Articulatory variability of clutterers

Mariam Hartinger
Zentrum für Allgemeine Sprachwissenschaft (ZAS), Berlin, Germany

Christine Mooshammer
Christian-Albrechts-Universität Kiel, Germamy

In order to investigate the articulatory processes of the hasty and mumbled speech of clutterers, the kinematic variability was analysed by means of electromagnetic midsagittal articulography (EMMA). In contrast to stutterers, clutterers improve their intelligibility by concentrating on their speech task. Variability is an important criterion in comparable studies of stuttering and is discussed in terms of the stability of the speech motor system. The aim of the current study was to analyse the spatial and temporal variability in the speech of three clutterers and three control speakers. All speakers were native speakers of German. The speech material consisted of repetitive CV-syllables and foreign words, because clutterers have the most severe problems with long words which have a complex syllable structure. The results showed a higher quotient of variation for clutterers in the foreign word production. For the syllable repetition task, no significant differences between clutterers and controls were found. The extremely large and variable displacements were interpreted as a strategy that helps clutterers to improve the intelligibility of their speech.

1   Introduction

Speech production underlies a certain amount of natural variability, a property that it shares with all other human motor behaviour. Hence, a specific motor task cannot be replicated in an absolutely identical way. This kind of token-to-token variability has been attributed to neural noise corrupting the control signals (Perkell & Nelson 1985). Knowledge about normal variability of speech motor abilities is limited and the cut-points between normal and so-called pathological motor behaviour is arbitrary (van Lieshout et al. 2004). In speech therapy it is common to classify the behaviour of speakers as pathological if it is different. This is the reason why the variability in the speech production of stutterers "… has always been a hot topic in research …" (van Lieshout et al. 2004, 329).

In this current study, our aim is to investigate the spatial and temporal kinematic variability of the fluency disorder cluttering. Up to now no unique definition exists for this fluency disorder and the transition between "normal" and "pathological" is not fixed. For example, the speech rate of clutterers is described as extraordinarily fast and hasty. Normal speakers, however, are also capable of speaking very quickly but their speech is usually still intelligible and not classified as pathological. Fast speakers do not exhibit as many speech errors as clutterers (Sick 2000). In the literature, there are hardly indications of how many speech errors like elisions, substitutions etc. are "normal" and how many are typical for cluttering. For normal speakers Levelt (1991) describes 1 error per 1000 words.

The symptoms of cluttering also seem to be rather divers. Scherer (2003) denotes it individual and situational variability. In particular, in emotional situations the symptoms appear very clearly. However, the more clutterers concentrate on their own speech, the more intelligible their utterances become. The improvement of fluency by concentration is an important criterion to differentiate between clutterers and stutterers. Up to now there is no unique definition of cluttering.

As was pointed out by Caruso et al. (1988) a higher variability in stuttering reflects an inherently unstable situation, that is based on underlying neuromotor control problems. Up to now the etiology of cluttering is non-specific. Impairments of the central nervous system (Wirth 1994) were as well discussed as the disability of formulating speech (Braun 1999) or as the auditory processing deficit (Molt 1996). Following the conclusion of van Riper (1990) it is possible that not only stuttering but also cluttering is caused by tiny lags and disruptions in the timing of the speech production process. The reasons for cluttering, as well as the intra-individually inconstant pathology and experimental studies on stuttering, discussed above lead to the expectation of a higher articulatory variability for clutterers as compared to controls. Therefore, our hypothesis is that clutterers exhibit higher temporal and spatial token-to-token variability of articulatory gestures as compared to control speakers.

2   Methods

Electromagnetic midsagittal articulography (EMMA) was used for recording the articulatory abilities of three clutterers and three controls. In comparison with the X-ray technology EMMA is a medically harmless experimental method which allows the measurement of two-dimensional articulatory movements simultaneously with a high spatial and temporal resolution.

In the field of speech pathology, EMMA has been used to investigate the fluency disorder stuttering (e.g. van Lieshout et al. 1993, Ward 1997, McClean

& Runyan 2000, McClean et al. 2004), as well as analyses of swallowing disorders (Kretschmer 1996) and motor disorders like dysarthria (Jaeger et al. 2000) and aphasia (Katz et al. 1990).

2.1   Subjects

Six German native speakers took part in this experiment. There were two male and one female clutterers and two male and one female normal speakers. The clutterers were diagnosed during speech therapy and had no other speech disorders. The subjects were between 21 and 36 years old. In the following the subjects will be identified with initials that refer to the group (P = clutterer[1], N = Normal speaker) and the gender (M = male, W = female) they belong to.

2.2   Material

The subjects were instructed to produce two different tasks: repetitive CV-sequences and test words with carrier phrases. The first task was to repeat simple CV-sequences as fast and intelligible as possible within a 10 sec interval. The syllables consisted of /pa/, /ta/ and /ka/; and each syllable train was repeated 2 times.

For the second task foreign words were embedded in the carrier sentence „Sage ... bitte" (say ... please) in order to elicit more natural utterances. The words consisted of a minimum of 5 and a maximum of 8 syllables. Foreign words were chosen because clutterers have special problems with longer words with complex syllable structures. All words contained the sequence /nali/ in which the final vowel /ɪ/ or /i/ was either stressed and lax as in /dimɛnzionaˈlɪstɪʃ/ or unstressed and tense as in /dimɛnzionaliˈziːrən/ (bold letters indicate lexically stressed vowels). The vowel /a/ was realised as pre-stressed 1 (p1) or pre-stressed 2 (p2). Each of the 5 word pairs appeared in the p1- and the p2-condition.

Each word was repeated 10 times in randomised order. Overall each speaker produced 100 sentences (50 x [naˈlɪ], p1, 50 x [nali], p2). Because during the recording session of speaker NM1 one sensor came off, only 35 repetitions of each word were used for further analysis.

The subjects were asked to speak as fast and at the same time as intelligible as possible. The aim of this instruction was to check how sensitive clutterers are to the perception of their own intelligibility.

---

[1]   in German = "Polterer"
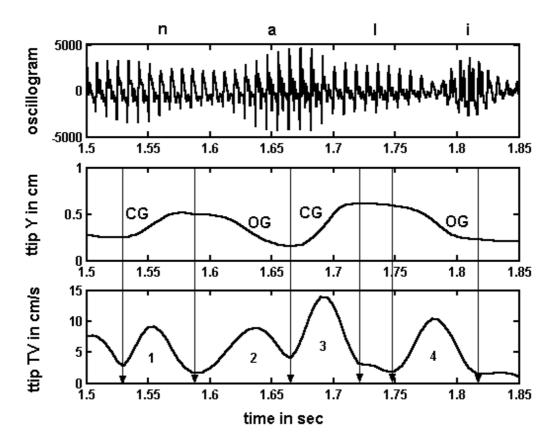
2.3 EMMA recording

For the present study kinematic data were recorded by means of the Articulograph AG 100 (Carstens Medizinelektronik Göttingen 1992). During the recording the speaker wears a plexiglass helmet to which three transmitter coils are attached in an exactly defined distance to each other. They generate electromagnetic fields with three different frequencies. A sensor coil moving in this electromagnetic field induces a current, the strength of which is approximately inversely proportional to the distance between the transmitter and the receiver coils. with an accuracy of about a couple of tenth millimeters.

Sensors were attached midsagittally to the jaw, the lower lip and four tongue positions with a distance of approximately 1cm to each other. The tongue tip sensor was placed about 1cm behind the tongue tip. Reference sensors were fixed to the gums of the upper incisors and to the nasion. Two more coils were used to measure the occlusion plane. After recording, the resulting movement signals were rotated to the occlusion plane and the origin of the new coordinate system was located at the lower edge of the upper incisors. For data analysis, the movement signals were differentiated and all movement, velocity and acceleration signals were smoothed with a cutoff frequency of 20Hz. The tangential velocity signal was calculated from the x- and y-movement of the sensors.

2.4 Data segmentation

Figure 1 represents an EMMA display with tongue tip movements during the production of /nali/. The acoustic signal is shown in the upper panel, in the second panel the vertical movement of the tongue tip is represented, the third panel shows the tangential velocity signal. During the articulation of the sequence /nali/ 4 gestures were produced. The vertical lines in the figure indicate the first closing gesture (CG) towards /n/. The following gestures are the opening gesture (OG) towards the vowel /a/, the second closing gesture towards /l/ and the second opening gesture towards /i/.

By means of the tangential velocity signal (TV in cm/sec), shown in the third panel in figure 1, the start and end points of each gesture were defined by the left and right minimum, surrounding the velocity peak.

Figure 1: EMMA-display with the articulatory labelling criteria for the foreign words; panel 1: oscillogram, panel 2: vertical movement of the tongue tip sensor and panel 3: tangential velocity, vertical lines mark the 4 gestures

For the syllable repetition task one opening gesture towards /a/ and one closing gesture towards the stops /p/, /t/ or /k/ was measured for each syllable.

3   Results

For the analysis we focussed on the opening gestures from the nasal to the vowel /a/. All statistical calculations are based on the 10 repetitions of each foreign word.
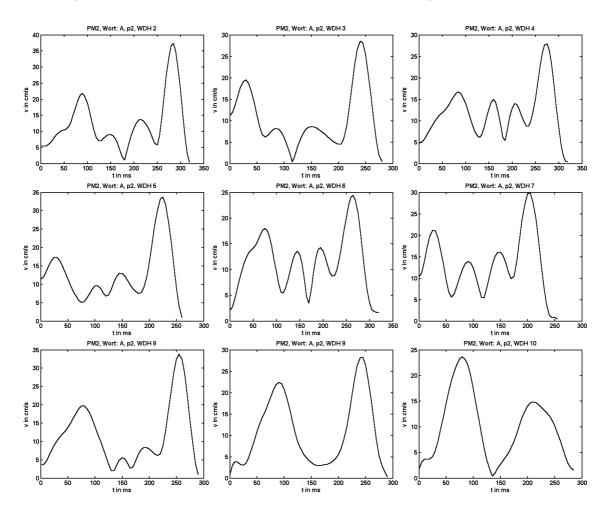
3.1   Qualitative analysis of movement patterns

As can be seen from table 1 a great amount of the data of the clutterers could not be analysed due to reduced gestures and target undershoot. An extreme amount of reductions- presumably because of high speech rate and mumbled speech-was found for speaker PM2, for whom it was impossible to analyse nearly half of his data in the p1-condition. This reduction phenomenon is presented in figure 2.

In line with the question of the perception of their own intelligibility one clutterer (PM2) sometimes noticed his fast speech rate. Speaker PW3 hardly noticed elisions e.g. realising [kɔnstitsi̯onaˈlɪstɪʃ] instead of [kɔnstitutsi̯onaˈlɪstɪʃ].

Table 1: Frequency of data that could not be analysed of /nali/ in the p1- and p2-condition

condition	Controls			Clutterers		
	NM1	NM2	NW3	PM1	PM2	PW3
p1		1 (2%)	1 (2%)	7 (14%)	21 (42%)	
p2	2 (5.7%)	13 (26%)	1 (2%)	16 (32%)	11 (22%)	6 (12%)

The following figures demonstrate the reduction phenomenon, exemplified for the tongue tip tangential velocity signal during /nali/ of speaker PM2 for nine of the ten repetitions of the word "emotionalisieren" in the p2-condition.

Figure 2: Tangential velocities of the tongue tip movement of PM2 during the nine repetitions of /nali/ in „emotionalisieren" in the p2-condition

In comparison to the segmented /nali/ shown in figure 1 it is noticeable that the second and third gesture always presents target undershoot, i.e. the two medial velocity peaks are much smaller than the first and the fourth, whereas in figure 1 the opposite is the case. In the last two displays these gestures were completely reduced.

3.2 *Quantitative assessment of articulatory variability*

Based on means and standard deviations the coefficient of variation (CV) was calculated for the durations and amplitudes of the opening gesture (OG) from the nasal towards the vowel /a/ in /nali/ in the p1- and p2-condition [CV = (standard deviation * 100)/mean]. Table 2 shows a higher coefficient of variation for the clutterers compared to the control group, e.g. the values of PM2 scatter with 19.7% and 22.6% over a much wider range from the mean than normal speakers with a maximum of 12.3%.

Table 2: Coefficient of variation (CV) in percent for the duration of the OG of /nali/ in the p1-/p2-condition

speaker		Controls			Clutterers		
		NM1	NM2	NW3	PM1	PM2	PW3
CV in %	p1	9.5	12.3	9.9	19.0	19.7	13.9
	p2	10.6	8.9	8.9	21.8	22.6	11.0

3.3 *Displacements*

The following figure also shows the frequency of the analysed foreign words. Speaker PM2 who produced only 29 instead of 50 analysable repetitions shows a smaller standard deviation than the other clutterers. However, the coefficient of variation (figure 4) indicates that his amplitudes vary to a similar amount as speaker PM1. Comparing the standard deviations of clutterers and controls in figure 3, it becomes clear that PM1 and PW3 produced the opening gesture with larger amplitudes which were also much more variable than the amplitudes of controls.
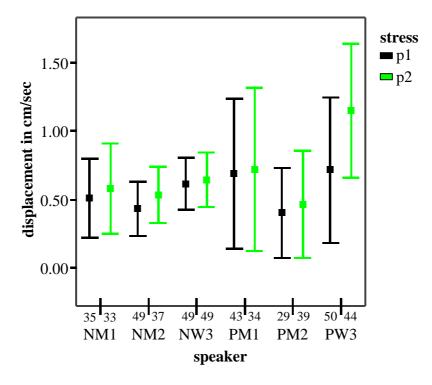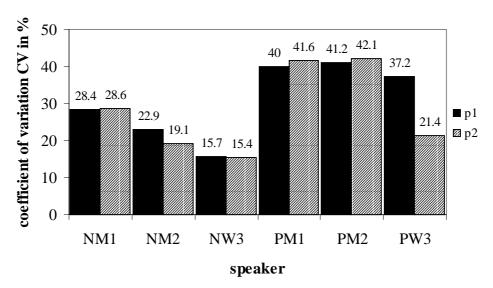
Figure 3: Standard deviation of the displacements of the OG of /nali/ in the p1/p2-condition

Figure 4 shows the coefficient of variation for the amplitudes of the OG of /nali/ in the p1-/p2-condition. As can be seen here, especially the results of the male clutterers are much more variable than speakers of the control group.

Figure 4: Coefficient of variation (CV) in percent for the displacements of the OG of /nali/ in the p1-/p2-condition

3.4 Syllable sequences

For the syllable repetition task, clutterers and controls did not differ significantly in the variability of the spatial and temporal domain.

3.5 Quantitative assessment of acoustic variability

Based on acoustic labelling the durations of the whole /nali/ sequences were calculated. As can be seen in figure 5 the coefficient of variation is higher for the group of clutterers than for normal speakers. Only PW3 in the p2-condition shows a similar value as the female control speaker.
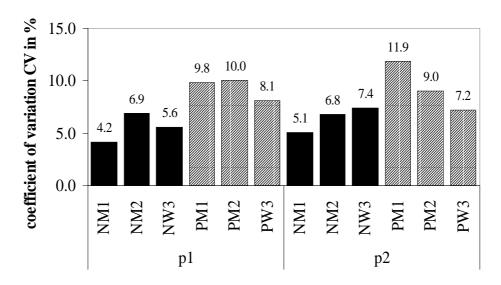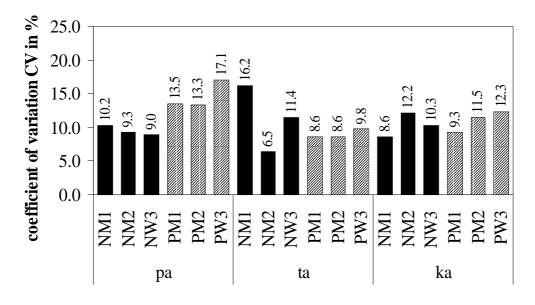

Figure 5: Coefficient of variation (CV) in percentage for the duration of the sequence /nali/

In agreement with the kinematic results, the syllable repetitions showed no prominent differences in the acoustic temporal variability between the two groups. Figures 6 illustrates the coefficient of variation, calculated for the syllable duration of /pa/, /ta/ and /ka/. Only in the production of /pa/ the durations are more variable.

Figure 6: Coefficient of variation (CV) in percentage for the duration of the syllable repetitions /pa/, /ta/ and /ka/

4   Discussion

Remarkably high coefficients of variation were found for the spatial and temporal data of clutterers. Against expectations clutterers showed not only variable displacements but also very large articulatory movements. On the one hand the displacements were partly larger than in normal speakers and were interpreted as strategy for improvement of the intelligibility. However, on the other hand more instances of movement data had to be excluded for the clutterers because of gestural reduction and extreme target undershoot.

Concerning the temporal results, again a higher variability was found for the clutterers. However, in comparison to the displacements the durations varied to a smaller degree. This means, that clutterers exhibit more variation in the spatial than in the temporal domain.

More variable results were found only for the fluent utterances of the clutterers. Their speech production altogether showed obviously a higher variability due to speech errors, break-ups within words and incomplete articulation.

The results of the present study are consistent with the findings on stuttering (Zimmermann 1980, Caruso et al. 1988, van Lieshout et al. 1993, Gracco 1994, Jäncke 1994, Ward 1997). The increased variability in stutterers was interpreted as a consequence of the usage of different articulatory strategies in order to avoid disfluencies (Ward 1997). For achieving the instructed intelligibility in the present study clutterers also used different strategies. They varied the speech rate and the displacements in order to speak clearly. In

comparison to normal speakers they might be more uncertain about their own speech production because they are aware of their problems and often do not know how to improve their speech quality. They control and vary their speech again an again. For one person (PW3) a high number of break-ups within words was observed. This might indicate that she always expects speech errors and therefore she stops speaking although no errors occurred.

Zimmermann (1980) explains the instability with an imbalance in the afferent-efferent nerve impulses and the critical spatial-temporal relationship in the command execution. This is one reason why research and treatment of cluttering should focus more strongly on the processes in the central nervous system. Not only for producing intelligible speech but also for the execution of other gross and fine movements is it necessary that the motor system with the motor cortex, basal ganglia, cerebellum works without any disruptions. Before reaching the muscles the commands run through the cerebellum, where the movements are modulated. The basal ganglia determine parameters like displacement, direction, velocity and strength of the movements.

In order to understand the processes in the brain it would be a worthwhile investigation to use functional magnetic resonance imaging (fMRI) for clutterers. A very recent study with right-handed stutterers has shown a right-hemispheric hyperactivity (Neumann et al. 2005). After speech therapy they observed a more widespread activation in the speech and language areas on the left hemisphere.

Both, in the articulatory and acoustic analysis of the present study, a higher variability for clutterers could not be observed for the syllable repetitions. The speech production system seems to work more stable during the syllable repetition task than during the articulation of foreign words with complex consonant clusters. Stop-vowel sequences are the most frequent syllable structures in general and are also the first linguistic units during the speech development of human beings (for a recent overview see Ackermann et al. 2005). This could explain why clutterers produced the syllable repetition task with ease.

An additional explanation of the observed variability patterns is the effect of speech rate. According to Fitts' (1954) law an increase in speech rate causes an increase in movement variability. Ward (1997) who analysed utterances in a normal, fast and slow speech rate confimed this relationship. In the present study a higher speech rate for clutterers was hypothesized, but this did not turn out to be true in all cases. Furthermore clutterers showed a high temporal variability which makes it difficult to discern the effect of speech rate and intrinsic variability. There must be other explanations for the higher coefficients of variation.

Another point of view concerns the linguistic complexity and is well documented in van Lieshout et al. (2004). Findings from Smith & Kleinow (2000) showed a greater movement variability for all speakers, not only for stutterers in syntactically more complex utterances. In terms of the present results it might be possible that the production of the foreign words were more difficult for the clutterers than for normal speakers. In a previous study van Lieshout (1995) explained the complexity of long words from a speech production view. According to him the brain has to prepare more production units for long words than for short words. While formulating, storing or executing these commands there are a lot of sources of errors. It should also be considered that complex words exhibit less common articulatory and prosodic patterns (van Lieshout 1995). This seems to be the case for the described test corpus. Because clutterers have problems in general with long and complex words further investigations should follow this question, for instance by means of fMRI.

For speech therapy is it advisable to practise a very clear speech with large articulatory movements. This strategy seems easier to realise for clutterers than the task to reduce speech rate.

References

Ackermann, H., Hertrich, I., Mathiak, K. (2005) Neurobiologische Grundlagen der Sprachlautwahrnehmung: Klinische und funktionell-bildgebende Befunde. *Sprache, Stimme, Gehör*, 29, (3): 112-120.

Braun, O. (1999) *Sprachstörungen bei Kindern und Jugendlichen: Diagnostik, Therapie, Förderung*. Stuttgart: Kohlhammer.

Carstens (1992) Articulograph AG100. *Elektromagnetisches Artikulations-Meßsystem*. Benutzerhandbuch. Carstens Medizinelektronik GmbH.

Caruso, A. J., Abbs, J. H., Gracco, V. L. (1988) Kinematic analysis of multiple movement coordination during speech in stutterers. *Brain*, 111: 439-455.

Fitts, P. M. (1954) The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47: 381-391.

Gracco, V. L. (1994) Some organizational characteristics of speech movement control. *Journal of Speech and Hearing Research*, 37: 4-27.

Jaeger, M., Hertrich, I., Stattrop, U., Schönle, P.-W., Ackermann, H. (2000) Speech disorders following severe traumatic brain injury: Kinematic analysis of syllable repetitions using electromagnetic articulography. *Folia Phoniatrica et Logopaedica*, 52: 187-196.

Jäncke, L. (1994) Variability and duration of voice onset time and phonation in stuttering and nonstuttering adults. *Journal of Fluency Disorders*, 19: 21-37.

Katz, W., Machetanz, J., Orth, U., Schönle, P.-W. (1990) A kinematic analysis of anticipatory coarticulation in the speech of anterior aphasis subjects using electromagnetic articulography. *Brain and Language*, 38, (4): 555-575.

Kretschmer, I. M. (1996) *Untersuchungen zur Analyse von Sprech- und Schluckbewegungen mit Hilfe der elektromagnetischen Artikulographie*. dissertation, Universität Tübingen.

Levelt, W. J. M. (1991) *Speaking*. 2. edition, Cambridge, Massachusetts: MIT Press.

McClean, M. D., Runyan, C. M. (2000) Variations in the relative speeds of orofacial structures with stuttering severity. *Journal of Speech, Language and Hearing Research*, 43, (6): 1524-1531.

McClean, M. D., Tasko, S. M., Runyan, C. M. (2004) Orofacial movements associated with fluent speech in persons who stutter. *Journal of Speech, Language, and Hearing Research*, 47, (2): 294-303.

Molt, L. F. (1996) An examination of various aspects of auditory processing in clutterers. *Journal of Fluency Disorders*, 21: 215-225.

Neumann, K., Preibisch, C., Euler, H. A., von Gudenberg, A. W., Lanfermann, H., Gall, V., Giraud, A.-L. (2005) Cortical plasticity associated with stuttering therapy. *Journal of Fluency Disorders*, 30: 23-39.

Perkell, J.S., Nelson, W.L. (1985) Variability in production of the vowels /i/ and /a/. *Journal of the Acoustical Society of America*, 77: 1889-1895.

Scherer, A. (2003) Poltern und Stottern als Ausdruck der emotionalen Befindlichkeit: ein Erfahrungsbericht aus sprachtherapeutischer Sicht. *Sprache, Stimme, Gehör*, 27: 88-91.

Sick, U. (2000) Spontansprache bei Poltern. *Forum Logopädie*, 4, (14): 7-16.

Smith, A., Kleinow, J. (2000) Kinematic correlates of speaking rate changes in stuttering and normally fluent adults. *Journal of Speech, Language and Hearing Research*, 43, (2): 521-36.

van Lieshout, P. H. H. M. (1995) *Motor planning and articulation in fluent speech of stutterers and nonstutterers*. dissertation, Nijmegen Institute for Cognition and Information.

van Lieshout, P. H. H. M., Alfonso, P. J., Hulstijn, W., Peters, H. F. M. (1993) Electromagnetic articulography (EMA) in stuttering research. *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München (FIPKM)* 31: 215-224.

van Lieshout, P. H. H. M., Hulstijn, W., Peters, H. F. M. (2004) Searching for the weak link in the speech production chain of people who stutter: A motor skill approach. In: B. Maassen, R. Kent, H. Peters, P. H. H. M. van Lieshout & W. Hulstijn (eds.) *Speech Motor Conrol in Normal and Disordered Speech*. Oxford: Oxford University Press: 313-355.

van Riper, C. (1990) Final thoughts about stuttering. *Journal of Fluency Disorders*, 15: 317-318.

Ward, D. (1997) Intrinsic and extrinsic timing in stutterers' speech: Data and implications. *Language and Speech*, 40, (3): 289-310.

Wirth, G. (1994): *Sprachstörungen, Sprechstörungen, Kindliche Hörstörungen*. 4. edition, Köln: Deutscher Ärzte-Verlag.

Zimmermann, G. (1980) Stuttering: A disorder of movement. *Journal of Speech and Hearing Research*, 23: 122-136.