental change. The altered gestures are indicated by arrows: an arrow behind the gesture indicates a decrease of temporal extent of the gesture; An arrow before the gesture indicates a shift of the gesture to the left, i.e. an increase in overlap with other gestures. The horizontal extension of the boxes represents the time interval of gestural activation. Association lines are indicated by vertical dotted (case syllable boundary) or dashed (all other cases) lines. Time intervals of consonantal obstructions (closures) are indicated by shaded areas within the boxes of gestural activation. Time intervals of gottal or velic closing movements following an opening gesture are indicated by dashed lined boxes if necessary.

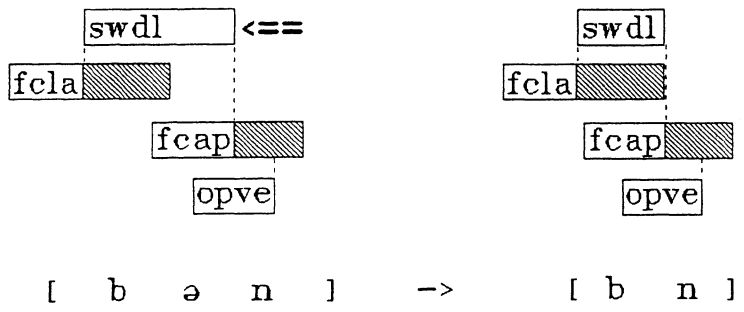


Figure 5a Gestural alteration process for elision of [ə] in "eben".

Elision of [ə] in "eben" ([bən]->[bn]) is reached by a decrease in temporal extent of the dorsal labial schwa-forming gesture {swdl} (fig. 5a). According to the gestural phasing rules this process also leads to an increase in temporal overlap of the labial and apical full-closing gestures {fcla} and {fcap}. The segmental elision occurs if the closure intervals of both closing gestures (shaded areas in both gestural activation intervals) overlap. It should be noted that segmental elision of [ə] occurs without a full reduction of the temporal extent of the schwa-gesture. The schwa-gesture still exists in the reduced form. This gesture is only hidden by the occlusions of the consonantal gestures.

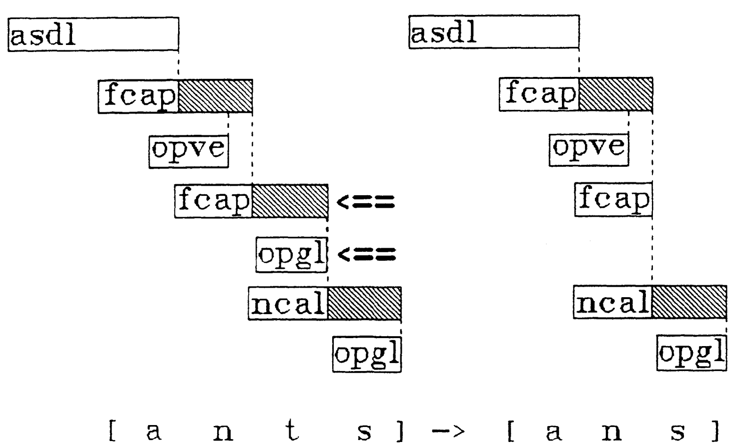


Figure 5b Gestural alteration process for elision of [t] in "Glanz".

Elision of [t] in "Glanz" ([nts]->[ns]) is reached by decrease in temporal extent of the apical full-closing gesture {fcap} and glottal opening gesture {opgl} (fig. 5b). The glottal opening gesture may remain unreduced if the reduction of temporal extent of the apical full-closing gesture does not lead to a shift of the temporal location of the glottal opening gesture.

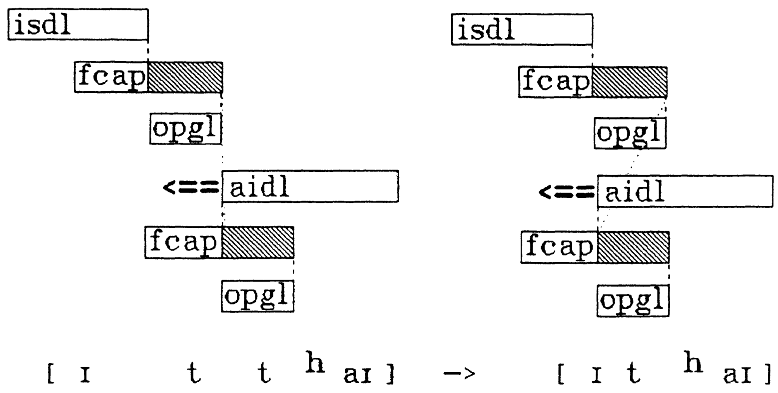


Figure 5c Gestural alteration process for reduction of double consonants in "mitteilen".

Reduction of double consonants in "mitteilen" ([tt]->[t]) is reached by increase in temporal overlap of the first and the second syllable, i.e. by increase in overlap of the vocalic gestures of the first and second syllable: the dorsal-labial short /i/-forming gesture {isdl} and the dorsal-labial /ai/-forming gesture {aidl} (fig. 5c). According to the gestural phasing rules this leads to a complete temporal overlap of the apical full-closing gestures {fcap} and their appertaining glottal opening gestures {opgl}, i.e. a complete temporal overlap of the initial consonantal closing gestures and their appertaining opening gestures of the second syllable with the final consonantal gestures of the first syllable.

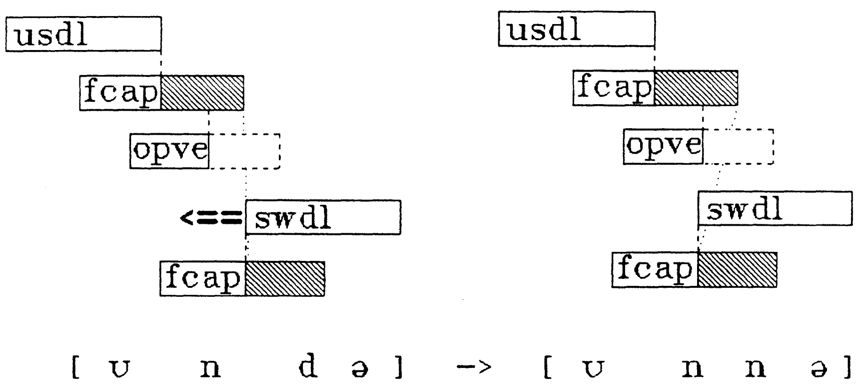


Figure 5d Gestural alteration process for progressive assimilation of nasality in "Bundes".

Progressive assimilation of nasality in "Bundes" ([nd]->[nn]) is reached by increase in temporal overlap of the first and the second syllable, i.e. by increase in overlap of the vocalic gestures of the first and second syllable: the dorsal-labial short /u/-forming gesture {usdl} and the dorsal-labial schwa-forming gesture {swdl} (fig. 5d). According to the gestural phasing rules this leads to a complete temporal overlap of the apical full-closing gestures {fcap}, i.e. a complete temporal overlap of the initial consonantal closing gesture of the second syllable with the final consonantal gestures of the first syllable.

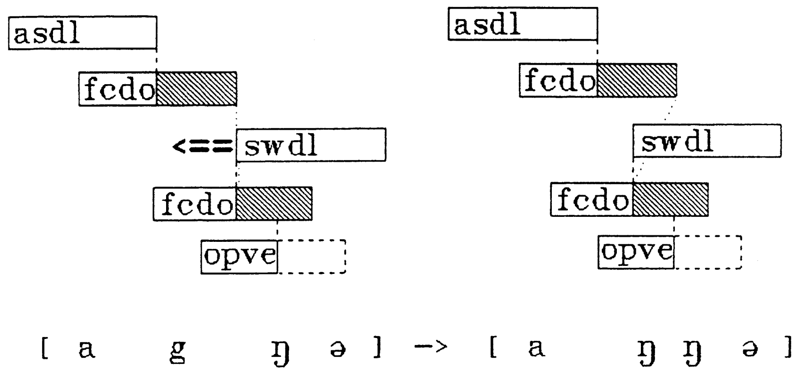


Figure 5e Gestural alteration process for regressive assimilation of nasality in "Agnes".

Regressive assimilation of nasality in "Agnes" ([gŋ]->[ŋŋ]) is reached by increase in temporal overlap of the first and the second syllable, i.e. by increase in overlap of the vocalic gestures of the first and second syllable: the dorsal-labial short /a/-forming gesture {asdl} and the dorsal-labial schwa-forming gesture {swdl} (fig. 5e). According to the gestural phasing rules this leads to a complete temporal overlap of the dorsal full-closing gestures {fcdo}, i.e. a complete temporal overlap of the initial consonantal closing gestures of the second syllable with the final consonantal gestures of the first syllable.

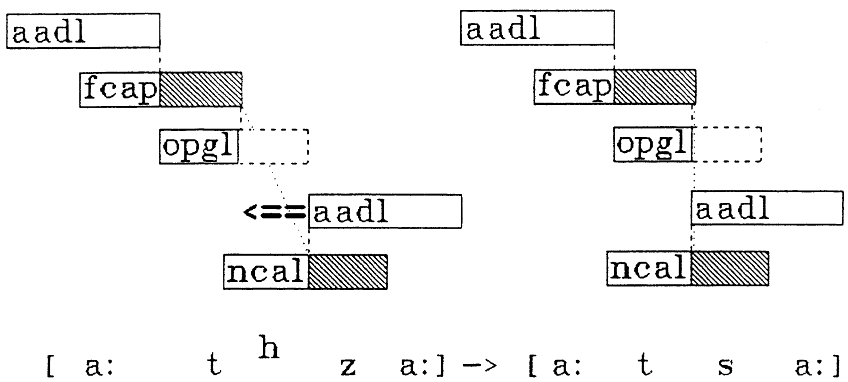


Figure 5f Gestural alteration process for progressive assimilation of voicelessness in "ratsam".

Progressive assimilation of voicelessness in “ratsam“ ([tz]->[ts]) is reached by increase in temporal overlap of the first and the second syllable, i.e. by increase in overlap of the vocalic gestures of the first and second syllable: both dorsal-labial long /a/-forming gestures {aadl} (fig. 5f). But in this case the temporal overlap is not robust as in the above cases of overlap of syllables. According to the gestural phasing rules this temporal overlap leads to a partial temporal overlap of the alveolar near-closing gesture {ncal} with the apical full-closing gesture {fcap}, i.e. a partial temporal overlap of the initial consonantal closing gestures of the second syllable with the final consonantal gestures of the first syllable.

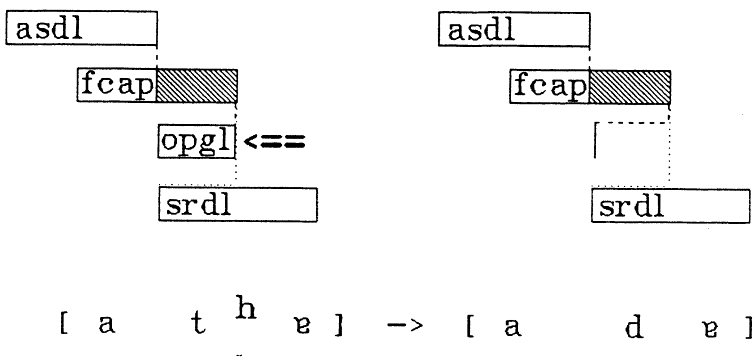


Figure 5g Gestural alteration process for sonorization in “hat er“.

Sonorization (i.e. assimilation of voice) in “hat er“ ([t^h]->[d]) is reached by decrease in temporal extent of the glottal opening gesture {opgl} (fig. 5g). This decrease need not lead to a total reduction of the glottal opening gesture. The segmental change is also reached if a rudimentary glottal opening gesture remains. If the phonological gestural concept is extended by introducing glottal closing gestures in order to ensure the occurrence of glottal vibration (i.e. of voicing) this process is equivalent to an increase in overlap of the gottal opening gesture and the glottal closing gesture of the vocalic portion of the second syllable. In the present gestural concept voicing results from glottal underspecification. But an extension of our gestural approach by introducing glottal and velic closing gestures in the case of sonorants and obstruents has been suggested (Geuman and Kröger 1995).

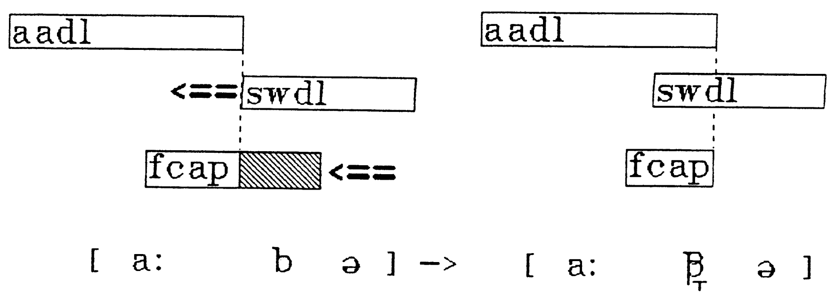


Figure 5h Gestural alteration process for reduction of degree of opening in “ich habe“.

Reduction of degree of opening in "ich ha<u>b</u>e" ([b] -> [β̞]) is reached by a decrease in temporal extent of the apical full-closing gesture {fcap} (fig. 5h). This decrease quantitatively leads to a release phase value of around 180 degrees. Thus the decrease of temporal extent leads to an omission of the consonantal closure interval. Since this makes the transition portions of the vocalic gestures of the second syllable audible, a rearrangement of the timing of the vocalic gestures is necessary. Thus the increase in overlap of the dorsal-labial schwa-forming gesture {swdl} occurrs together with the dorsal labial long /a/-forming gesture {aadl}.
For all cases given above, discrete segmental changes can be generated by continuous gestural alteration processes, which confirms a main hypothesis of Articulatory Phonology. But three categories of segmental changes remain which cannot be generated by a continuous gestural alteration process even if the identification of the hypothetical gestural process is no problem: the regressive assimilation of manner and the progressive and regressive assimilation of place. We will focus here on regressive assimilation of place. In the case of the examples "mit mei<u>n(e)m</u>", "mit je<u>d(e)m</u>", and "mit fe<u>tt(e)m</u>" the gestural alteration process is easy to identify: decrease of temporal extent of the schwa-gesture of the last syllable. After the elision of [ə] we expect the segmental change associated with regressive assimilation of place: the change [nm]->[mm], [dm]->[bm], and [tm]->[pm], i.e. a change from "apical" to "labial" for the last but one consonant. But the transcriptions of the generated reduced forms do not indicate this change in all cases. In many transcriptions the changes [nm]->[nn], [dm]->[dn], and [tm]->[tn] occur. In order to clarify these findings we performed quantitative listening tests.

3.2 Listening tests for "mit meinem", "mit jedem", and "mit fettem"

Since so far the transcriptions were done by a single person, we performed a quatitative listening test. This seems to be important especially in the case of regressive assimilation of place for the forms "mit meinem", "mit jedem", and "mit fettem". As stated above, the trained phonetician does not perceive clearly regressive assimilation of place in all cases if a decrease of temporal extent of the schwa-gesture of the last syllable is introduced (fig. 6a). It is our hypothesis, that a further gestural process must be introduced to Articulatory Phonology: *swap of the gesture-executing articulator*. In the case of our examples this is a swap of the apical full-closing gesture of the last syllable for a labial full-closing gesture (fig. 6b). This is a discrete gestural process and will be labelled *gestural or articulatory reorganisation.*

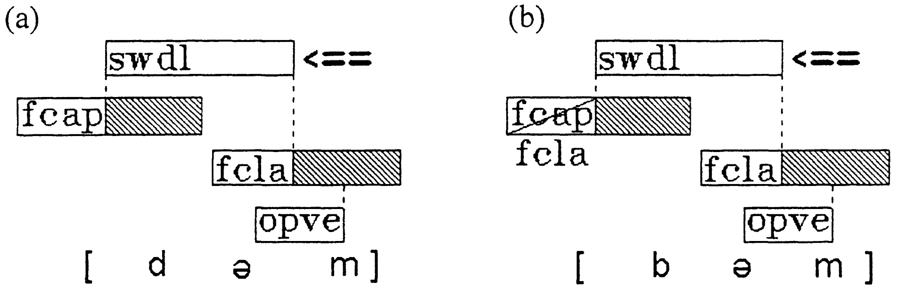
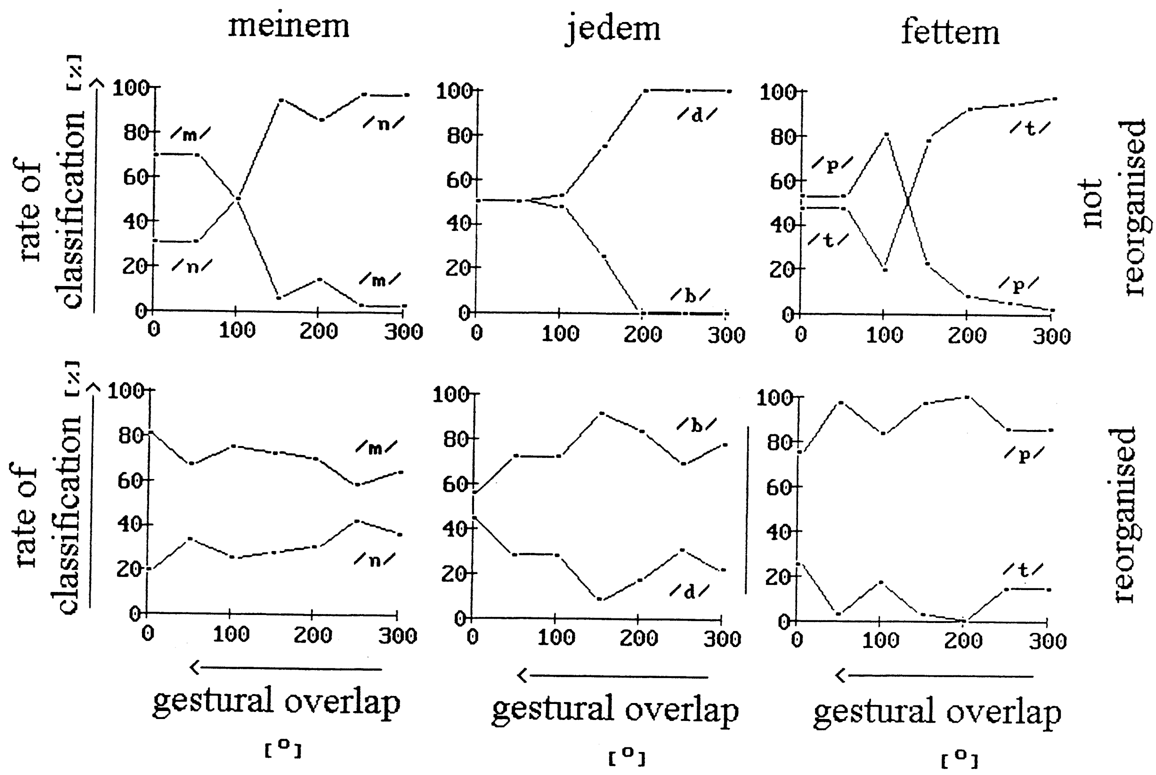


Figure 6 Continuous gestural alteration process (arrow) **(a)** without and **(b)** after reorganization for "mit je<u>d(e)m</u>".

In order to show that a continuous gestural alteration process is not sufficient to generate regressive assimilation of place, we performed quantitative listening tests. We started with six forms, i.e. the reorganised and not reorganised forms of "mit meinem", "mit jedem", and "mit fettem". For each form we performed the continuous gestural alteration process indicated in figure 6, i.e. decrease of temporal extent of the schwa-gesture of the last syllable. This process leads to an increasing degree of gestural overlap of the initial and final constriction-forming gestures of this syllable. Seven stimuli covering the whole range of the continuous gestural alteration process from no overlap to maximum gestural overlap of the constriction-forming gestures were generated for each of the six forms. The total of 42 stimuli was presented in random order with three repetitions of each stimulus to a group of 12 listeners (students of phonetics). Two tests were performed using the same stimuli but with different tasks: classification of the last but one consonant (test 1) and classification of the last consonant (test 2) to the category "labial" or "apical".

The rates of classification are given in figure 7. In the case of no reorganisation we find a tendency from apical to labial with increasing gestural overlap. But this tendency for regressive assimilation of place is not robust. Even in the case of full gestural overlap (left side of the diagrams) we find a maximum classification rate for "labial" of only around 50% to 70% for these three forms. But in the case of test 2 we find a robust tendency from "labial" to "apical", i.e. a tendency *against* regressive assimilation of place. Consequently, regressive assimilation of place has not been perceived clearly in the case of the non-reorganised forms. On the other hand, in the case of reorganisation a high rate of "labial" is perceived, regardless of the degree of gestural overlap. This result was to be expected since in the case of reorganisation we have only labial closing gestures on the side of articulation. But it excludes any immanent tendency towards "apical". So gestural reorganisation is necessary in the case of these forms in order to generate regressive assimilation of place.

(a)



(b)

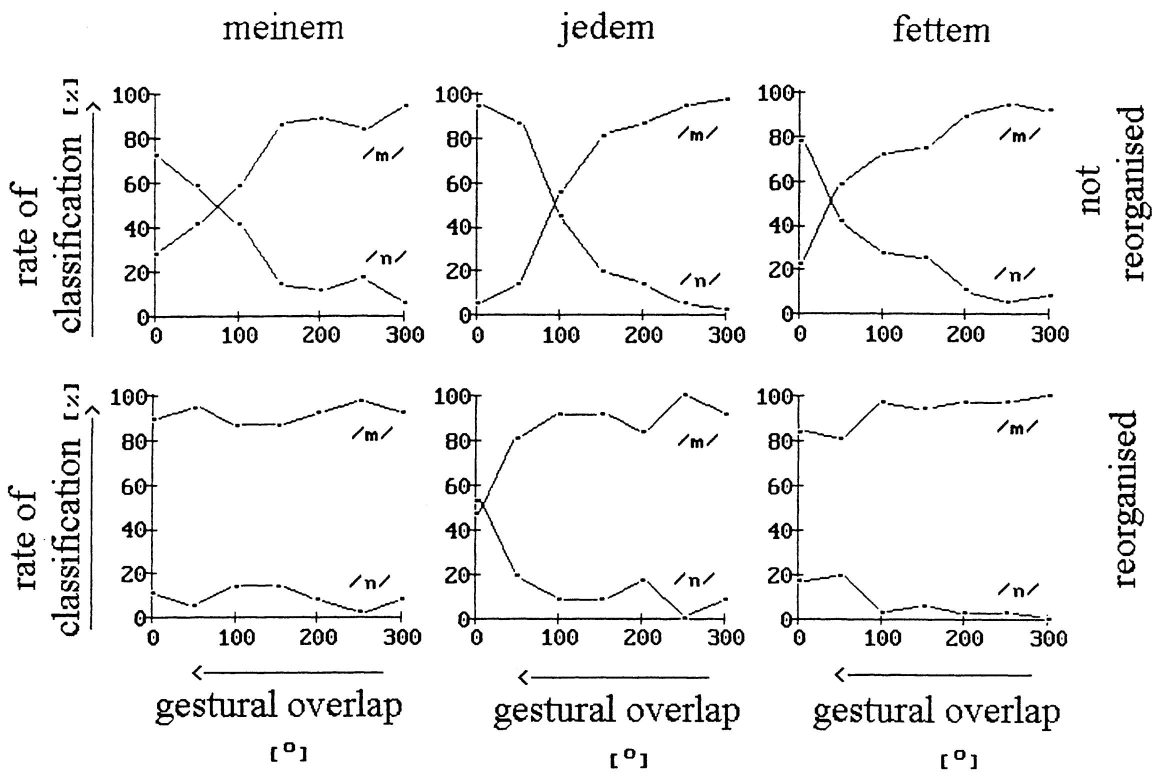


Figure 7 Rates of classification as function of gestural overlap (association phase values; increasing overlap from the right to the left side) for classification of the last but one (**a**; test 1) and the last (**b**; test 2) consonant for the three forms in the reorganised and the not reorganised case.

4 A gestural theory of reduction

It has been elucidated so far that different gestural processes are involved in reduction: two continuous gestural processes - i.e. increase in gestural overlap of gestures and decrease in temporal extent of a gesture - and one discrete gestural process: gestural reorganization. We assume one unique underlying driving force for all continuous and discrete gestural processes in reduction: *minimization of articulatory effort* (Lindblom 1990).

In our concrete gestural approach articulatory effort can be expressed quantitatively. Articulatory effort of an utterance can be defined as the sum-total of articulatory effort of all gestures occurring within an utterance. And effort of each single gesture is proportional to the duration of the appertaining gestural activation interval. We will find that this simple quantitative model of articulatory effort is capable of motivating many of the gestural processes in reduction.

In a more complex quantitative model of articulatory effort we can differentiate between effort of movement, i.e. effort of the transition portion, and effort of the gestural hold portion. The effort of the transition portion increases with increase in length of its time interval and with the distance of the target position to the initial articulator position - i.e. increases with the gestural movement amplitude - and increases with decrease in eigenperiod - i.e. increases with increase in strength of gestural activation (Kröger et al. 1995 and Kröger 1997).

In order to show that all gestural processes involved in reduction lead to a decrease in articulatory effort, we have to analyze the change in articulatory effort of an utterance (a word) for each gestural alteration process.
(1) Decrease of the temporal extent of a gesture: In this case articulatory effort decreases assuming that all other gestural parameters are constant. This follows directly from our quantitative expression for articulatory effort.
(2) Increase in temporal overlap of two gestures: (2a) If two gestures act on the same articulator, increase in overlap leads to a reduction in temporal extent of the preceding gesture. In this case gestural overlap is modelled by separating the time interval of overlap continuously for both gestures. In the first part of the interval of overlap mainly the first (or preceding) gesture remains active while in the second interval of overlap mainly the second (or following) gesture is active. This leads to a shortening of the temporal extent of both gestures. In our simple quantitative gestural approach (Kröger 1993) we introduced a dominance of the following gesture if two gestures overlap on the same articulator. This leads to a shortening of the temporal extent of the first (or preceding) gesture. But in both cases, we get a decrease in articulatory effort. This holds for all cases, i.e. for increasing the temporal overlap of tract-shaping gestures (i.e. increase in overlap of syllable cores; see chapter 3.1), for increasing the temporal overlap of constriction-forming gestures acting on the same articulator and of opening gestures acting on the same articulator. (2b) If two gestures act on different articulators the decrease in articulatory effort from other sources. Here we must differentiate three main groups (chapter 3.1): (A) Increase in overlap of constriction-forming gestures occurring in a syllable initial consonant cluster with those occurring in a syllable final consonant cluster. This process presumes the decrease of temporal extent of the tract-shaping gesture of this syllable (e.g. in the case of elision of [ə], example "eben" chapter 3.1). (B) Increase in overlap of constriction-forming gestures occurring in a syllable final consonant cluster with those of the syllable initial consonant cluster of the following syllable. This case occurs only if the constriction-forming gestures act on the same articulator (e.g. for reduction of double consonants in "mitteilen", for regressive assimilation of nasality in "Agnes" or for sonorization in "ratsam", chapter 3.1). For constriction-forming gestures acting on different articulators this case leads to a decrease in articulatory effort, if overlap of tract-shaping gestures is reached leading to a decrease in temporal extent of these gestures. This condition is satisfied if tract-shaping gestures overlap more strongly. (C) Increase in overlap of constriction-forming gestures occurring in the syllable initial or final consonant cluster. This case occurs only if the constriction-forming gestures act on the same articulator (e.g. for elision of [t] in "Glanz", chapter 3.1). For constriction-forming gestures acting on different articulators this case leads to a decrease in articulatory effort if overlap of tract-shaping gestures is reached which leads to a decrease in temporal extent of these gestures.
(3) Swap of gesture-executing articulator: This discrete gestural alteration process reduces the articulatory effort of an utterance in definite cases: (3a) Swap of constriction-forming gestures from tongue tip to lips or tongue body. In these cases we have a swap from the apical to the labial or dorsal tier. Since the tract-shaping gestures (i.e. the vocalic gestures) are always active on these tiers, this swap leads to an increase in overlap of gestures on the labial or dorsal tiers and subsequently to a reduction in articulatory effort. It is important to mention that the condition of swap from apical to labial or dorsal covers all cases of assimilation of place occurring in German (see chapter 3.2). So only shifts from apical to labial or from apical to dorsal occur but not vice versa. (3b) Furthermore it should be mentioned that the occurrence of glottalization and glottal stops (Kohler 1994) can be interpreted as a swap of a closing gesture

from the velum to the glottis in the gestural framework (e.g. in "Stu<u>nd(e)n</u>" [ndn] -> [nʔn] or as glottalization: [ndn] -> [nn̰n]). The glottal tier exhibits phonatory gestures without temporal gap: Closing gestures with medium adductive force produce voicing and opening gestures produce voicelessness (In the case of our simple model voicing results from gestural underspecification on the glottal tier). Consequently this swap to the glottal tier leads to an increase in gestural overlap of the (former velic) closing gesture with phonatory gestures. Furthermore, in the more complex model of articulatory effort we can assume that a swap of a closing gesture from velum to glottis leads to a decrease in gestural movement amplitude and thus to a decrease in effort of movement.

5 Conclusions

We have illustrated some advantages of Articulatory Phonology. A close relationship between phonological description and its phonetic realization is reached. The gesture as central unit of this theory can be seen as a phonetic unit (unit of action, goal-directed articulator movement) as well as a phonological unit (unit of phonological contrast). Articulatory Phonology is able to ascribe a lot of different discrete segmental changes (elisions and assimilations) to simple continuous gestural alteration processes: decrease in temporal extent of a gesture and increase in temporal overlap of gestures. But some segmental changes (e.g. assimilation of place in German) can be realized only by the discrete gestural process called "swap of gesture-executing articulator" or "gestural reorganization". It has been emphasized that the underlying driving force for all gestural processes mentioned here is the minimization of articulatory effort. In future work we will try to establish a complete rule system for the specification of concrete gestural processes in given utterances. This, however, will make it necessary to incorporate both articulation and perception into Articulatory Phonology.

6 Literature

Browman, C.P., Goldstein, L. (1989): "Articulatory gestures as phonological units", *Phonology* **6**, 201-251.

Browman, C.P., Goldstein, L. (1990): "Tiers in articulatory phonology, with some implications for casual speech", in: J. Kingston, M.E. Beckman (Eds.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech* (Cambridge University Press, Cambridge), S. 341-376. Also in: *Haskins Laboratories Status Report on Speech Research* **SR-92** (1987), 1-30.

Browman, C.P. (1991): "Consonants and vowels: Overlapping gestural organization", *Proceedings of the XIIth International Congress of Phonetic Sciences*, Vol.1, 379-383.

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

Fujimura, O. (1981): "Temporal organization of articulatory movements as a multidimensional phrasal structure", *Phonetica* **38**, 66-83.

Fujimura, O. (1986): "Relative invariance of articulatory movements: An iceberg model", in: J.S. Perkell, D.H. Klatt (Eds.), *Invariance and Variability in Speech Processes* (Lawrence Erlbaum, Hillsdale, New Jersey), 226-242.

Fujimura, O. (1992): "Phonology and phonetics - A syllable-based model of articulatory organization", *Journal of the Acoustical Society of Japan (E)*, **13**, 39-48

Geumann, A., Kröger, B.J. (1995): "Some implications for gestural underspecification as a result of the analysis of German /t/-assimilation", *Proceedings of the 13th International Congress of Phonetic Sciences,* Vol. 3, 374-377.

Kohler, K.J. (1994): "Glottal stops and glottalization in German", *Phonetica* **51**, 39-51.

Kohler, K.J. (1995): *Einführung in die Phonetik des Deutschen* (Erich Schmidt Verlag, Berlin), 2. edition.

Kröger, B.J. (1993): "A gestural production model and its application to reduction in German", *Phonetica* **50**, 213-233.

Kröger, B.J., Schröder, G. Opgen-Rhein, C. (1995): "A gesture-based dynamic model describing articulatory movement data", *Journal of the Acoustical Society of America* **98,** 1878-1889.

Kröger, B.J. (1997): "Ein quantitatives Konzept des artikulatorischen Aufwandes", in: K. Fellbaum (Ed.), *Elektronische Sprachsignalverarbeitung. Studientexte zur Sprachkommunikation* **14**, 248-255.

Liberman, A.M., Mattingly, I.G. (1985): "The motor theory of speech perception revised", *Cognition* **21,** 1-36.

Lindblom, B. (1990): "Explaining phonetic variation: A sketch of the H and H theory", in: W.J. Hardcastle, A. Marchal (Eds.), *Speech Production and Speech Modelling* (Kluwer Academic Press, Dordrecht), 403-440.

Löfqvist, A., Yoshioka, H. (1984): "Intrasegmental timing: Laryngeal-oral coordination in voiceless consonant production", *Speech Communication* **3**, 279-289.

Öhman, S.E.G. (1967): "Numerical model of coarticulation", *Journal of the Acoustical Society of America* **41**, 310-320.

Saltzman, E.L., Munhall, K.G. (1989): "A dynamical approach to gestural patterning in speech production", *Ecological psychology* **1**, 333-382.