Thinking *soap* but Speaking 'oaps'.
The Sound Preparation Period: Backward Calculation from Utterance to Muscle Innervation[1]

Nora Wiedenmann
Amalienstr. 71 Mgb, Munich, D-80799, Germany
n.wiedenmann@gmx.net

Abstract

In this article's model—on speech and on speech errors, dyscoordinations, and disorders—, the time-course from the muscle innervation impetuses to the utterance of sounds as *intended* for canonical speech sound sequences is calculated *backward*. This time-course is shown as the sum of all the known physiological durations of speech sounds and speech gestures that are *necessary* to produce an utterance. The model introduces two internal clocks, based on positive or negative factors, representing certain physiologically-based time-courses during the *sound preparation period* (*Lautvorspann*). The use of these internal clocks shows that speech gestures—like other motor activities—work according to a simple serialization principle: Under *non*-default conditions, alterations of the time-courses may cause *speech errors* of sound serialization, *dyscoordinations* of sounds as observed during first language acquisition, or *speech disorders* as pathological cases. These alterations of the time-course are modelled by varying the two internal-clock factors. The calculation of time-courses uses as default values the sound durations of the *context-dependent* Munich PHONDAT Database of Spoken German (see Appendix 4). As a new, human approach, this calculation agrees mathematically with the approach of Linear Programming / Operations Research. This work gives strong support to the fairly old suspicion (of 1908) of the famous Austrian speech error scientist Meringer [15], namely that one mostly thinks and articulates in a different serialization than is audible from one's *uttered* sound sequences.

Keywords: speech production; speech muscle-innervation serialization; sound serialization; speech gesture timing; speech errors; speech pathology; speech disorders; speech physiology; earliest possible sound-related reafferences; ERPs; robotics.

1. Introduction

The acoustical time-course of an utterance is mostly taught as the fundamental part of basic phonetic or articulatory lectures on speech—it is qualitatively and quantitatively known in detail. However, the time-course (after speech intention) *between* each sound-related earliest impetus (which is necessary for muscle innervations to produce the different articulatory movements) and the resulting beginning of each sound in a sequence of sounds has not yet been sufficiently investigated. The mathematics used in this article models the utterance mechanisms required to articulate sequences of speech sounds according to approaches of *Linear Programming* (Kelley, 1961 [9]; Wille, Gewald, Weber, 1966, 3rd edn. 1972 [23]). This approach here may represent a paradigm change in phonetics, which could affect first language acquisition and speech pathology.

2. Model

The basics of modelling speech mechanisms required to articulate sequences of speech sounds is well-known from the literature on phonetics. The utterance of sounds shows certain different default durations of sounds (and certain different pauses between them): There are long durations and short durations. These duration differences exist in a similar way for the time-course already

[1] A first-draft version of this paper was presented at the 31st Annual Meeting of the Gesellschaft für Angewandte Linguistik at the University of Bremen, FRG, in Sept. 2000, based on the doct. diss. (1997, Univ. Munich, and its *Lautvorspann* ('sound preparation period'); see Wiedenmann (1999, p. 15)), with the first introduction of two internal clocks.

before the beginning of an uttered sound. This time-course is termed here the *sound preparation period* (see Wiedenmann, 1999, p. 15: Lautvorspann [22]).

To be prepared for articulation, some sounds require long *preparation periods* due to a relatively complex system of innervation of the large number of involved muscles necessary to produce the articulatory movements; other sounds need only a short *preparation period* due to a simple muscle-innervation system in which only a few muscles' *jobs* (see below) are needed. According to Hardcastle, 1976, pp. 134-7, [7] the most complex system consists of about 25 muscles for the speech gestures of the fricative [s] [2]. *In contrast*, the dorsovelar plosives [g] or [k] are very simple to pronounce: According to Laver, 1977, [11] the "soft palate" or "the velic system is probably the simplest of all the muscle systems" used in producing articulatory movements [3].

Research results provide muscle-innervation time-courses of sounds for certain consonants; see, for instance, Kühnert, 1989 [10], with her electromyographic (EMG) data of the bilabial plosive [p] showing about 250 ms from the muscle-innervation impetus until the resulting plosion burst as the *sound preparation period*; see also, for instance, the empirical EMG data—related to the audio-signals—for some speakers in Kühnert, 1989, and the data of Löfqvist and Yoshioka, 1981 [13].

For the sake of clear terminology, it has to be emphasized that such an innervation impetus must *not* be understood as an impulse (that is, a very high, sharp up and down (or peak) of voltage—during a very short time—as physiologically known in brain neurosciences), but rather as the relatively smooth starting of muscle-innervation voltage necessary to produce an articulatory movement (see Kühnert, 1989, taking in each case the accompanying audio-signal as a reference to the voltage).

To demonstrate the *default* time-course relations for an utterance, it is only necessary to work with the physiological system's *extreme* values (that is, minimum and maximum values): very long innervations versus very short ones. To produce a [b] (or a [p]), one needs about 250 ms (see Kühnert's EMG data) from the earliest muscle-innervation impetus until the resulting utterance of this plosive sound; however, a dorsovelar plosive [g] or [k] implies perhaps only 80 ms (see Laver's "simplest of all the muscle systems", thus entailing a very short innervation period; here, generally disregarding the *voicing* gesture (glottal gesture) of a sound). For this model, it is *only important to know* that *differences in duration* do exist in the *sound preparation period* in a similar way as in the very *sound* duration.

Fig. 1a shows the default time relations for the utterance [pak] with its *sound preparation periods* p' until [p] and k' until [k]; from these data, one can see the long time needed to produce the [p], followed by a relatively short time period for the [k]—the [k] just in time after an [a]. This [a] is not shown with its gesture time-relations in all diagrams but rather with its duration. For the sake of simplicity, the diagrams are restricted to *consonantal gestures* and to *vowels with consonantal gestures* (for instance, i.a. the lip-rounding gesture as necessary for bilabials) of an [u] or an [o]. The earliest necessary impetus of different muscle innervations p' for a [p] is followed by the utterance-initial sound [p] itself, which, in turn, is followed by the earliest necessary impetus of muscle innervations k' for a [k], followed by the intended utterance-final sound [k] itself (see Figure 1):

[2] Concerning this complex [s]-producing system, Shattuck-Hufnagel and Klatt [17] (1979, p. 43) show relatively high error rates in their Confusion Matrix of English speech errors. In Wiedenmann (1999), the relative sound frequencies of erred German consonants were taken into account (see tables in Wiedenmann [20] (1998a, pp. 33-38): Weighted with these sound frequencies (based on 32,000 sounds of spoken language), the data show relatively high frequencies for dissimilations and non-dissimilative speech errors with [z] and [s]; see Wiedenmann (1999, distribution histograms of pp. 78, 80, 93, and 95/6).

[3] Laver (1977, pp. 145-6) speaks of the (paired) muscles: "palatal tensor", "palatal levator", "palatoglossus"—attached to "the sides of the back of the tongue"; "each different position of the tongue will entail a different state of contraction of all the palatal muscles"—, and the "superior pharyngeal constrictor".

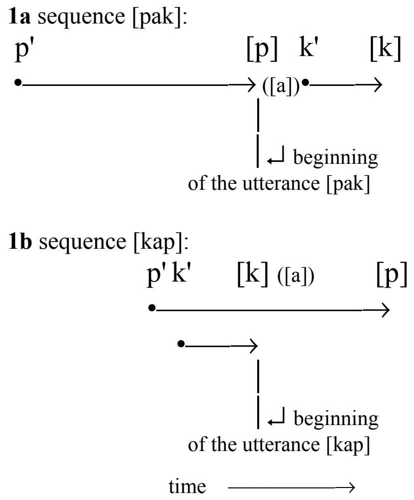

Figure 1. Sound preparation period and default utterance as a function of time (true to scale according to 250 ms for p' until [p])

A stroked letter represents the starting point for the necessary muscle innervations of a sound; thus, the line between the beginning of the arrow (the earliest muscle-innervation impetus, indicated by a dot) and the corresponding tip of the arrow (the beginning of the sound utterance, the acoustic realization of a sound) represents the *sound preparation period*. Every sound duration is the time interval between its arrow tip and the arrow tip of a possibly next sound. **1a** The entire production time for [pak] between p' and the burst of the last sound [k] shows *no overlap of arrows*; **1b** in contrast to [pak], an overlap of (the arrows symbolizing) the different muscle innervations (*sound preparations*) for the utterance [kap] *does exist*. A sound preparation period of an utterance ends here with the beginning of the final consonant (as in Kühnert (1989), the end of [pak] or [kap] is the burst for the final [k] or [p]).

Figure 1b shows the time relations of the *inverse* utterance [kap]. Compared with the time p' until [k] for the entire utterance [pak] of Fig. 1a, the time p' until [p] for the *inverse* entire utterance [kap] of Figure 1b is a relatively short period, namely *only* the time between p' (the *earliest necessary* impetus for muscle innervations for [p]) and the utterance-final (or word-final) [p] itself. The time period p' until the utterance-*final* [p] for [kap] contains also (and it runs *parallel* to) the *sound preparation period* k' until the burst of this [k]. In this example [kap], this overlap in time requires that all the speech gestures to articulate a [p] and a [k] be—from a time perspective—independent of each other. In other words, these sounds do not use the *same* muscles at the *same* time (as, for instance, in sound dissimilations).

For the utterance [kap] in Fig. 1b, one has to take into account that—sound-related—the time period from p' to [p] as well as the time period from k' to [k] are the same, as can be seen in Fig. 1a, as for the utterance [pak].

Now, the *backward calculation* [4], which is carried out by the programme used for this article's model, will be explained, beginning with Fig. 1a (*without* the use of any internal clock):

The sequence [pak] ends at the [k], and the arrow tip at [k] corresponds to the instant of this sound's burst (and this is—as in Kühnert, 1989—the *reference point*, here, to begin with the *backward calculation*). The mean duration value of an [a] articulated before a [k] is known; with

[4] The programme for *backward calculation* takes as utterance—in this version—a maximum of 10 IPA symbols (of all German sounds, including the initial sound of French *journal* for certain loan words and the German allophones [R] and [r]; in broad transcription, but without characters for vowel lengths). Concerning the specific uttered sounds, the programme works on audible sounds (on their *durations*; not on formant transitions). Based on these sounds, the programme delivers also the non-audible time structures (periods of muscle innervations for the speech gestures) which are necessary for the entire sound-producing system. Thus, it brings together the audible utterance and the inaudible articulatory movements, which are nevertheless important in its time-course because of the production of *kinaesthetic-proprioceptive reafferences*. The combination of uttered sounds and their muscle innervations results in certain time structures, which are described by a network of arrows. In a future programme version, input IPA characters for the sounds of languages other than German (always in view of the language-specific phonotactics) could be added to the background databases to indicate the durations of these sounds and also the duration of their muscle innervations for all articulatory movements.

this value (of the PHONDAT Database; see Appendix 4), one has to go to the left on the time axis, and the found instant is the *beginning of the utterance* [pak], where (again) the arrow tip corresponds to the instant of the burst to make the [p] audible. Now, one has again to go to the left (backward!) on the time axis with the *sound preparation period* of the word-initial [p]. Thus, the entire duration of this isolated utterance [pak] is p' until [k].

Fig. 1b (also *without* the use of any internal clock): The sequence [kap] ends at the [p], and the arrow tip corresponds to the instant of the bilabial burst (the reference point to begin with the *backward calculation*). The mean duration value of an [a] before a [p] is known with this value (of the PHONDAT Database). One has to go to the left on the time axis, and the found instant is the *beginning of the utterance* [kap], where (again) the arrow tip (see one line under the [k]) corresponds to the instant of the burst of the [k] to make this sound audible. Now, one has to go to the left (backward!) of this arrow tip for the [k] on the time axis with the *sound preparation period* of the word-initial [k]. This step ends at k', indicated in the line above. In the line with the arrow tip for the bilabial [p], one has to go backward with the *sound preparation period* of the word-final [p]. This step ends in p'. Thus, as opposed to Fig. 1a, the entire utterance duration for this isolated utterance [kap] is the time from p' until [p]. Note that in this case, the entire duration for this inverse utterance [kap] is only as long as the *sound preparation period* for the [p]. The sound with the longest *sound preparation period* may determine the lengthy process of an entire utterance's *preparation period*, even when this sound has to be uttered *utterance-finally*.

The time periods in Fig. 1a and 1b express in both cases speaker-individual time-courses. Using the PHONDAT Database of *context-dependent* sounds concerning German phonotactics, a table was built containing for each sound the PHONDAT sound durations' *mean* values (to simulate another speaker-individuality (or the speech of a specific personality), the German sound duration values in the table used by the programme could be changed according to certain speaker idiosyncrasies; as a base, see Appendix 4). For *backward calculation*, *the programme uses these mean values* as *default* values for sound durations. (For speech errors with the sounds [p] versus [k], [g] versus [b], and vice versa, see Appendix 3 with data in several languages.)

Now, what are the consequences of the two *internal clocks* on this model? What should then be taken into consideration in view of *backward calculation*?

In Fig. 2b, both *internal clocks* were applied to show how variations in the time structures of the *default* utterance [sop] could explain the dyscoordination [ops] of a child.

(For the following explanations, refer to an utterance-specific example in Fig. 3. The two internal clocks—Cl-Inn on the ordinate, Cl-Imp on the abscissa—give an overview of the possible variations of an utterance and of its default values (around the default value 1 in the middle of the diagram of both internal clocks). The coloured areas (respectively the hatched areas in overview diagrams for the clock factors of other figures) stand for changed time-courses which lead to speech errors or disordered speech.)

The one internal clock (Cl-Inn on the ordinate of Fig. 3) allows to vary the duration of the *sound preparation period* (that is, the *length* of the entire *arrow*): A clock factor less than the (default) factor of 1 means *acceleration*, or *shortening* of all the sound preparation periods (a *shortened* length of all the *arrows*). A clock factor greater than 1 means *retardation*, that is, *lengthening* of all the sound preparation periods (an *increased* length of all the *arrows*).

The other internal clock (Cl-Imp on the abscissa of Fig. 3) allows to vary the time between the sequence of *muscle-innervation impetuses* (the *distance* between the *arrow beginnings*): A clock factor less than the (default) factor of 1 means *acceleration*, or *shortening* of all the *distances* between the *arrow beginnings* (a *shortened distance* between all sequences of *arrow beginnings*). A factor greater than 1 means *retardation*, or *lengthening* of all the *distances* between the *beginnings* of the arrow (an *increased distance* between all sequences of *arrow beginnings*).

In Fig. 2b, *both* clocks were applied: Note (in comparison with Fig. 2a, which is the default case) both the *shortened arrows* and the *shortened distances between the beginnings* of the arrow.

With the shortened or enlarged *sound preparation periods* (arrows) and the shortened or enlarged time between the *muscle-innervation impetuses* (distances between the *arrow beginnings*), the procedure of *backward calculation* can be carried out in the same way as without clock-changed time-courses (to indicate the specific *Clock Factors* of an utterance, see also in Fig. 4 (and also in the diagrams of the *Appendix 1*) the two rectangular *labels* for Cl-Inn and Cl-Imp at the bottom. For the overview diagram, the programme uses the two labelled values as an input to determine the location of the little (red) circle (an *operating point*; see in Fig. 4, above on the right hand of the IPA symbols of the programme's utterance input); in the case of Fig. 4, this operating point lies in the left hatched area: Cl-Inn and Cl-Imp (as labelled at the bottom) < 1).

Generally, the time consideration in this article also shows that *hidden* gestures *can be co-produced* with certain gestures of the utterance, that is, they can overlap in time as mentioned above. On hidden gestures, see Browman [1] and Goldstein [6], 1987. In certain cases, some gestures do not result in a sound, for instance, when articulating the tip-alveolar gestures for the [t] in the intended utterance *a perfect memory*. Such an overlap in time for an articulatory occlusion (as in the mentioned example) for the later following [m] only occurs because the [m]-occlusion is *more* to the front of the vocal tract than the occlusion for the [t]; see Goldstein, 1989, p. 94 [6], and his remark on a "completely hidden" sound gesture as well as his assertion that "the acoustic output" will be determined "by the entire ensemble of concurrent gestures". However, in the case of the German utterance "im Himblick auf" instead of the canonical *im Hinblick auf* ('in view of'), the hidden gestures for the tip-alveolar [n] (the *preparation period* for the [n] is *shorter* than that for the [m]) may be present during the *preparation period* of the gestures for [mb], but in this case without audible results. For, in the case of a speaker's "Himblick"—perhaps a lifelong heard and learnt (and in this way *intended*) utterance—, the "gestural complexes" (see Goldstein, 1989, p. 95 [6]) "may contrast not only in the set of gestures involved, but also in their organization". Such *organization* in time is presented here.

The muscle-innervation impetus serialization (see Figure 1a and b) for the utterance [pak] is p'k'. However, for the inverse utterance [kap], the required impetus serialization is also p'k'—and *not* k'p'. This serialization results from the values of the two extremely different *sound preparation periods*: a long one for [p] and only a short one for [k].

In the Figures 1 and 2, the line between the *beginning* of the arrow (the earliest muscle-innervation impetus) and the *tip* of this arrow (that is, the moment of the resulting sound utterance) represents the entire *sound preparation period*. Thus, this line merely *simplifies* a *concert* of speech gestures (innervated muscles) for every sound (see as previously mentioned Goldstein's "gestural complexes"). Without this simplification, the modelling of the serialization of all needed sound-specific different muscle jobs (for instance, the orbicularis oris muscle for lip-rounding) requires a graph-theoretically complex deterministic *job*-oriented network (see Wille, Gewald, Weber's (1966) Programme Evaluation and Review Technique diagram, Table II, 1, and Kelley, 1961, p. 298). In such time-course network, a *critical path* (according to an Operations Research approach) of muscle jobs can be determined. In this critical path, no (further) time-course *compression* is possible, so that the sum of these muscle jobs cannot further reduce the *sound preparation period* represented by the arrow in the diagrams.

But what happens, and what alterations result *if the time periods* between the different earliest impetuses for muscle innervations (between different *beginnings* of the arrow for sounds in the figures), *or the innervation durations themselves would change*?

This model considers only *linear* time-alteration mechanisms. (In phonetics, *prepausal lengthening* and *utterance-final lengthening* are well-known retardations of speech—*non*-linear default phenomena, but not dealt with in this paper.) Two *internal clocks*, working independently of each other, are introduced into the model—an innovation in the field of articulatory phonetics. The first internal clock is used for varying the timing of the different impetuses for muscle innervation (the time between the different beginnings of the arrow in the diagrams); the second internal clock is used for varying the duration of each innervation process itself (the arrow length in the diagrams).

Both internal clocks can represent either a *retardation* (clock factors >1), or an *acceleration* (clock factors <1).

3. Consequences of the model concerning alterations in speech serialization

The four different states of these two internal clocks (working independently of each other and linearly with time) and their combined actions show, on the one hand, the production of intended speech and of involuntary speech errors, and, on the other hand, the production of pathological speech in the widest sense.

Variations of the internal clocks can provide material to explain data found in the English literature: Leonard and McGregor, 1991 [12], described the logopaedic case of a child with developmental sound dyscoordinations. During a certain period of first language acquisition, this child had not yet *learnt* the specific and correct muscle-innervation *time-course* needed to articulate (among other fricatives) the *word-initial* prevocalic fricative [s] as in the intended child utterance [sop] for the word *soap* (without diphthongization of the vowel in this state of speech development). Instead of [sop], this child was only able to utter [ops]. The following Fig. 2a shows the default case (that is, both internal clocks' factors are 1) of an *adult* uttering [sop]:

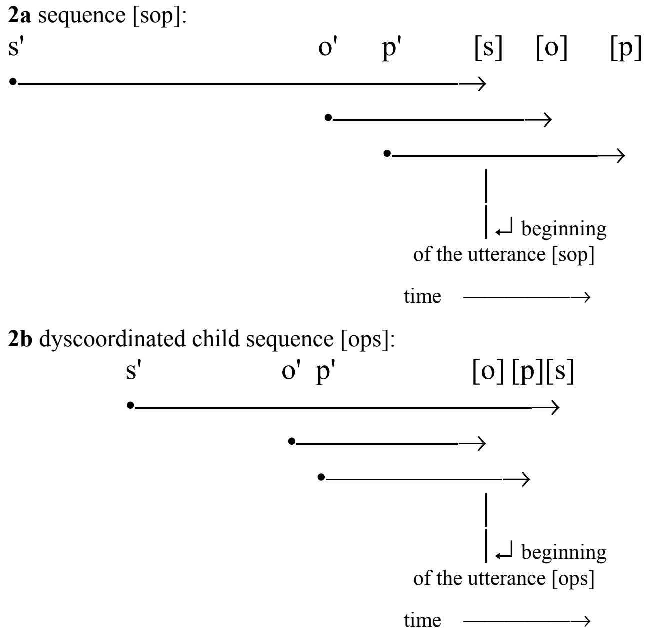

Figure 2. Sound preparation period and utterance as a function of time (true to scale according to 250 ms for p' until [p])

2a the default (adult) utterance [sop] is shown with the default factor 1 of both internal clocks (see text below) and with its default sound duration beginning at every arrow tip; **2b** the internal clocks show, firstly, an acceleration factor of muscle-impetus sequencing of .5 (i.e., the time difference between the beginnings of the two arrows, between s' and o', is only half of that in 2a, for instance) and, secondly, .9 as an acceleration factor of muscle innervation. Thus, the time difference between s' and [s] (i.e.: the length of the arrow) is only 90% of that in 2a, for instance; and for the other sounds, analogously. Therefore, the muscle-innervation impetus s' comes too late to produce [s] in time, and the result is [ops], resulting in sound durations reduced.

In Fig. 2a, the default (adult) serialization of the s'- and then of the o'-impetus (labial gestures for [o]) is followed by the p'-impetus (assuming the *sound preparation period* of the very complex fricative [s] to be the longest *sound preparation period* (according to Hardcastle, 1976) compared with the *preparation periods* of the other two sounds).

Fig. 2b shows one possible time-course of a *dyscoordinated* (child) utterance. As to the muscle innervations for the vowel [o] (see Fig. 2a), the child begins too late (related to o') with the innervation for the [s]. See the two internal-clock factors in Fig. 3: acceleration factor of muscle-impetus sequencing of .5 and acceleration factor of the process of muscle innervation of .9 (i.e., the time period between s' and o' in Fig. 2b is only half of that in Fig. 2a, and the time period between s' and [s] in Fig. 2b is only 90% of that in Fig. 2a, and thus also for the two sounds [o] and [p]; compare the length of the lines in diagram 2a versus those in diagram 2b).

Alterations of the two internal-clock factors—i.e., retardation or (here) acceleration relative to the default case of Fig. 2a—result in a changed sound serialization of the default utterance due to the altered values of all muscle-impetus sequencing periods and also due to the altered values of all muscle-innervation periods (see Fig. 2b). All these altered values represent the sound preparation period. These alterations show the involuntary and non-intended child utterance [ops], as described in Leonard and McGregor, 1991 [12], which also lead to *non*-default sound durations. These involuntary and non-intended utterances occur in spite of an, in principle, *correct* muscle-impetus sequencing (as to be learnt by a child for [sop]; see the *beginnings* of the arrows; cf. Fig. 2a). These aforementioned impetus sequencings and sound durations follow from the mathematical model.

The diagram of Fig. 3 with its four quadrants shows as an overview the *possible* factor combinations of the two internal clocks, whereby the range of factors for each clock is limited[5]. This diagram is shown for the same utterance example [sop]: The (coloured) areas (on the left of the white area) represent the two possibilities of *sound reversals* ([ops], as dyscoordinated by the child of Fig. 2, or [osp]).

The size of the two areas for sound reversals in Fig. 3 depends on the degree of deviation from the two internal clocks' *default* factors of 1, as well as on the intended *default* sound durations of the utterance. In this paper, these durations have been taken from PHONDAT, the Munich University Database of Spoken German. From these data, a phonotactic table was constructed for the sound durations, and this table (see Appendix 4) was used by the programme. The larger the white (in some overview diagrams the non-hatched) diagram area around the default factors of 1 (see the point 1.0; 1.0 in the diagram of Fig. 3, respectively the little (red) circle in the overview diagrams with hatched areas), the more *ideal* the utterance is and the less the danger of deviations from an *intended* utterance.

An *ideal* utterance would be accompanied by a canonical or intended sound serialization, whereby only slight alterations of sound durations in a wide range of the clock parameters occur. In this sense, an ideal utterance (or an *ideal* word) is robust against disturbances. The German word *Baum* ('tree'; see the diagram in the Appendix 1: Fig. 5) as an isolated utterance is an example of such an ideal utterance, because of *similar* muscle-innervation durations for [b] and [m]. The programme used for the model makes it possible to investigate the ideal-state degree of every utterance or word (see also the diagram (in the Appendix 1) of an *ideal* word: Fig. 6: German *dort* ('there')); to compare the ideal-state degree with these two figures, see for instance the English child utterance [sop] in Fig. 3:

[5] Concerning the ranges of the two clocks in Fig. 3, it has to be discussed whether the maximal value 1.5 on the ordinate and the maximal value 3.0 on the abscissa may occur in reality as a very high retardation (see the diagrams in Appendix 1 (Figures 11/12) for the intended German *mich gebryft* (instead of the canonical German *mich geprüft* ('examined me') resulting in the error utterance "mich begryft" of an old man of 94).

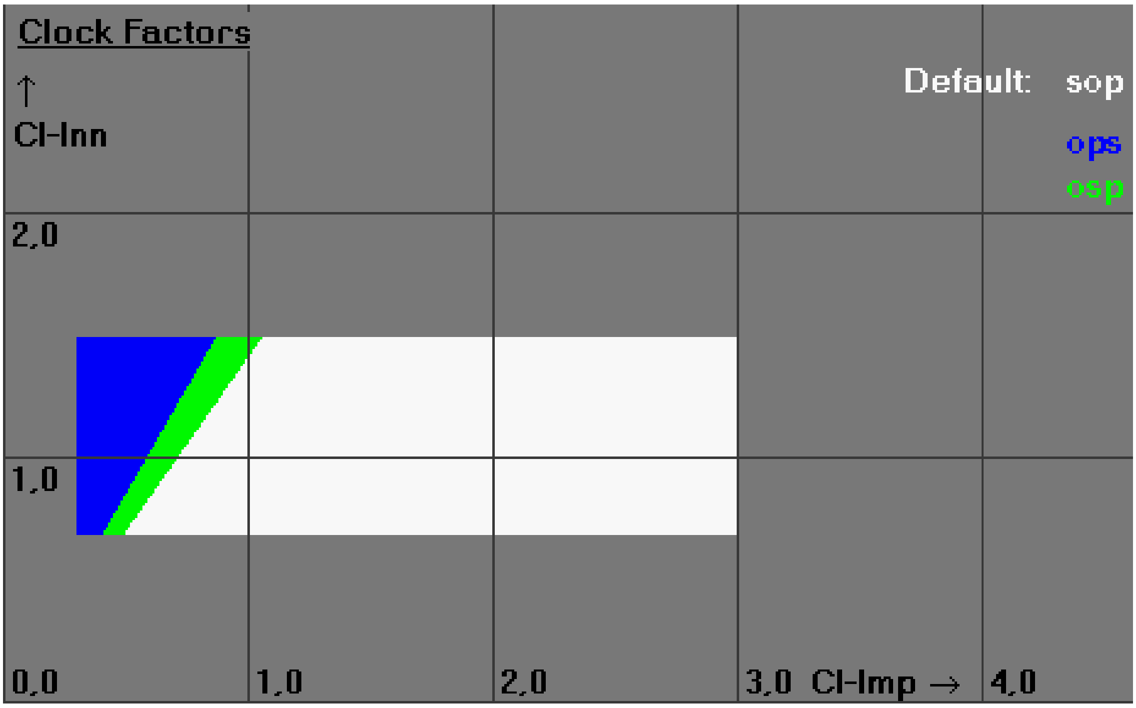

Figure 3. Two Internal Clocks and a specific intended utterance (an overview diagram)
The example is an intended *soap* ([sop]) with the two possible reversals [ops] (blue area on the left; see Fig. 2b) and [osp] (the (green) area on the left of the white area), whereby the point 1.0; 1.0 represents the default case of the utterance (both factors are 1); on the ordinate: clock factors of varied muscle-innervation periods; on the abscissa: clock factors of varied muscle-impetus sequencings. The white area represents possible utterance variations of the specific utterance, namely only changed sound durations; in this white area, the *sound serialization* corresponds to the child-intended [sop].

The boundaries between the diagram areas—white or the differently coloured—mathematically result from phase transitions; their number and gradient are *utterance-specific*.

Examples of more complex utterances, including consonant clusters, are the German words *Fracht* ('freight'), *Brief* ('letter'), and *Frachtbrief* ('waybill'; to compare time relations and overlapping muscle innervations, the latter German compound is at first shown as articulated with an uvular [R] and then with a tip-alveolar [r]) in syllable-initial consonant cluster positions—see the Appendix 1, Figures 7-10; Figures 13 and 14 illustrate the time relations of the German words *Kommunist* / *Kommunisten* ('communist' / 'communists'). All these examples of more complex utterances (and for values of the two internal clocks in higher ranges) show sound serializations that are very unusual in the field of speech error research. In pathology however, similar, almost unpronounceable sound serializations—disorders—do occur.

4. Discussion

The internal clocks' acceleration could be interpreted as a speaker's excitement; retardation could also explain certain pathological speech disorders, after a stroke or in higher age. Many combinations of the internal-clock factors stand for serialization errors, which may be interpreted in a *pathological* sense or simply as *speech errors*. The utterance [sop], *soap*, as well as Elsen's (1991) [4] case of German *Frau* ([fRao]; 'woman') are not *ideal words* (not easy to speak; see the child utterance [Rvao] instead of [vRao] in Fig. 4).

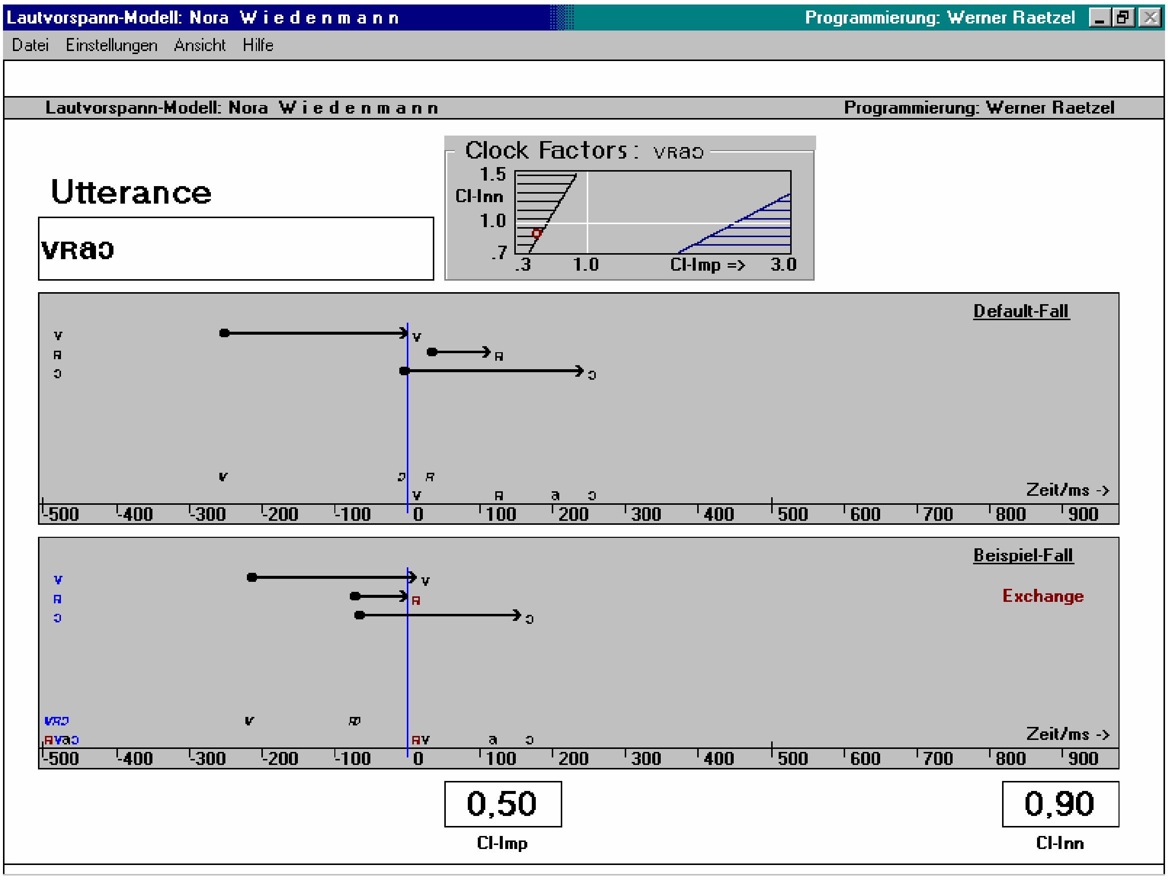

Figure 4. Diagram on the figure's top: *Two Internal Clocks* and a specific intended utterance in early first language acquisition

The diagram shows Elsen's German child example *Frau* ([vRao]; 'woman') and the two reversals [Rvao] (left hatched area; the clock factors are shown by the little (red) circle) and [vaoR] (right (blue) hatched area, not described here). Below: Two time-course diagrams for the *child-intended* utterance [vRao] as examples of sound preparation periods: at first above, the child's default case [vRao], labelled *Default-Fall* (uttered from 0 to 240 ms); below, the dyscoordinated child utterance [Rvao], labelled *Beispielfall* ('case of example'; uttered from 0 to 160 ms; the anticipated sound is highlighted in red). This dyscoordinated [Rvao] is produced with a stronger muscle-innervation overlap. This is mathematically based on the two internal-clock factors Cl-Imp and Cl-Inn: acceleration factor of muscle-innervation impetuses Cl-Imp of .5, acceleration factor of the duration of muscle innervation Cl-Inn of .9, as indicated in the two rectangular labels below the abscissa. In both cases (default / dyscoordination), the utterance begins at time 0 (positive times to the right, beginning at the (blue) vertical line). The time (as well as the name in IPA symbols) of each uttered sound is seen closely above the abscissa. Time periods in the negative area (to the left of the (blue) vertical line) show the needed sound-specific *preparation periods*: The time of the start of a *sound preparation period* is shown by the sound's name in italics.

When logopaedicly *trained* by a speaker, these non-ideal words have no chance of giving this speaker a quick feeling of success. However, in contrast to these two words, [pak] is a *relatively ideal* word: It is robust against false serialization caused by retardations (in a normal or pathological way, according to the clock factors) in the time network. However, there are utterances, which—*even for small alterations of the internal-clock factors*—result in deviations from the *intended* form and thus produce sound reversals.

In Fig. 4, in the case of the innervation-impetus sequence R' and o'—nearly a clash in time (see there the diagram at the bottom, at about -90 ms)—, the uvular [R] and the labial [o] need *different* muscles.

However, in some cases of sound sequences, one could think of a *clash* of two *innervation impetuses* of the *same* muscle needed to build speech gestures of *different* sounds or also of quickly following *same* sounds. This case is phonologically known as *progressive dissimilation* (i.e., *forward masking*, as already described in 1969 by MacKay [14]; see Wiedenmann's dissertation (1999), p. 127). An example of *progressive* dissimilation is the German utterance "Schluck Mi_ch" (with _ for the deleted sound) instead of the intended *Schluck Milch* ('gulp of milk'): A certain muscle needed for articulating the second [l] could not be quickly innervated a second time, due to a certain used muscle's refractory phase.

(However, and *not* this article's subject, the time-course *before* (and, in turn, overlapping with) *muscle innervations* of an utterance surely corresponds to the time-courses of *ideation*, the *imagining*, and the *planning* of this utterance—perhaps with similar serialization mechanisms in the time network (as shown for the time-course of muscle innervation and articulation). These mechanisms could be analogous to those shown here for changed time-courses of the *sound preparation period*; they could thus lead to *regressive* dissimilation (i.e., MacKay's (1969) *backward masking*; see Wiedenmann 1999, p. 138f), in some cases.)

This speech *production* (Lautvorspann) model shows a connection to speech *perception*, too: The *earliest* possible time to perceive sound-related *reafferences* is determined by the moment of the word's or the syllable's *earliest* sound-related muscle-innervation impetus. These reafferences can be measured via event-related potentials (ERPs), which are positive or negative, so-called waves: brain responses as voltages in microvolts.

These *earliest* sound-related reafferences of muscle innervations could also explain certain rare empirical data cases of a speaker breaking an intended utterance as soon as he, the speaker, has *felt* that a certain speech error is *imminent*. An example of such a case (see the Corpus of German Speech Errors by Wiedenmann 1992 [19], p. S77; see Wiedenmann 1998 [21], p. 75; chapter 3.9) is the German utterance

> ... *Spezifik – Spezifikation* ... ('specific—specification'; s. Wiedenmann, [19], p. S42: speaker Th.V., 1989/12/20 17h). The collector's comment is that the speaker (in NW's translation) had "early reafferently felt his [imminent] speech error". The collector's additional note of 1989 is "Preartic. Editing?". For the last-named term, see also Dell and Reich [2], 1980, p. 280; for more details, see Wiedenmann, [20], p. 185, with further data.

This example's trained speaker could already have *felt* that in the next moment he would have uttered an [o] for an erred *Spezifiko...*. However, this [o] with a lip-rounding for the German labial vowel [o], as perhaps anticipated from the accentuated word-final syllable *-tion* (in German articulated as: [tsjon]), should have followed only one syllable later. Kinaesthetic-proprioceptive reafferences—muscle-induced—could have started at the moment of the first necessary o'-muscle-innervation impetus (possibly only a few milliseconds after this impetus, as soon as a certain threshold of the voltage has been reached, that is, about 300 ms before the very resulting [o]-utterance). These muscle-induced reafferences could have initiated the speaker's feeling of lip-rounding, thus indicating to him that this feeling occurrs in the false prearticulatory moment, namely too early and instead of the right reafference for his intended sound [a]. What some authors call *Prearticulatory Editing* will be this (system-inherent) initiation of *reafferences*—the literature on neuro-physiological ERP measurements is very extensive.

(A speech error in the same sense (with a *repair on the fly*: early proprioceptive reafferences of bilabial muscle movement?) is the following one:

> *Herr Thurnher, der (B)—österreichische ˊBundeskanzler hat ...* ('Mr Thurnher, the (B)—Austrian ˊBundeskanzler has ...'; speaker Tina Mendelsohn, TV 3sat; her interview with Chief Editor Thurnher of the journal Der Falter, 2008/05/05, 19.26h / re-heard 2008/05/06, 9.11h; at the moment of the (B), the break was already *visible* by her cheeks clearly blown up for the plosion of the [b]. However, the plosion did not occur, that is, the lips did not open for the burst; therefore, the [b] did *not* get audible, and her *repair on the fly* followed rather immediately; Speech Error Corpus Wiedenmann, unpubl. part).)

The empirical data of the Perception-Centre research (i.e. 'P-Centre'; see Janker, 1995, [8] i.a.) show that these *reafferences* will be a very fact—and an important factor for understanding the perception of the syllable to be dealt with: In experiments, subjects *felt* the centre of a syllable at a rather early moment. Janker's centre of the syllable was positioned by the subjects at a rather early moment, so that the above-mentioned reafferences could possibly provide an explanation for such an *early* syllable centre. It could depend on early reafferences of a syllable's *sound preparation period* (or: Lautvorspann; see Wiedenmann, [22], passim)—before the syllable's uttered *first sound*.

5. Conclusions

The cause of involuntary sound reversals in speech presented here is only *one* of multiple causes for speech errors, but nevertheless an important one. Gestures are related to an intended *target* of speech, to the event of a sound utterance, or also to the event of a human being's or an animal's touching or grasping. The thoughts and the modelling discussed in this article on the *sound preparation period* could provide a basis for studying target-related gesture time-coordination of man (as well as of *robots*: The travel time (Laufzeit period) of *electrons* to wander from a central movement-steering electronic apparatus to the target-related robot members is to a certain extent analogous to the *sound preparation periods* of speech sounds, and comparable with the target-related movement *preparation period* of human foot-tapping).

Concerning patients, this model could also provide a basis for a better *logopaedic* work: Regarding an intended utterance, *speech training* should go from the easy to the difficult, and thus, it should start, if possible, with *ideal* words—during children's first language acquisition as well as in speech pathology—to help in adult rehabilitation.

The fairly one hundred years old *suspicion* of the Austrian linguist and speech error collector Meringer says:

> *... es ist keineswegs ohne weiteres klar und sicher, dass die Teile des Satzes in der Reihenfolge gedacht werden, in der sie im gesprochenen, grammatisch richtig angeordneten Satze zum Vorschein kommen*—(in NW's translation:) By no means and without further explanation, it is clear and sure that the parts of a sentence are thought in the same serialization as they appear in a spoken, grammatically well-formed sentence (Meringer [15], 1908, pp. 21/22; on his categorization by means of anticipations, postpositions, etc.).

The results of this investigation here give strong support to this old suspicion that one does not always think and act (articulate via muscle innervations) in the same serialization as it is audible by one's uttered sounds.

No other authors in the field of *speech production* or in the—via refferences—closely related field of *speech perception* have considered—to my knowledge—the systemic *timing aspects* presented in this *sound preparation period* model of the utterance.

Appendix 1: Figures 5-14

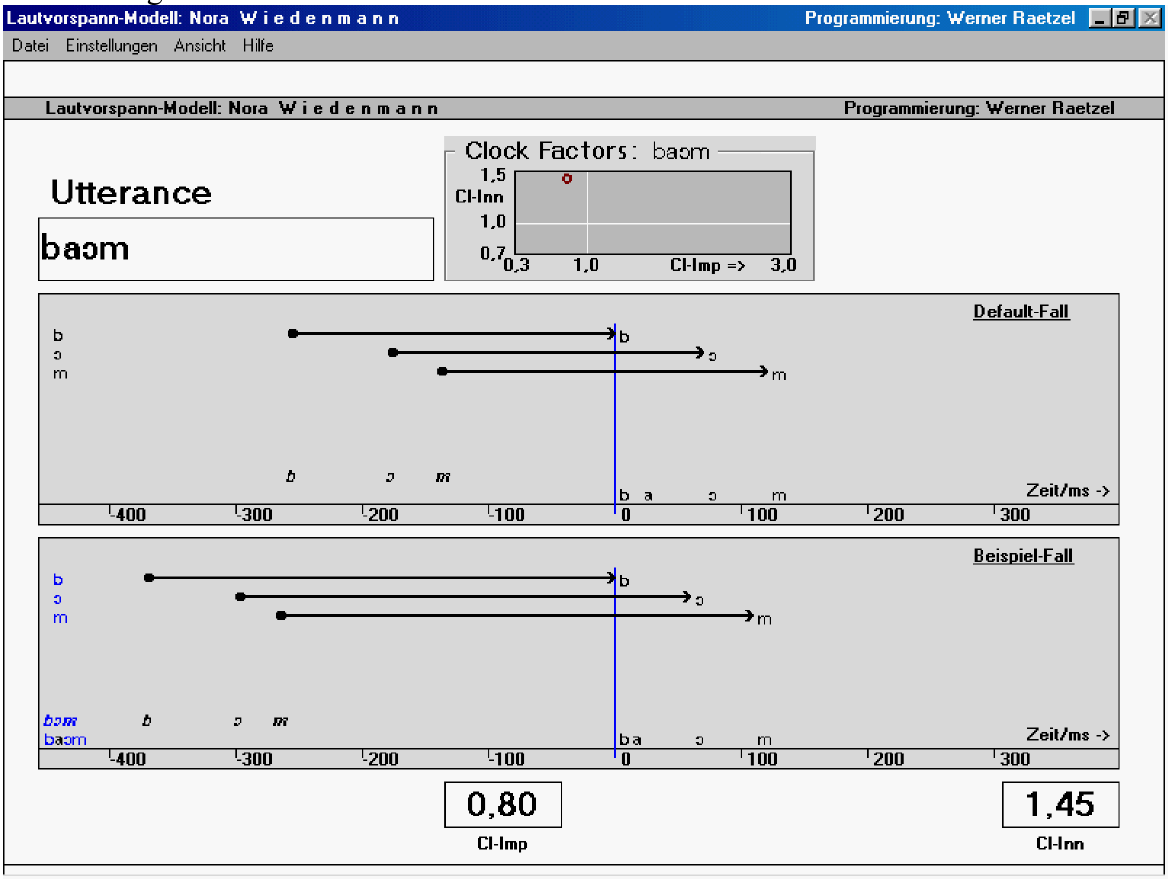

Figure 5. See above the *ideal* German word *Baum* ('tree').

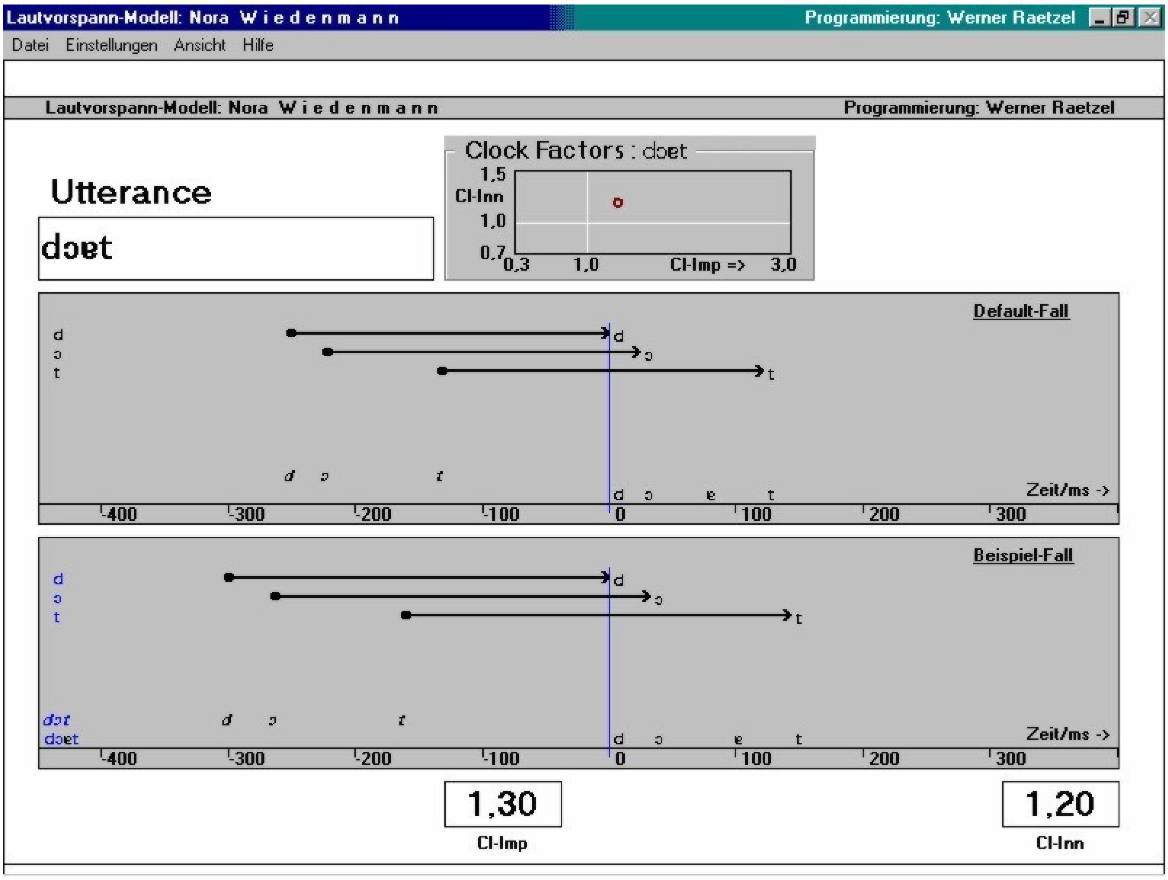

Figure 6. See above the *ideal* German word *dort* ('there').

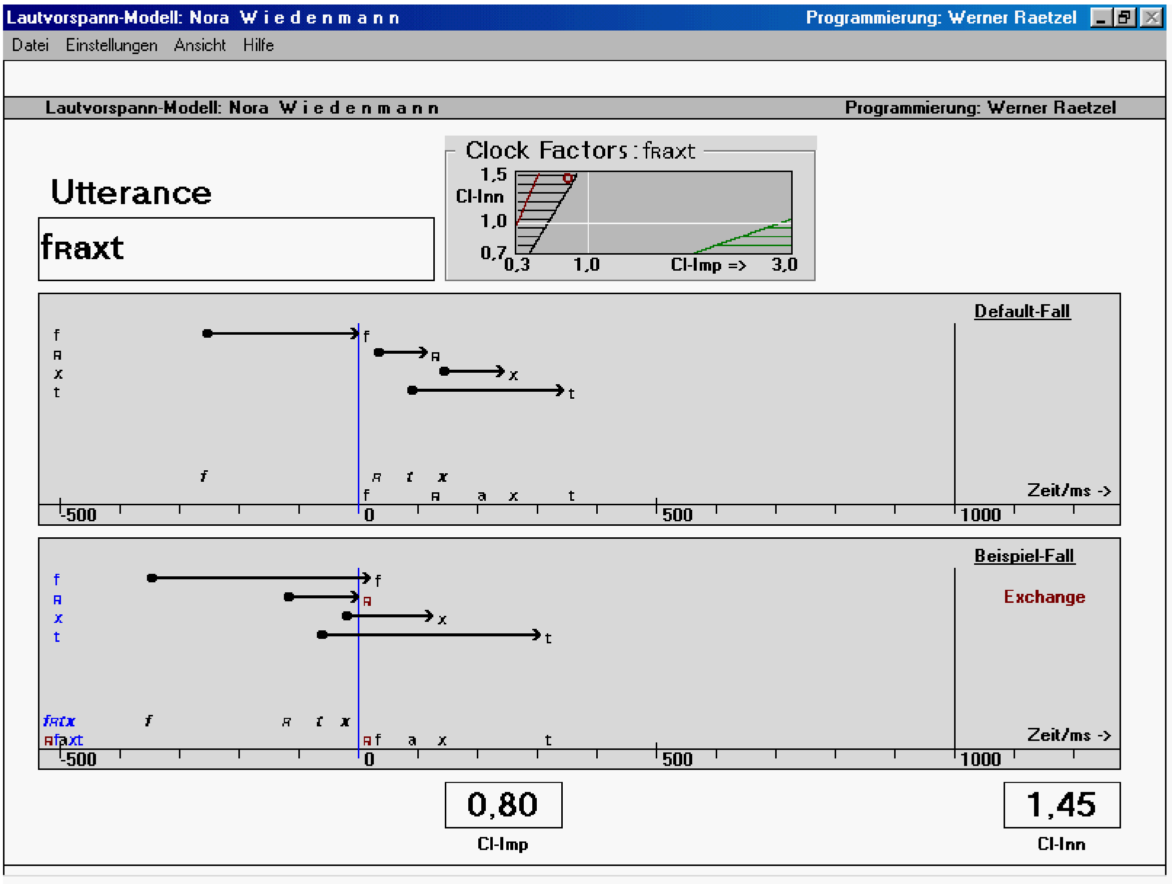

Figure 7. See above the German word *Fracht* (‘freight’), articulated with an uvular [R].

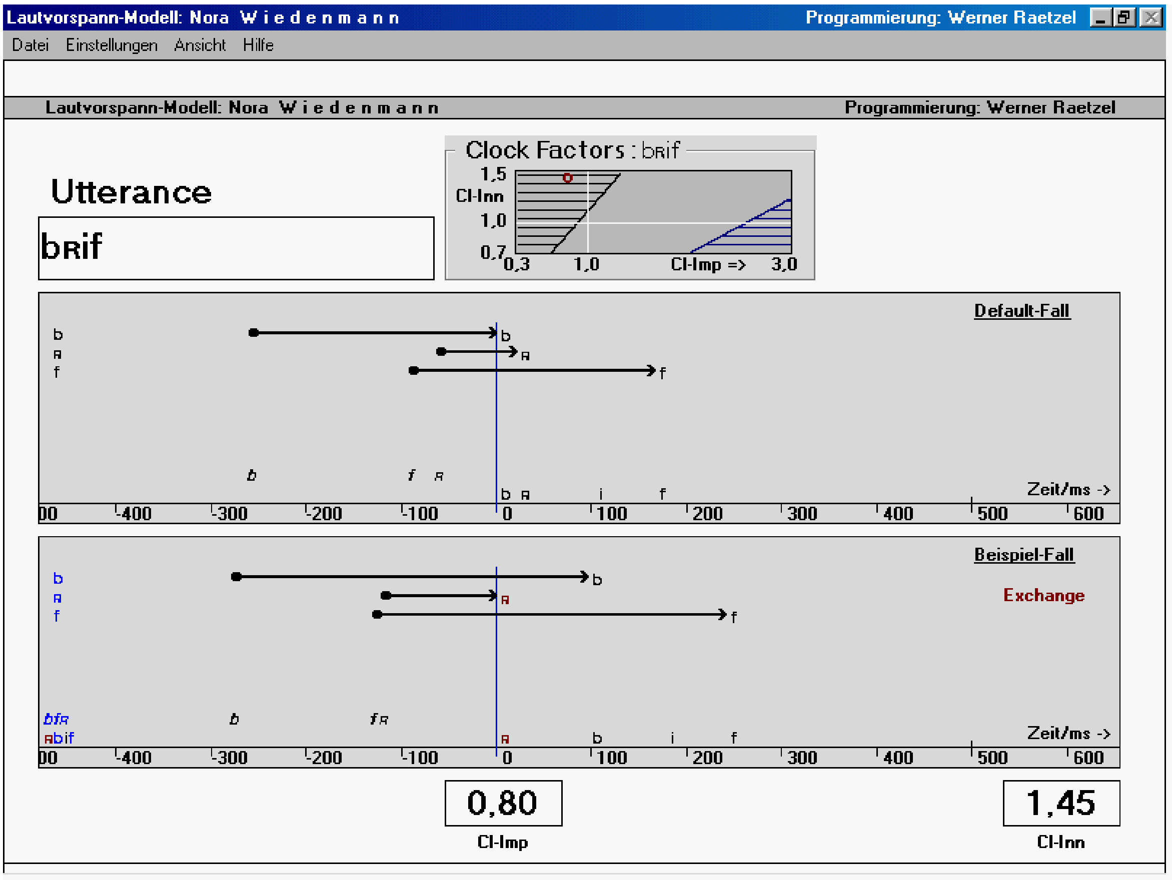

Figure 8. See above the German word *Brief* (‘letter’), articulated with an uvular [R].

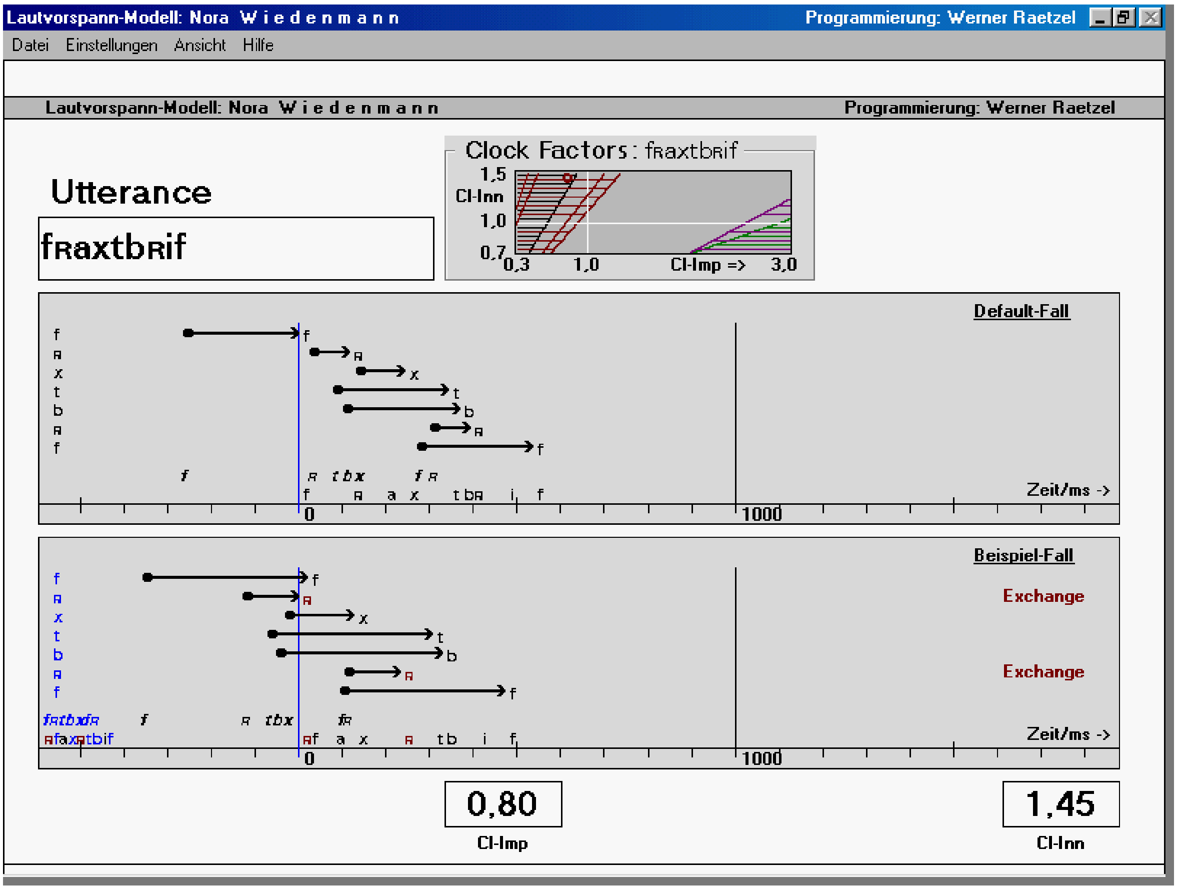

Figure 9. The German word *Frachtbrief* ('waybill'), articulated each time with an uvular [R]; sound exchanges (under the *same* (labelled) clock alterations as in Fig. 10): "Rfachrtbief"; the time differences depend on the r-gestures.

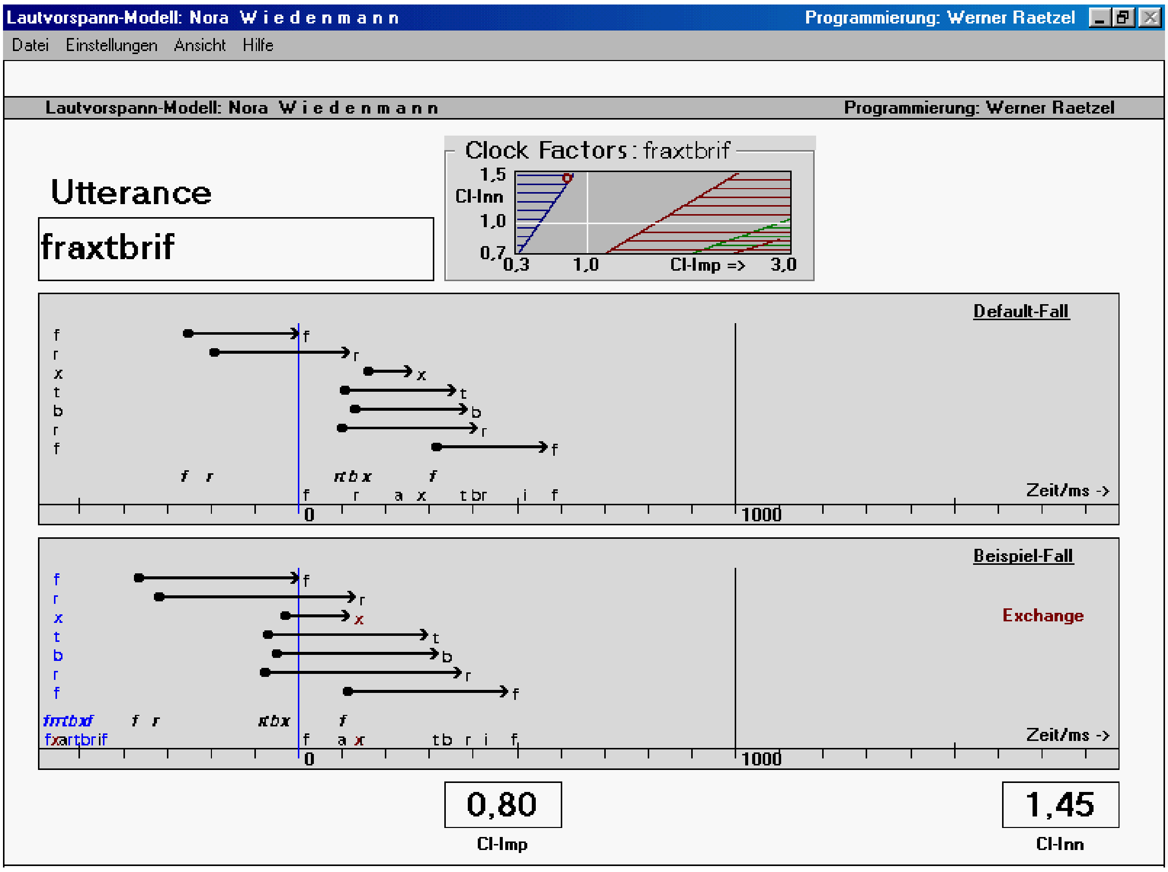

Figure 10. The German word *Frachtbrief* ('waybill'), articulated each time with a tip-alveolar [r]; sound exchange (under the *same* (labelled) clock alterations as in Fig. 9): "Fachrtbrief"; the time differences depend on the r-gestures.

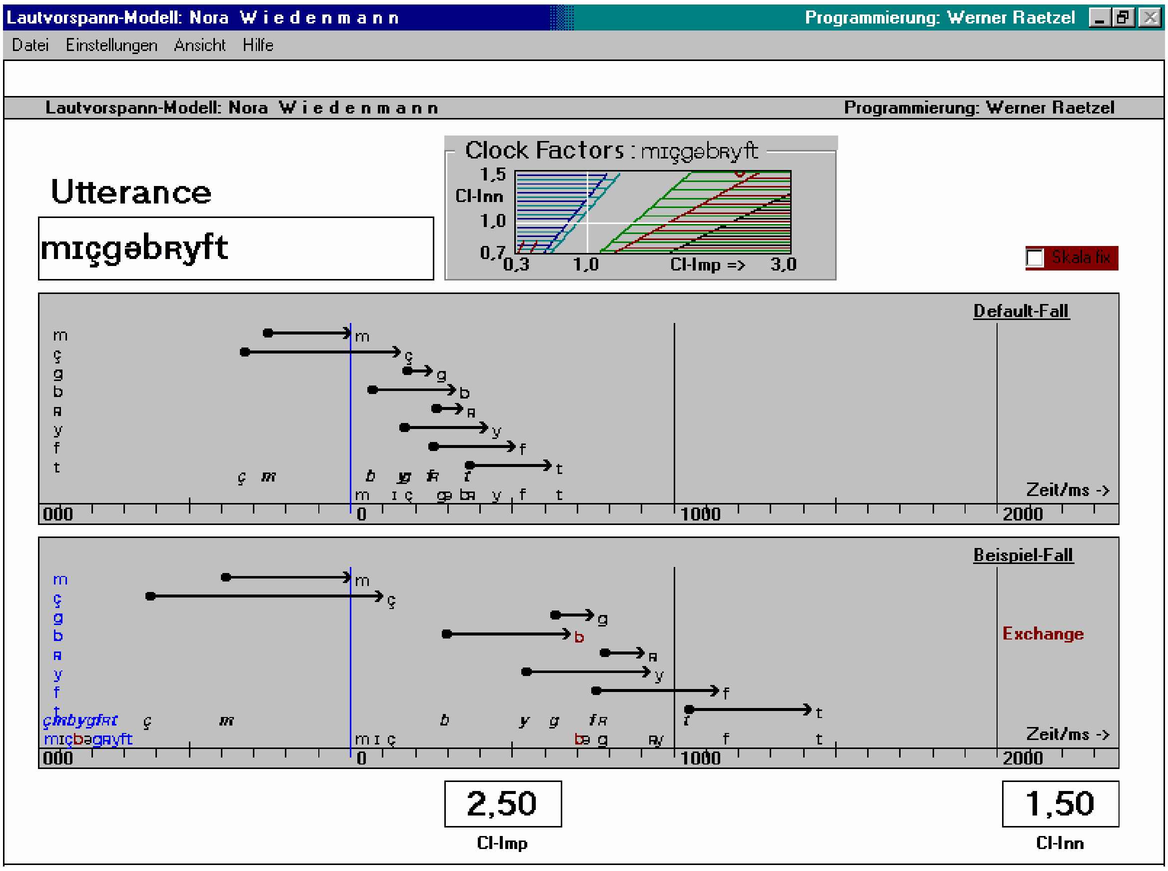

Figure 11. See above the German utterance *mich gebrüft* (canonical *mich geprüft*; 'examined me'), articulated with an uvular [R]. The sound exchange under the retardation as labelled is: "mich begrüft".

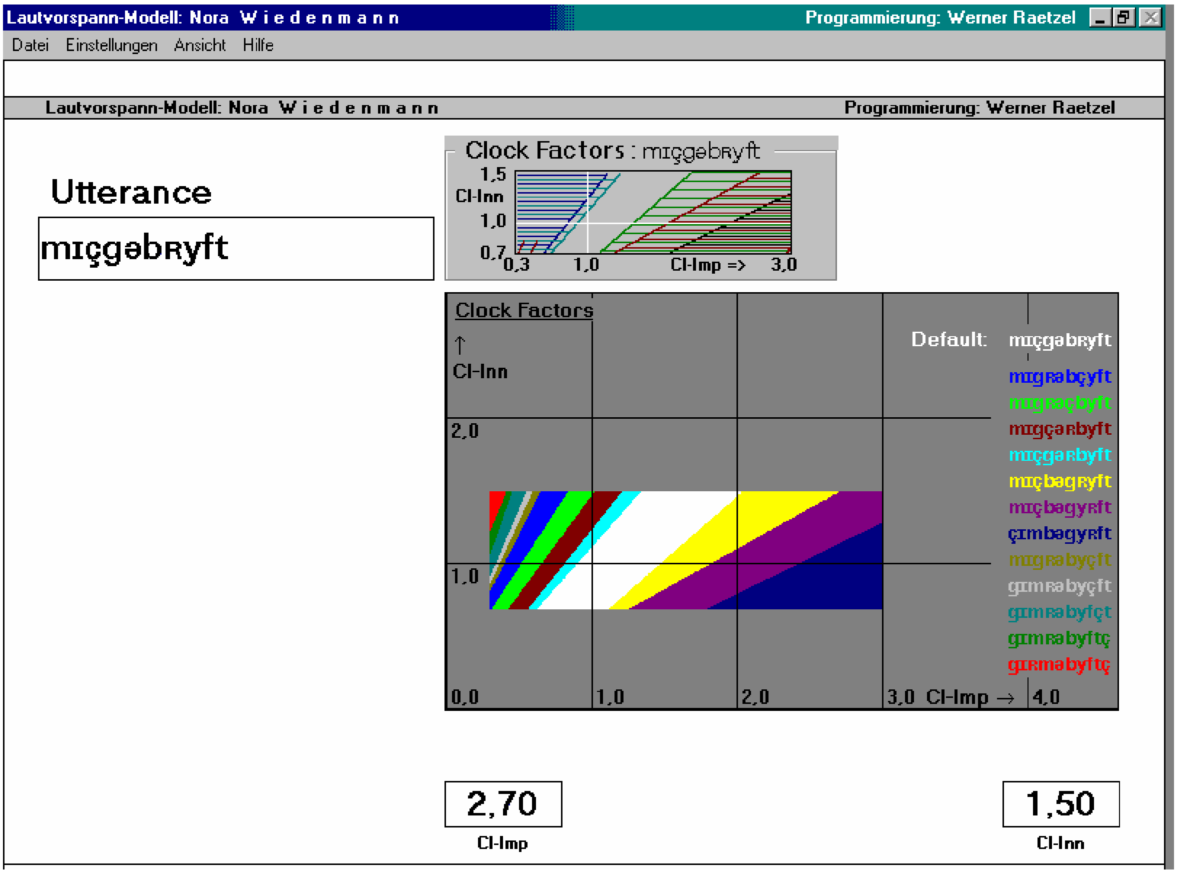

Figure 12. The German utterance *mich gebrüft* (canonical *mich geprüft*; 'examined me') and the coloured areas of possible variants of sound serializations (besides the sound exchange "mich begrüft" under the labelled retardation).

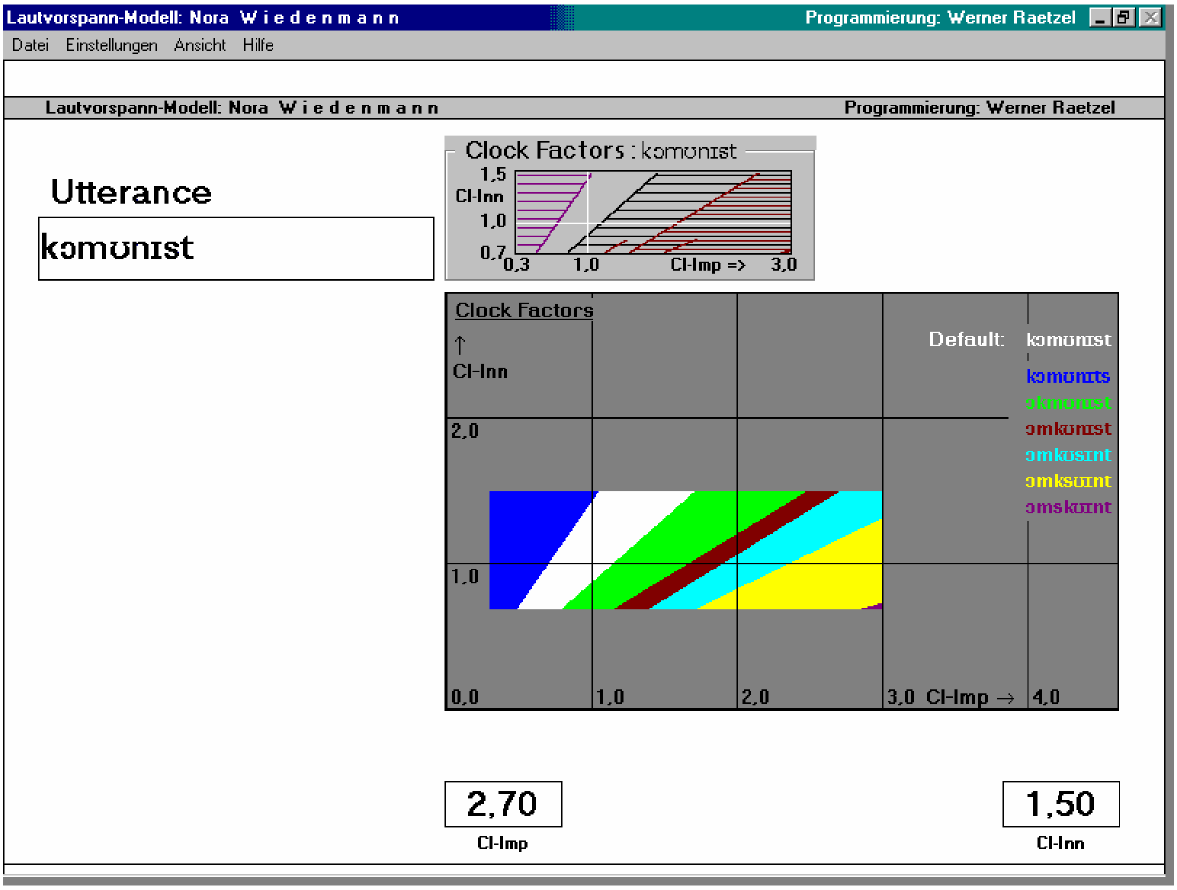

Figure 13. See above the German utterance *Kommunist* ('communist') and the coloured areas of numerous possible variants of sound serializations (besides the sound shift "Ommkunist" under the labelled retardation).

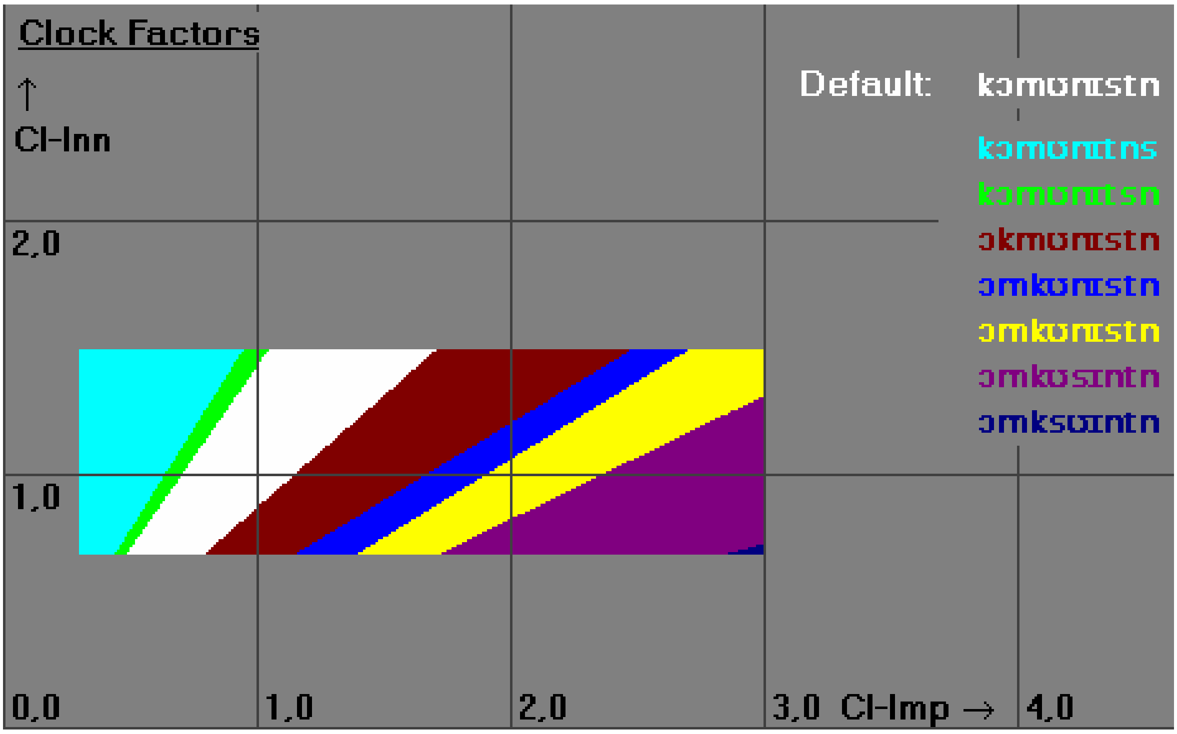

Figure 14. See above an overview of the German utterance *Kommunistn* (canonical *Kommunisten*; 'communists') and the coloured areas of numerous possible variants of sound serializations.

Appendix 2

Comments on the 'Lautvorspann' / Sound Preparation Period Programme SPEAK
(programmed in Visual Basic by Ing. Werner Raetzel)

Programme's Purpose

This programme was written for *Windows 95* and *Windows 98* and requires a screen with a minimum of 600 x 800 pixels. It simulates a simple articulatory-temporal model, which can be used to explain the genesis of certain sound speech errors (sometimes called phonological speech errors) as well as phenomena of a child's first language acquisition, and of speech pathology.

For the temporal serialization of the 'Lautvorspann' (the *sound preparation period*, i.e., the period between the earliest impetus for muscle innervations and the beginning of the corresponding sound utterance), two mechanisms are implemented in this model (see the *Comments on the Mathematics of the Wiedenmann Sound Preparation Period Model*): The first mechanism consists in a retardation / acceleration of the Clock Frequency ('Taktfrequenz') underlying the *impetus* serialization for the muscle innervations of one sound followed by its next sound in the utterance (this Clock Frequency is called *Cl-Imp*); the second mechanism consists in a retardation / acceleration of the Clock Frequency that is the basis for the duration of the sound preparation period (called *Cl-Inn*), which, in turn, is the duration of the muscle innervations.

The programme shows graphically the results of these retardations (respectively accelerations) of *Cl-Imp* (respectively *Cl-Inn*) for utterances consisting of several sounds (max. 10). In particular, the programme shows that with sufficient alteration of the Clock Factors ('Taktfrequenzen')—which are taken as *independent* from each other—it may come to a change in the serialization of the beginnings of the corresponding sounds, this change of serialization being, for example, an involuntary (that is, not intended) speech error. Extreme alterations of the Clock Factors can lead to sound serializations, in the same way as found in utterances of a child during the first language acquisition, or in the cases of speech pathology.

Usage of the Programme

<u>Preliminary Remark</u>
The area at the lower part of the screen is used to change (by moving the vertical or horizontal scroll bar) the two model parameters—the Clock Factors *Cl-Imp* and *Cl-Inn*. At its very beginning, the programme works with the default case, i.e. both Clock Factors being 1: no speech alterations.

<u>Detailed Instructions</u>

The programme has two variants, which can be selected when it is started: One variant distinguishes between vowels and consonants (because the production of vowels and consonants originates in separate areas of the brain), and the other variant does not. Only the first variant is described here.

The sounds of the utterance to be investigated are entered into the text field called *Utterance*. The programme changes the colour of the entered letters to green to indicate that it has not yet processed them.

By pressing *Enter*, the sound symbols (in some cases via the special *IPA-Symbol Keyboard* on the screen) are passed to the programme to start the simulation. The programme indicates this by changing the colour of the letters in *Utterance* to black.

In the *Utterance* field, vowels without consonantal (e.g. without labial) speech gestures are ignored by the programme as sounds, in so far as they are not shown as arrows (from the thick dot of the earliest beginning of the muscle innervation up to the tip of this arrow as the beginning of this sound's utterance). However, they are shown by the vowel symbol in the sound serialization.

When now the simulation starts automatically, the programme displays two time diagrams—one above the other (in real time relations, i.e., true to scale). The diagram on top shows the default case (both Clock Factors are 1). The bottom time diagram shows an example case *only when* the values of the Clock Factors in the displayed parameter area are changed. As the values for these Clock

Factors are preset by the programme to the default value, both time diagrams are initially identical.

Changing one of both Clock Factors automatically initiates the update of the **bottom** time diagram as an example case. The difference to the default case is well visible because the time axes of both diagrams are aligned. The utterance beginnings (the arrow tips) of the **first** sound of an entire utterance are aligned (shown by the blue vertical line). All sounds along with their sound symbols are shown closely above the time axis and vertically aligned with their corresponding arrow tips.

By specifying the option *Skala fix, nicht angehakt* (*scale is fix, not clicked*), the programme chooses the scale for the time axis based on the duration of the entire utterance. It also tries to maximize the scale, though this may make the comparison of the two diagrams somewhat difficult.

If it comes to a changed sound serialization, and, in particular, to a reversal of the utterance beginnings of two sounds, then the sound that is *too early* (that is the anticipated sound) is shown in red and supplemented by the word *Exchange* (in red). The serialization of the muscle-innervation impetusses, which may be different from the sound serialization, is shown with letters in *italics* in the left area of the example diagram (i.e. before the beginning of the entire utterance's first sound; marked by the blue vertical line).

<u>Options and Change of the Clock Factors</u>:

- Changing the Clock Factors: Scroll in known manner. (Constantly pressing on one of the two triangles produces a quick change of the factor (tenths of the factor), pressing on one of the two grey areas beside the scroll bar causes a slow change (only hundredths).)

  This way, all values can be set for a clock factor. These values must lie between the limit values shown under the scroll bar. These limit values can also be changed by *substituting* the corresponding displayed limit.

- Specifying the default value 1.00 / 1.00: Click on the rectangular labels of *Cl-Imp* and *Cl-Inn*.

- Specifying an arbitrary example value: Scroll for a coarse setting; for a fine setting, do as described above. (A single click with the mouse or touchpad instead of holding the left-button of the mouse or touchpad when clicking, enables a stepwise and slow change.)

<u>A change of the under *Utterance* selected entire sound utterance which should be illustrated as a diagram</u> leads to—after an acknowledgement of this new input via *Enter*—an update of **both** diagrams.

To show the different results of the same *Clock Factors* for different utterances—one case may result in a speech error, another may not—, the chosen *Clock Factors* are not automatically reset to the default values. The display of the default case should be done **before** a new entire utterance (as described above) is input: via click on the corresponding labels.

Display of the Diagram of the Clock Factors via menu *Ansicht* (*View*)

Buttons are located next to the diagram of the Clock Factors (with hatchings, and accessible via menu *Ansicht* (*View*), and then *Clock Factors*). By using these buttons, it is possible to return to the default case (a central +) or go to certain four example values: the 4 marked edge values (arrows in the 4 directions).

Two buttons, *S* and *R*, are located above the diagram. Selecting *S*, the programme *saves* the current setting of clock values; in the diagram, the current setting is shown by a little red ring; the programme indicates the completion of the saving by showing a blue dot in the centre of the little red ring. The blue dot remains fixed in its position in the diagram, even if the corresponding little red ring wanders in the Clock Factors' diagram (as a consequence of selecting other values). Selecting *R* (*retour*), one can return to the values of the saved blue dot.

When the user presses *S*, the programme carries out similar actions as for *R* before actually saving: Via *R*, it returns to the pair of example values which was prepared in the database of all basic data under the values *Cl-Imp / Cl-Inn for a preparatory case of example values*, for instance 1.30 / 1.20 (see the rectangular labels).

Clicking (with a mouse or touchpad) into a limited hatched area (or into a certain region in the default area) instructs the programme to calculate—for this *working point*—the corresponding values of muscle innervation (*Cl-Inn*) and muscle-innervation impetus (*Cl-Imp*) to show the time diagram of this example. This way, it is possible to choose any point in the Clock Factors' diagram to demonstrate the serializations.

Other Functions of the 'Lautvorspann' / Sound Preparation Period Programme SPEAK

Display of the sound inventory: menu *Datei* (*File*), and then menu *Inventar anzeigen* (*Display of Inventory*). *dVok* is the *duration* of a vowel, *dFore* the *duration* of the sound preparation period (i.e. the 'Lautvorspann').

Indication of the basic data: menu *Datei* (*File*), and then menu *Basisdaten* (*Basic Data*). Every single value may be replaced here via *Überschreiben* (*Replacement*) and thus changed (for instance according to *newer results of research*; the values used in this programme are based on the values of the database PHONDAT of the *Institute für Phonetik und Sprachliche Kommunikation der Ludwig-Maximilians-Universität München* (that is, of the Munich University Database of Spoken German; Christoph Draxler), of 1998; in the programme, they describe a individual time behaviour of articulation).

Representation: Vowels without consonantal (for instance, labial) speech gesture (for instance an [a]) are not shown (in this programme variant) with arrows in the diagrams. Capitalization of characters in the field *Utterance* partly results in the representation of IPA characters. Vowels including labial speech gestures (for instance an [o]) are handled mainly like consonants. All sounds are saved in tables and are ready to be used via the menu *Datei* (*File*), and then menu *Basisdaten* (*Basic Data*): Together with each sound, the tables contain the duration of its *Sound Preparation Period* ('Lautvorspann') *dFore* as well as the duration of the very sound.

Measuring Time Distances: As a change of the Clock Factors' default values implies a change of the time distances in the entire net of sounds, the programme was designed to enable the *measurement* of every arbitrary time distance (in milliseconds) in the time diagrams. For this purpose, the cursor is positioned at the beginning of the length to be measured and the cursor is dragged (that is, by holding the left button of a mouse or a touchpad pressed) up to the endpoint of the length to be measured. After the left button of the pointing device is released, the distance between the two cursor points (in milliseconds) is shown.

Comments on the Calculations of the Wiedenmann Sound Preparation Period Model

To articulate every speech sound of an entire utterance with the right beginning (beginning of the sound utterance, t_{utt}), the corresponding muscle-innervation impetus time (muscle-innervation impetus, t_{imp}) has to be set in the right manner, namely *timely*. (This is also difficult for a child at the very beginning of the first language acquisition of sounds—respectively in cases of speech pathology, when the time net with a probably changed Clock Frequency has to be learnt anew, respectively has to be trained.

The time distance / duration (for a certain sound) between the muscle-innervation impetus and the corresponding beginning of the sound utterance is called 'Lautvorspann' (*sound preparation period*, *Spp*).

Spp = Sound preparation period = sound beginning - muscle-innervation impetus = $t_{utt} - t_{imp}$ (1)

The time distance / duration between the muscle-innervation impetus for a sound and the impetus for the following sound is not only dependent on the articulatory-temporal qualities of a certain sound itself but also on the articulatory-temporal qualities of its following sound.

$t_{imp}(i+1) - t_{imp}(i) = Dist_{imp}$(sound, following sound) (2)

Two remarks on the values $Dist_{imp}$(sound, following sound):
Firstly, they cannot be determined directly, but can only be calculated from

$Dist_{imp}$(sound, following sound) = $Dist_{utt}$(sound, following sound) + Spp (following sound) - Spp (sound) (3)

Secondly—according to the hypothesis shown here—, $Dist_{imp}$(sound, following sound) is dependent on the variable 'Taktfrequenz' (Clock Frequency; Clock Factor). Note that, compared to its default value, this frequency may be retarded (factor > 1) or accelerated (factor < 1).

Naturally, this has consequences—and even very complex ones—for the values of $Dist_{utt}$(sound, following sound). Therefore, also the values set into equation (3) for $Dist_{utt}$(sound, following sound) always can only be understood as values concerning the 'Taktfrequenz' (Clock Frequency) for the default case. Equation (3) has to be stated more precisely as follows:

$Dist_{imp,\ Default}$(sound, follng.sound) = $Dist_{utt,\ Default}$(sound, follng.sound) + Spp(follng.sound) - Spp(sound) (3')

Moreover, the following equation applies:

$Dist_{imp}$(sound, following sound) = Cl-Imp x Default-$Dist_{imp}$, Default(sound, following sound) (4),

with Cl-Imp as the Clock Factor corresponding to the sequence of muscle-innervation impetusses.

The **purpose of this programme** is to depict the effects of changed clock factors.
In contrast to $Dist_{utt,\ Default}$, it is possible that—as shows equation (3')—the values of

$Dist_{imp,\ Default}$(sound, following sound)

become negative. That is the case when

$Dist_{imp,\ Default}(S_i, S_{i+1}) < 0$ for $Spp(S_{i+1}) > Spp(S_i) + Dist_{utt,\ Default}(S_i, S_{i+1})$ (5),

with S_i for *sound* and S_{i+1} for *following sound*.
In such a case, the serialization of the respective beginnings of the sound utterance can change, provided that the Clock Frequency is sufficiently retarded (factor > 1).

Appendix 3

Empirical data: Speech errors of bilabial versus dorsovelar plosives

(´ is always word accent; all translations by N.W.)

Cases cited in: **Wiedenmann, Nora: A corpus of German speech errors.** Forschungsber. Inst. Phonet. Sprachl. Kommun. Univ. München FIPKM *30* Suppl.: S1-S77 (1992a).

p. S44, speaker Christine M.: *Bekannte, die sind jetzt irgendwo am Nordpa – Nord – Nordkap* (‘Good friends, they are now somewhere near the North Pole’; reversal, broken off, plus multiple repair on the fly; nearly: *Nordpak*; German *Pa(c)k*: to remind of *Pack* (‘stack’ / ‘riffraff’), or of *packen* (‘to pack’); *-pak* a metathesis crossing the syllable position: from final to initial position; sentence end)

p. S44, speaker Steffi O.: *Be´girge – Ge´birge* (‘mountains’; *ge-* and *be-* frequent German morphemes; *-birge* ref. to *Berg* (‘mountain’); *-girge*: German nonsense morpheme; reversal of [b] versus [g])

p. S44, speaker Creutzberg: *indem man dann das Ge´briffsgebäude – Be´griffsgebäude aufbauen darf.* (‘while then one may construct the formation of the concept’; reversal, instead of intended *Be´griff* (‘concept’); *ge-* and *be-* frequent German morphemes; *-briff*: a German nonsense morpheme)

Cases cited in: **Wiedenmann, Nora: Versprecher: Phänomene und Daten. Mit Materialien auf Diskette** (Wissenschaftsverlag Edition Praesens, Wien 1998b):

p. 35, speaker Meringer: ... *der Bru – der Gruber* (‘the Bru—the Gruber’; reversal, broken off, plus repair on the fly; nearly: *Bruger* (nonsense word); *Gruber* is a German surname)

p. 36, speaker Meringer: *geb – begreiflich* (‘stun—understandable’; reversal, broken off, plus repair on the fly; nearly: *gebreiflich* (nonsense word); *ge-* and *be-* frequent German morphemes; -*breif-*: German nonsense morpheme)

p. 165, speaker M.: *So! Jetzt werd ich mal meine ´Schuhe gep – be´gutachten!* (‘Now, I will take a look at my shoes’; reversal, broken off, plus repair on the fly; nearly: *gebutachten*, i.e. nonsense compound because of the German nonsense morpheme *but* [bu:t], instead of *gut* ([gu:t]; ‘good’); [p] instead of [b] because of break: voice ended early; *ge-* and *be-* frequent German morphemes; sentence end)

p. 165, speaker M.: *´Was!? ´Pappkap – ?* (‘´What!? card(board) cap—?’; contact metathesis of [p] versus [k], instead of *´Packpapier* (‘wrapping paper’), broken off, not corrected; *Pappkap* is a metathesis crossing the syllable position of *Packpapier*: from initial to final position)

p. 226, speaker N.: ... *sei eine normale Blüh – Glühbirne* (‘be a normal bight—light bulb’; nearly the German nonsense compound *Blühgirne*: *Blüh...:* ‘bloom’; *Girne*: German nonsense word; reversal, broken off, plus repair on the fly; sentence end)

p. 226, speaker N.: *Das Be – Gepäck* (‘The gu—luggage’; reversal, broken off, plus repair on the fly; nearly the nonsense word *Bekäck* (the written *-käck* to remind of: *Kacke* (‘crap’); *ge-* and *be-* frequent German morphemes; for an illustration of this utterance’s *sound preparation periods* see the figure on p. 15 in Wiedenmann, N.: Versprecher: Dissimilation von Konsonanten. Sprachproduktion unter spatio-temporalem Aspekt (Niemeyer Verlag, Tübingen 1999))

p. 226, speaker N.: *Da gab‘s noch keine Ba – Gastarbeiter* (‘At that time, there were no foreign workers’; reversal, broken off, plus repair on the fly; nearly the German nonsense compound: *Bas*t (‘raffia’) plus the nonsense word *-argeiter*; sentence end)

p. 226, speaker N.: ... *ganz unbe – ungeplant* (‘really unplanned’; reversal, broken off, plus repair on the fly; nearly the nonsense word: *unbeklant*; *be-* und *ge-* frequent German morphemes; sentence end)

p. 230, speaker N.: *Ist mir kein Ge´briff!* (‘Doesn’t mean anything to me!’; reversal, instead of intended *Be´griff* (‘concept’); *ge-* and *be-* frequent German morphemes; *-briff*: a German nonsense morpheme; sentence end)

p. 230, speaker N.: *Wenn Du Caspal – Pas´cal kennst, ...* (‘When you know Caspal—Pas´cal then ...’; reversal of [k] versus [p] resulting in the German nonsense word *Caspal*)

p. 230, speaker N.: *..., parce qu’il a voulu m’exclip – m’expliquer ...* (‘because he wanted to explain to me ...’; reversal of [k] versus [p], broken off before an *excliper*, plus repair on the fly; this error (by the German speaker N.) resulting in the English ‘clip’ as used in French for ‘ear clip’ / ‘film clip’)

p. 230, speaker N.: *ganz schönes Pak – Kapi´tal!* (‘really a pretty pack—capital!’; reversal, broken off, plus repair on the fly; German *Pa(c)k*: to remind of *Pack* (‘stack’ / ‘riffraff’), or of *packen* (‘to pack’); but [pak] nearly would have been resulted in the German nonsense word *Pakital*; sentence end)

p. 230, speaker N.: *... ´Pängk – ´Camping* (‘Pang[k]—Camping’; reversal, broken off, plus repair on the fly; nearly (the German nonsense word) *Pängkim*?; the German spelling *-ng-* means the dorsovelar nasal consonant)

p. 259, speaker N.: *Da musse* [i.e.: *musst du*] *nur die Kortop – die Portokosten einkalkulieren* (‘In this case, you’ve only to reckon on the postage (*Porto-*) expenses (*Kosten*)’; reversal, broken off, plus repair on the fly; the German nonsense compound *Kortoposten*: the German nonsense word *Korto-*, and nearly continued by *-Posten*, i.e.: ‘job’ / ‘guard’)

Cases cited in: **Rossi, Mario, Peter-Defare, Évelyne: Les lapsus ou comment notre fourche a langué** [16] (Presses Universitaire de France, Paris 1998):

p. 120, No. 310-2499: *il veut faire de son pays une Cologne — une Pologne qui intègrera l’OTAN* (‘he wants to make out of his country a Cologne—a Poland which will integrate the OTAN’; French *Cologne* for the German city *Köln*; nearly the reversal of [k] versus [p] *Cologne pi* because of *Pologne qui*?)

p. 149, No. 741-339: *c’est la fin d’une écope — d’une époque* (‘that’s the end of an eecop—of an epoch’; reversal of [k] versus [p] resulting in the French nonsense word *écope*; French *époque* with a final schwa (if not: *écope* as metathesis crossing the syllable position, from final to initial position; sentence end; the French error categorization term is *interversion* for reversal / metathesis)

Cases cited in: **del Viso, Susana [3]: Errores espontáneos del habla y producción del lenguaje. Apendice: Corpus de errores espontáneos del habla (esp. castellano).** Tesis doctoral. Facultad de Psicología, Departamento de Psicología Básica, Procesos Básicos (Universidad Complutense de Madrid, Madrid 1990).

p. A-80, 12: *¿Cuándo vais a hacer las campartas?* (‘When will you finish with the *campartas*?’; reversal of [k] versus [p]: intended: *pancartas*; coarticulation before [p] in the Spanish nonsense word *campartas*)

p. A-80, 15: *relata una camparta para la libertad* (‘reports a *camparta* for the liberty’; reversal of [k] versus [p]: intended: *pancarta*; coarticulation before [p] in the Spanish nonsense word *camparta*)

p. A-82, 04: *Ese agua ya no es repuquerable después de un ... aproximadamente* (‘This water isn’t any longer reclaimable since ..., approximately’; reversal of [p] versus [k]: intended: *recuperable* (‘reclaimable’); *repuquerable*: Spanish nonsense word)

p. A-82, 13: *excedente de puco – de cupo* (‘quotation surplus’; reversal of [p] versus [k]; *puco*: Spanish nonsense word)

p. A-83, 02: *Gary Pooquer* (instead of *Gary Cooper*; reversal of [p] versus [k]; *Pooquer* ([pu:ker]): Spanish nonsense word)

p. A-84, 15: *Hay muy poca gente impac – incapaz de ...* ('There's very few people which is unable to ...'; reversal of [p] versus [k]; coarticulation before [p] in the Spanish nonsense word *impac...*)

p. A-84, 16: *Para copas, ya me caben en la carpeta* ('For only few, (they) will just fit into the briefcase'; intended: *pocas* ('few'); reversal of [k] versus [p]: *copas*: 'cups' / 'glasses')

p. A-84, 17: *La verdad es que son un copo – un poco incómodos* ('The truth is that they are a little awkward'; reversal of [k] versus [p]: *copo*: 'flake' / 'tuft')

p. A-88, 04: *Cuesta lo mismo llamar de allá p'acá que de apá c'allá* ('That's the same: to say *de allá p'acá* as to say *de acá p'allá*'; intended: *de allá p'acá que de acá p'allá* ('from there till at home as (to say) from at home till there'); sentence end; reversal of [p] versus [k]: *apá c'allá* instead of the intended *acá p'allá*)

Cases cited in: **Fromkin, Victoria A.: Speech Errors as Linguistic Evidence** [5] (Janua Linguarum, Series Maior, 77. Studia memoriae Nicolai van Wijk dedicata); Appendix 'A Sample of Speech Errors', p. 243-269 (Mouton, The Hague 1973).

p. 254: *patterkiller* (Engl. nonsense compound because of *patter-*; reversal of [p] versus [k] instead of intended *catterpillar*; written by Fromkin to show the speaker's pronunciation: *patter-* / *catter-* for the meant *caterpillar*)

p. 254: *canpakes* (Engl. nonsense compound because of *-pakes*; reversal of [k] versus [p] instead of intended *pancakes*)

Acknowledgements

I thank Ing. Werner Raetzel, Munich, for programming my Lautvorspann Model (in Visual Basic; see my short description of his programme SPEAK in Appendix 2) following my proposals and ideas as well as for creating the layout of all the figures in the body of the article and in the Appendix 1. In addition, I would also like to thank Alwin Fill and Francisco Nunes who helped me to improve my English style.

I also particularly thank the anglicist Prof. Bernhard Kettemann, Graz (Austria), and his wife Ilse for providing me with their speech error collection, which was one of the fundamentals of my research in speech errors during more than a decade, and which was important, too, for this investigation. I also thank PD Christoph Draxler (Institut für Phonetik und Sprachliche Kommunikation, now Phonetik und Sprachverarbeitung (Speech Science), Munich University) for giving me the necessary parts of the PHONDAT Database of Spoken German to put in the *context-depending default sound durations*. Finally yet importantly, I thank Prof. Mario Rossi, Aix-en-Provence (France), for his fruitful comments on an earlier version of this paper, and Prof. Adrian Simpson and Prof. Valentino Braitenberg for their encouraging words on this version.

References

[1] Browman, C.P., and Goldstein, L. (1987). Tiers in articulatory phonology, with some implications for casual speech. *Haskins Laboratory Status Report Speech Research*, SR-92, 1-30.

[2] Dell, G.S., and Reich, P.A. (1980). Toward a unified model of slips of the tongue. In V.A. Fromkin (Ed.), *Errors in linguistic performance. Slips of the tongue, ear, pen, and hand* (pp. 273-286). London: Academic Press.

[3] del Viso, S. (1990). Errores espontáneos del habla y producción del lenguaje. Appendix: Corpus de errores espontáneos del habla (español castellano) (pp. A1-A245). Tesis doctoral. Facultad de Psicología, Departamento de Psicología Básica, Procesos Básicos. Madrid: Universidad Complutense de Madrid.

[4] Elsen, H. (1991). *Erstspracherwerb. Der Erwerb des deutschen Lautsystems.* Wiesbaden: Deutscher Universitätsverlag.

[5] Fromkin, V.A. (Ed., 1973). *Speech Errors as Linguistic Evidence* (Janua Linguarum, Series Maior, 77. Studia memoriae Nicolai van Wijk dedicata), Appendix: *A Sample of Speech Errors* (pp. 243-269). The Hague: Mouton.

[6] Goldstein, L. (1989). On the domain of the quantal theory. *Journal of Phonetics*, 17, 91-97.

[7] Hardcastle, W.J. (1976). *Physiology of speech production. An introduction for speech scientists.* London: Academic Press.

[8] Janker, P.M. (1995). Sprechrhythmus, Silbe, Ereignis. Eine experimentalphonetische Untersuchung zu den psychoakustisch relevanten Parametern zur rhythmischen Gliederung sprechsprachlicher Äußerungen. *Forschungsberichte Institut für Phonetik und Sprachliche Kommunikation Universität München (FIPKM)*, 33, 1-259.

[9] Kelley, J.E., Jr. (1961). Critical-path planning and scheduling: Mathematical basis. *Operations Research*, 9(3), 296-320.

[10] Kühnert, B. (1989). Das Timing der artikulatorischen Gesten bei der Produktion rhythmisch gleichmäßiger Sequenzen. *Forschungsberichte Institut für Phonetik und Sprachliche Kommunikation Universität München (FIPKM)*, 27, 141-209.

[11] Laver, J.D. (1977). Neurolinguistic aspects of speech production. In Gutknecht, C. (Ed.), *Grundbegriffe und Hauptströmungen in der Linguistik* (pp. 142-155). Hamburg: Hoffmann und Campe.

[12] Leonard, L.B., and McGregor, K. (1991). Unusual phonological patterns and their underlying representations: A case study. *Journal of Child Language*, 18, 261-271.

[13] Löfqvist, A., and Yoshioka, H. (1981). Interarticular programming in obstruent production. *Phonetica*, 38, 21-34.

[14] MacKay, D.G. (1969). Forward and backward masking in motor systems. *Kybernetik*, 2, 57-64.

[15] Meringer, R. (1908). *Aus dem Leben der Sprache: Versprechen, Kindersprache, Nachahmungstrieb. Festschrift der k. k. Karl-Franzens-Universität Graz aus Anlaß der Jahresfeier am 15. XI. 1906.* Berlin: B. Behr's Verlag.

[16] Rossi, M., and Peter-Defare, É. (1998). *Les lapsus ou comment notre fourche a langué*. Paris: Presses Universitaire de France.

[17] Shattuck-Hufnagel, S., and Klatt, D.H. (1979). The limited use of distinctive features and markedness in speech production: Evidence from speech error data. *Journal of Verbal Learning and Verbal Behavior*, 18, 41-55.

[18] Wiedenmann, N. (1992). A corpus of German speech errors. *Forschungsberichte Institut für Phonetik und Sprachliche Kommunikation Universität München (FIPKM)*, 30 Suppl., S1-S77.

[19] Wiedenmann, N. (1992). *Versprecher und die Versuche zu ihrer Erklärung. Ein Literaturüberblick*. In Series FOKUS, Kühlwein, W., and Vater, H. (Eds.), *Linguistisch-Philologische Studien*, Nr. 9. Trier: WVT Wissenschaftlicher Verlag Trier.

[20] Wiedenmann, N. (1998). Versprecherdaten und Konsonantenvorkommenshäufigkeiten. Gesprochene Sprache aus Versprechersammlungen im Gegensatz zu geschriebener Sprache der KAEDING-Zählung. *Sprache & Sprachen*, 22, 16-39.

[21] Wiedenmann, N. (1998). *Versprecher: Phänomene und Daten. Mit Materialien auf Diskette*. Wien: Wissenschaftsverlag Edition Praesens.

[22] Wiedenmann, N. (1999). *Versprecher: Dissimilation von Konsonanten. Sprachproduktion unter spatio-temporalem Aspekt*. Tübingen: Niemeyer Verlag.

[23] Wille, H., Gewald, K., and Weber, H.D. (1972), *Netzplantechnik. Methoden zur Planung und Überwachung von Projekten. Vol. 1. Zeitplanung* (3rd ed.; 1st ed. 1966). München / Wien: R. Oldenbourg.