Evaluation of a Training Program Using Visible Speech in German Schools for the Deaf

Article *in* Journal of Deaf Studies and Deaf Education · February 1998

DOI: 10.1093/oxfordjournals.deafed.a014345 · Source: PubMed

Cordula Artelt
Leibniz Institute for Educational Trajectories

Evaluation of a Training Program Using Visible Speech in German Schools for the Deaf

Regine Becker
Humboldt University of Berlin

Cordula Artelt
University of Potsdam

The present study provides insight into several technical and computer-aided procedures that can be used with deaf students to improve their speech. A visible speech device combined with a training program was developed and evaluated in six German schools for the deaf. The teachers were initially given special instructions about the work with their students in accordance with our technical and training concept. The students were able to build up a knowledge of visible patterns of spoken sounds, syllables, and words. They used this knowledge to recognize speech errors and to improve their pronunciation. Teachers' experiences were mostly positive. As they said, this device made their work much easier. The results show that visible speech can be used effectively; however, it has to be embedded in a training program that systematically builds up phonological knowledge and considers symbolic meanings from signed and written language.

Research evidence indicates that the majority of deaf individuals do not have age-appropriate speech and language skills (Conrad, 1979; McAnnaly, Rose, & Quigley, 1987). The speech intelligibility of the majority of hearing-impaired individuals with severe to profound hearing loss is generally low (20%), as the results of speech instruction in the past show (Markides, 1970). Similar intelligibility scores were found in German schools (Becker, 1993; Becker, Schildhammer, & Ruoss, 1994) despite technical and educational improvements of the oral-aural approach.

This article deals with the development of a speech training curriculum in German schools. More efficient methods are necessary to improve speech intelligibility of deaf speakers by taking into account time and effort involved in speech training in the classroom.

Technical Aids for Speech Training

Because of the rapid development of instructional, interactive technology and signal-analyzing devices, computer-based speech training aids have become increasingly important instruments in the work with deaf students. Most of these computer-based resources (Arends, Povel, van Os, Michielsen, Claasen, & Feiter, 1991; Bernstein, Goldstein, & Mahshie, 1988; Lippmann, 1982; Watson & Kewley-Port, 1989) are feedback methods enabling external control over peripheral speech processes (like phonation, prosody, and articulation). The aim of these aids is to improve the intelligibility of deaf speakers. Speech parameters (like pitch, intensity, vowel articulation) are trained by compensating for the lack of auditory feedback through the use of vision and tactile perception. Speech training aids offered in German schools for the deaf provide visual (Esser, Nolte, & Printzen, 1983; Neppert, 1990) or vibrotactile (Breiner, 1990) representations of speech sounds. Furthermore, physiological methods for measuring and visualizing speech movements (e.g., lip and

The investigation was supported by a grant of the Federal Ministry for Research and Technology (registration no. 01VJ9301/3) awarded to Prof. Dr. M. Krause and Prof. Dr. K Eyferth at the Technical University of Berlin. We take responsibility for the content of this publication. We thank the participating teachers and students for their cooperation and engagement; Dr. Anita Schildhammer, who contributed to the conception and data collection; and Dr. Elke Brenstein and Andrew Barnes for their help with translation. Correspondence should be sent to Dr. Regine Becker, Ev. Geriatric Center of Berlin, Reinickendorfer Str. 61, D-13347 Berlin, Germany (e-mail: rbecker@ukrv.de).

tongue movement) are available in clinics (Klajman, Huber, Neumann, Wein, & Böckler, 1988; Schönle, Gräbe, Wenig, Höhne, Schrader, & Conrad, 1987).

Technical aids for deaf students can be differentiated according to their main area of application: improvement of speech production or perception.

Speech Production Training

Levelt (1989) describes articulation as the motor execution of a phonetic plan and the syllable as a unit of motor execution in speech. When executed, the syllables' phones are overlapping and interacting gestures over time. Each gesture consists of dynamic phonetic features, such as voicing and nasalizing (Levelt, 1989; p. 346). Speech training aids mainly support the development of self-regulation of speech processes. Via direct feedback of gestures or single phonetic features and the comparison with a target, one should be able to improve and consolidate speech production skills and speech motor control. Such improvements should be reflected in the development of spontaneous speech as well.

Most speech training systems are used for the modification of suprasegmental aspects of speech (intensity, pitch) and acoustic characteristics of single sounds (voicing, formant positions). The IBM SpeechViewer system is a typical device that provides exercises for phonation, pitch courses, voice-voiceless contrast, and vowel accuracy. Pratt, Heintzelmann, and Deming (1993) reported a study of six preschool children with hearing-loss who were trained using the vowel accuracy module of the IBM SpeechViewer. The subjects showed a treatment effect for at least one of the three trained vowels. In no case did the pronounciation of all three vowels improve through training. Problems arose with the inaccuracy of the feedback, the lack of instructional feedback, and the nonlinearity of the criterion-adjustment control provided by the computer system. The authors therefore concluded that additional external validations like listener judgments are needed to evaluate the improvement of speech production skills. Furthermore, it may be that training of single sounds cannot result in fluent speech because coarticulatory effects are ignored.

Watson, Reed, Kewley-Port, and Maki (1989) developed and evaluated ISTRA (Indiana Speech Training Aid), which provides a computer-based judgment of speech quality. Their results showed strong relations between computer-based and human judgments of speech quality. The speech training system contains a speech recognition processor that provides a goodness metric to the trainee based on his best tokens. Without setting normative data as a criterion, the trainee is guided systematically toward better speech productions.

Mahshie, Vari-Alquist, Waddy-Smith, and Bernstein (1988) presented a study of computer-based speech training and practice aids for profoundly deaf students. Acoustic and physiological measurements obtained from a microphone, an electroglottograph, and a pneumotachograph were used for assessment and training in a school or clinic. In the software development, especially, the needs of young children were taken into account. Different games were implemented to teach sustained vocalization, production of repeated syllables, and control of voice intensity and fundamental frequency (F_0). Speech training was conducted over a 15-month period with 15 deaf children ages 3 to 11 years. Promising results were presented related to the ease of use of the aid, the acceptance by the children of the aid and the displays, and increased motivation of the trainees for speech training.

In a study by Arends et al. (1991) prelingually deaf children ages 4 to 7 years were instructed in different subskills of speech (i.e., control of timing, voicing, loudness, pitch, and vowel production). The 22 children in the experimental group were trained by using the Visual Speech Apparatus (VSA), and the 16 children in the control group received speech training based on the maternal reflective method of van Uden (1977), a method focusing on a natural instructional approach to language. Speech training was given only on the basis of children's spontanously produced speech. The performance of the children was measured using the CID Phonetic Inventory and a specially developed test. The results showed that the experimental group achieved significantly higher scores for voice control and vowel production than the control group. The younger children especially benefited from the training with VSA. Segmental characteristics of longer speech units (syllables, words, and sentences) can be trained via a spectrographic display in real-time (Lippman, 1982; Stewart, Larkin, & Houde, 1976), which provides information about the changing spectral patterns of speech caused by articulatory movement and coarticulatory influences.

Speech Perception Training

Along with conventional hearing aids, there are vibrotactile aids to support speech perception of deaf subjects. Only 30% of phonemic information is visualized by articulatory gestures of the lips. The Minifonator, a vibrotactile device developed by Siemens (Weisenberg, 1989) is aimed at compensating insufficient lipreading information. Utterances spoken into a microphone are transformed into vibrations that can be perceived by the skin. The amplitude of the incoming acoustic waveform is used to modulate the amplitude of a broad-band vibratory carrier. This device allows discriminations of rhythm, stress, and duration patterns. Furthermore, spectral information like voice-voiceless contrasts is conveyed through the use of the broad-band vibratory carrier. The aid can be used to support speech perception, but also to improve speech production skills.

Most of the reported studies deal with the instrumental aspect of computer-based speech production and perception training. The efficacy of technical aids in speech training of deaf students, however, depends on the development of appropriate curricula. Further research must therefore focus on the learning conditions under which visual or tactile feedback helps to improve speech production skills.

Treatment Effects With Spectrographic Display

There is no linear relationship between the acoustic signal and the perceived speech code (Cole, Rudnicky, Zue, & Reddy, 1980; Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1968). The decoding of speech sounds is based on context-dependent, rule- and knowledge-driven parallel processes with which human speech perception operates. These special top-down processes are not taken into account if the acoustic stimulus is transformed into another sensory modality. Only parts of the linguistic information can be translated. A study by Cole et al. (1980) showed that after intensive training about 80% of the phonetic segments of unknown spectrograms were labeled correctly by a hearing subject. Segmentation and labeling processes were performed with the help of phonetic and phonotactic knowledge without the use of higher-level knowledge (syntactic or semantic). The symbolic content of the test utterances was only partly disclosed after the phonetic analyses were completed. This means that no on-line speech processing takes place and special instruction concerning the interpretation and learning of the visible speech patterns is necessary for the use of spectrographic aids in speech training. Furthermore, a special program for deaf children must allow them to build up phonetic and phonological knowledge and consider symbolic meaning from signed and written language.

A visual speech device providing spectrographic feedback with a training program for deaf students of different ages was developed and evaluated by our research group (Becker, 1993; Hobohm, 1993; Ruoss, 1992; Schildhammer 1996). In our pilot study Schildhammer (1996) investigated the influence of speech training with the visual speech device over a 12-month period (70 15-minute training sessions per child). Twenty-three prelingually deaf students with a mean hearing loss from 85 to 117 dB ages 9 to 15 years were investigated. The speech training included exercises for basic attributes (timing, voicing, and loudness) and articulatory skills at the word level (see *Training program*) and was held individually two times a week. The intelligibility of tape-recorded trained and untrained words (48 items per student) was judged by naive listeners both before and after the students' training. Differences in training results were found dependent on school grade level and speech production skills attained before the training started. In the lower grades (9 to 10,6 years) only the less intelligible students improved their speech production skills significantly with the visual speech aid (on the average from 25.2% to 38.4% intelligible phonemes). In the higher grades (12,6 to 15,3 years), however, the different intelligibility levels before training started had no effect on the improvements of the students' speech production skills after training. All of the older students showed significantly increased intelligibility ratings after participating in the training scheme. The less intelligible older students increased from a mean of 31.4% to 35.7% intelligibility and the more intelligible ones from 40.8% to 43.3%.

These differences in the results for younger compared to older students may have been caused by the fact that conventional speech lessons are usually given only between the first and the fourth grades. The initially less intelligible younger students may have improved speech intelligibility with the spectrographic display because of increased motivation and the provision of important information about the acoustic and articulatory properties of fluent speech, that lipreading and tactile cues are not able to offer. However, the more intelligible younger students seem to have already reached a certain level of speech production skills (on average 42% of phonemes were intelligible), and thus there was no significant improvement from practice with the additional visible speech exercises. The older students, however, had had regular speech lessons in the elementary school some years ago. Through the minimal amount of speech practice in the meantime, they have lost former articulatory skills. Therefore, all of them profited by the speech training with the visual speech aid because it allowed them to refresh previously attained but recently lost speech abilities.

Research Questions

As results from the pilot study showed, the visual speech aid can be used as an effective medium in the speech training of deaf students. Visual feedback of speech sounds in real time can improve speech production by increasing external control and providing a more efficient way of acquiring speech motor skills. The less intelligible younger deaf students and all of the older students benefited from the visual speech aid. The system has now been tested in five German schools over a 4-month period. Two major questions led this investigation. The first is related to *teachers' acceptance* and implementation of the program. What do teachers experience when using visible speech in speech training with deaf students of different age levels? How do they handle the technical equipment, the segmentation, and labeling processes of the spectrograms and the implementation of the whole training program?

The second question is related to *students' achievements*. What level of performance do students reach after training with visible speech as measured by the number of corrected speech errors? What is the relation between the results achieved during the training program in terms of modified speech motor skills and the degree of hearing loss? How do the individual error profiles and the various stages of the training program influence training achievements?

Table 1 Description of the sample (grade level, average age, and number of students, number of training sessions per teacher)

School	Grade level	No. students	Age (year; month) (*M*)	No. training sessions
Güstrow	1	7[a]	7;03	116
Güstrow	7	5[a,b]	13;02	50
Güstrow	10	5[a,b]	15;06	57
Bremen	7	3	15;02	10
Bremen	3	3	9;04	25
Bremen	5	7	11;11	46
Bremen	Kindergarten	5	4;06	39
Oldenburg	1	5	7;10	96
Oldenburg	10	7[b]	17;09	103
Dortmund	1	7	8;00	44
Dortmund	1	7	8;03	21
Dortmund	5	10	12;05	42
Berlin	2	4	8;06	—

[a]Two teachers per class.

[b]Subsample (see Table 2).

Method

Participants

Sixteen teachers from 5 German schools for the deaf participated in our study with a total of 75 students of different grade levels. Three of the 13 classes were instructed by two teachers. Table 1 gives a description of the sample. The whole research period (after introduction into the training program) lasted four months. As Table 1 shows, however, there were differences among teachers in the amount of lesson time devoted to using the device (total number of training sessions).

We assumed that the amount and regularity of

Table 2 Description of the subsample: grade level, degree of hearing loss at 500 Hz, 1 kHz, 2 kHz with and without hearing aid, number of sessions, and training stages of individual students

Student	Grade level	Age	500 Hz	500 Hz with	1 kHz	1 kHz with	2 kHz	2 kHz with	No. sess.	No. train. stages
K1	10	16;11	90	60	100	60	110	65	15	9
K2	10	17;03	80	—	90	—	100	—	14	9
K3	10	17;07	85	60	90	55	105	60	16	10
K4	10	18;03	75	65	95	65	100	70	15	9
K5	10	17;05	90	60	105	65	100	65	14	8
K6	10	18;00	85	65	100	70	105	70	14	9
K7	10	18;11	90	65	85	45	70	45	14	8
K8	7	13;01	95	50	110	45	120	45	10	8
K9	7	12;09	95	70	115	55	115	65	11	8
K10	7	13;07	100	65	110	55	110	60	10	8
K11	7	13;01	90	55	100	40	115	55	10	8
K12	7	13;04	95	55	110	45	115	50	9	8
K13	10	14;11	95	65	105	50	120	60	16	11
K14	10	15;08	105	60	110	45	110	55	16	11
K15	10	15;06	80	55	90	45	85	35	16	11
K16	10	15;06	90	65	90	50	85	55	17	11
K17	10	15;11	95	65	110	55	120	70	17	11

training sessions as well as the age of the students would have an effect on learning. A detailed analysis of learning achievements (number of speech errors and corrections) can therefore only be given for the older students of the sample because only with them was it possible to keep the learning conditions constant (see also *Experiences depending on age group*). Table 2 contains a description of this subsample consisting of 17 students of the higher grades. Each student in the selected subsample received on average 14 training sessions during the investigation.

In German schools systematic speech lessons are given in kindergarten and from first to fourth grade, but students of the higher grades do not receive any regular speech training. In the last 10 years the use of substitutional media (German Sign Language, signed German) and an emphasis on written language have been discussed with the aim of improving the acquisition of a formal language system (Prillwitz & Wudtke, 1988; Wisch, 1990). Several teachers are recently using signed German or other manual methods to facilitate communication and instruction. The students in our sample had received speech lessons that included multisensory stimulation for speech, a focus on the development of speech sounds and the syllable as the basic unit of instruction (except one class of the lower grades that was trained with the auditory global method).

Equipment and Procedure for Speech Training

The visible speech device provides spectrographic feedback in real time (Hobohm, 1993). At the moment, we work with a visualizing device consisting of a DOS PC (386 or higher), TIGA graphic card (TMS 34010), signal processing card (AT&T DSP 32c), color monitor, microphone preamplifier (with preemphasis 12 dB/octave), and two head microphones (Beyerdynamic M560 N[C]). A Fast-Fourier-Transformation (128 point-FFT, Hamming window, time resolution 8.4 ms) was used for speech analysis. The linear frequency scale runs from 50 to 5800 Hz. The intensity of energy in the speech signal is indicated by color values. High spectral energy is displayed in dark color values (purple, black) and low spectral energy is indicated by light color values (green, yellow) (see Figure 1). The duration of the speech segment that can be analyzed is 2.5 seconds, or 4.8 seconds when special hardware compo-

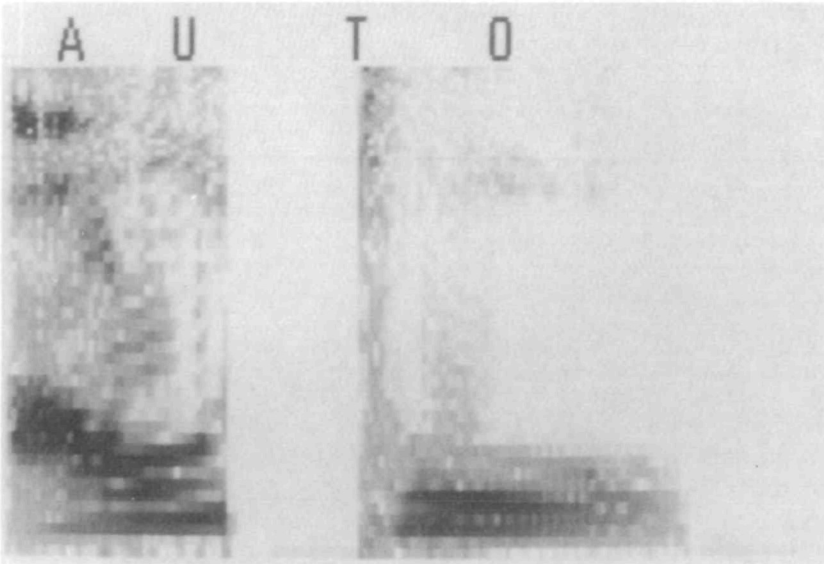

Figure 1 Spectrographic display of the German word "Auto"/auto/(car). In the original computer display, spectral energy is indicated by color values.

nents with extended graphic and memory powers are used.

Throughout the training, both teacher and student wear head microphones that are alternately activated. The monitor display is divided into two independently movable traces. The ongoing display can be stopped with a key stroke for comparisons of speech output. On the upper track, the teacher gives a reference pattern. He asks the student to label the single sound segments of any spectrogram produced. The student then repeats the word until it is similar to the reference pattern on the lower trace. Every time the teacher wants to inform the student about speech deviations or about correct speech production, he pauses the ongoing display. During speech training, manual methods of communication (signed German or fingerspelling) and written language are used to illustrate the meaning of training words and to give speech instruction.

Training sessions are held individually. A training session is not supposed to last longer than 15 minutes and takes place at least two times a week. At least one session is planned for each stage of the training. Depending on individual speech performance, a training stage can last for several sessions. Individual performance levels determine the amount of speech material and intensity of training within each training stage.

A menu is placed beneath the two movable display tracks with hot keys for the different functions of the program (e.g., labeling of speech segments, editing and saving functions, or changing of single display parameters). A short description of these menu commands can be recalled via a help function, which is also described in more detail in the users' guide.

Training program. In addition to the development of the visual speech aid, a training program for the use of spectrographic feedback in speech training with deaf students was developed and evaluated (Ruoss, 1992; Schildhammer, 1996; Schildhammer & Becker, 1993). The training program starts with exercises to train basic skills (breathing control, phonation, nasality, rhythm, phoneme contrasts, etc.) that are based on exercises proposed by Lindner (1984). Furthermore, the program includes 11 training stages (see Table 3) in which correct production of different vowels and consonants is taught in the context of words. The phonemes are introduced successively in a special order proposed by Kloster-Jensen and Jussen (1974) that is based in part on articulatory difficulty. Furthermore, acoustic characteristics are considered when teaching the visual similarities and differences of sound patterns in the different stages of the training program. In the initial training stages, only those combinations of consonants and vowels are introduced that are easy to dis-

Figure 2 Application of the visual speech device.

Table 3 Training program with 11 successive training stages for word production

	Phonemes	Words like . . .
Stage 1	/p t k/ /a o u/	Papa, Tag, Opa
Stage 2	/f ʃ s/	Schaf, Fuß
Stage 3	/n m ŋ/	Nuß, Oma
Stage 4	/h/	Haus, Hof
Stage 5	/x/	Fach, Tuch
Stage 6	/ç/ /i I e ɛ ø œ y Y aI/	Teich, Fisch
Stage 7	/b d g/	Biene, Geige
Stage 8	/v z j/	Junge Wiese
Stage 9	/l/	Lippen, Welle
Stage 10	/R/	Reise, Afrika
Stage 11	/pf ts ks/	Apfel, Katze, Kekse

tinguish on the computer display. For example, back vowels and voiceless plosives in the word "Opa" (grandpa) are easy to distinguish because of their very different acoustic features. In the following training stages, the discrimination of the sound patterns becomes progressively harder as more and more new phonemes are introduced, including phonemes with similar sound patterns that are more difficult to discriminate. With this structured speech instruction, deaf students can build up a greater phonetic and phonological knowledge. They learn to segment and label visible speech patterns of rehearsed and nonrehearsed words as well as to improve motor skills.

Speech production is associated with other language and speech modalities. Word meanings (symbolic relations) are worked out through the presentation of signs and written words. If the student fails to master specific phonemes within a word context, additional prompts are given such as tactile cues, speechreading the sound, and listening to the sound. Basic exercises with visual feedback produced by single sounds or syllables are also practiced.

Spectrographic features and speech errors. For different speech sounds, specific invariant acoustic patterns can be found in the spectrographic display, which can be categorized following a learning episode (Cole et al., 1980). As results of Ruoss and Schildhammer (1993) indicated, the deaf students' ability to segment and label visible speech patterns increased with specific speech training. After a 12-month period deaf students ages 9 to 15,6 years reached pattern recognition scores similar to those of adult experts who are very familiar with visible speech patterns (see Table 4).

To work with their students in a similar way, the teachers of our sample received special instruction in segmenting and labeling the visible speech patterns with the help of invariant phonetic cues. In our training program (Schildhammer & Becker, 1993), acoustic features of speech sounds are introduced sucessively via

Table 4 Mean recognition scores for deaf students ($n = 20$) ages 9 to 15, 6 years and experts ($n = 4$) for 12 visible speech patterns (words) (Ruoss & Schildhammer, 1993)

Correct answers (mean score in %)	Deaf students	Experts
Segmentation	94.8	96.4
Phoneme labeling	55.2	46.9
Within phoneme class	28.2	37.8
Word recognition	19.6	6.3

special test materials. Distinctions among phoneme classes (plosives, fricatives, nasals, vowels) or even recognition of single phonemes (e.g., /sch/) are learned with training.

Special instruction by the teacher is needed so that deaf students learn to recognize their own speech errors and learn to correct them with the help of visual feedback. For this purpose, teachers have to be familiar with the visual characteristics of different speech errors. For the present purposes, examples of visualized words from deaf speakers who were trained with visible speech (Ruoss, 1992) were used for illustration. We grouped these examples by contrasting an incorrect pronunciation of a word (pretraining) with an intelligible pronunciation of the word (posttraining) and the word spoken by a subject without hearing-loss (Schildhammer & Becker, 1993). Typical patterns of speech errors can be shown in this way (Becker, 1993; Becker & Druba, 1994), but even intelligible speech pronunciation by deaf subjects differs significantly from the given target. This difference appears to be the consequence of speaker-specific pitch characteristics. A speaker-independent display is not possible with Fast-Fourier-Transformation. Phonatory deviations (like falsetto pitch, breathiness, laryngeal strain), in particular, influence the visual patterns in a negative way. Further differences among speakers are caused by different midfrequencies and band-widths of formants based on different measurements of vocal tracts. In the context of speech training, teachers have to decide (also from listening to the speech) which spectrographic deviations are crucial speech errors or unimportant speaker-specific variations. The teachers therefore received intensive training on the device. During individual sessions with students, they were instructed to analyze and interpret speech distortions with the help of the visual speech aid. The teachers' knowledge about visual patterns of speech errors naturally increased with practice and supervision. Frequent communication between experts and teachers ensured that problems could be solved and any inappropriate use of the device could be corrected.

Procedures for Data Collection

Teachers' acceptance of the program were assessed through a questionnaire. The items referred to (1) the technical use of the device, (2) the segmentation and labeling of speech patterns, (3) the teachers' goals in using the device during speech lessons, (4) possible combinations with other training approaches, (5) evaluation of the speech training aid, and (6) ideas for improvement of the visual speech display.

A second measuring instrument was developed for the recording of the individual learning and speech performances. The teachers had to collect the number and type of training words, speech errors, and corrections for each learning stage of the training program. Phonetic transcriptions of the spoken words were made by the teachers to record speech errors. We used two measurements of word corrections. "Short-term corrections" are speech corrections that were observed to occur within one training session. In the next session a selection of these corrected words (not all) was presented together with new training words (of the same learning stage). Speech corrections were classified as "habitual" if they were maintained over two consecutive sessions.

Teachers' instructions and procedures were the same across the five participating schools. The visual speech device was installed at each school. Instruction of the teachers took place on two successive school days. It contained (1) a theoretical introduction to the speech training program, (2) practical examples and exercises in the use of the device, (3) introduction to the measuring instruments, (4) exercises in the segmentation and labeling of sound-patterns, and (5) exercises in the assessment of speech errors. Members of the research team visited schools regularly for supervision. Technical problems were solved, questions concerning practical work were answered, and, where

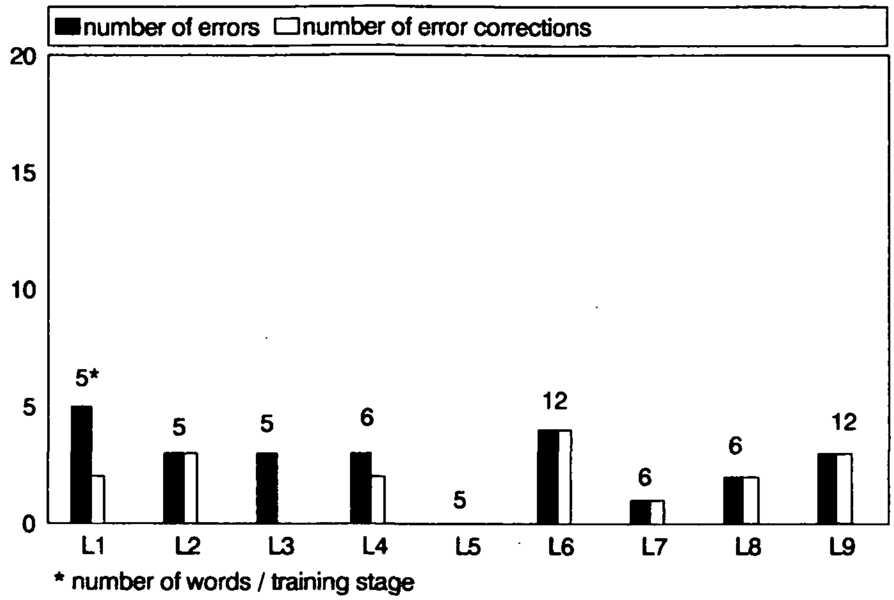

Figure 3 Absolute number of training words, speech errors and error corrections over nine training stages (L1 to L9) for one subject.

needed, theoretical issues regarding the program were elaborated. At least one individual training session per teacher was supervised and discussed.

Data collection was completed in 17 weeks, after which teachers were given the questionnaire concerning the acceptance and implementation of the program. The second measuring instrument was used by the teachers to record the individual learning and speech performances per student and training session. The data concerning a student's training achievement were thus collected for each training session given.

The sample was devided into two age groups. In the first group, there were 33 students (1st to 4th grade) ages 7 to 9 years and 5 preschool children ages 4 to 5 years. They were trained by eight teachers. In the second age group there were 37 students (5th to 10th grade) ages 12 to 19 years. Eight teachers worked with them.

Measurements

Teacher questionnaire (evaluation of the program). Descriptive analyses of the teacher questionnaire items were conducted separately for each group according to the different age levels and training conditions. Most of the questionnaire items were multiple choice items where more than one answer was possible. When teachers were asked to name other variables that they thought influence students' improvement of articulation, they were free to list those variables they thought were relevant (e.g., students' motivation, intelligence, degree of hearing loss). All analyses of the questionnaire items are based on the proportion of chosen items and free answers.

Assessment of students' achievement. During individual training sessions with each student the teachers decided whether training words were pronounced correctly or which speech errors occurred in a training word. Most of the teachers had no problems in segmenting the patterns after they had received instruction. Recognizing the invariant patterns and labeling the spectrograms were more difficult. The teachers learned step by step to segment and label the visible speech patterns with the help of invariant phonetic cues (see Ruoss & Schildhammer, 1993). The number of training stages undertaken with students varied strongly across schools and classes. Teachers used the measuring instrument after each training session to (1) list the words that were trained, (2) the specific speech

Table 5 Five error categories and their sources

Category	Source
Substitution	(a) Voicing error
	(b) Place of articulation
	(c) Manner of articulation
	(d) Place and manner of articulation
	(e) Vowel duration
	(f) Substitution, neutralization or nasalization of vowel and diphtong
Omission	(a) Vowel
	(b) Consonant
Addition	(a) Sound with related articulatory position
	(b) Neutral vowel between consonant blend
	(c) Neutral vowel at word final position
	(d) Neutral vowel at word medial position
	(e) Aspiration of vowel
	(f) Others
Phonation	(a) Inappropriate pitch level
	(b) Pitch fluctuation
	(c) Impaired loudness control
	(e) Pitch breaks
No classification	—

errors made, (3) the speech error corrections made with the help of the device, and (4) the stability of corrections (correct pronounciation of those words that were corrected during the previous session). See Figure 3 for a single case illustration of the relations between number of training words, errors, and error correction over nine training stages.

Students' *speech errors* could be categorized according to the following five error categories: substitution, omission, addition, phonatory deficits, and unclassified errors. More than one error could occur per word. Table 5 shows the different sources of speech errors. For short-term *corrections of speech errors* we formed three categories: corrected, partly corrected, and not corrected words. Their relative frequency per type of speech error and training stage was calculated. Furthermore, we differentiated between four types of *habitual corrections:* stable, partly stable, nonstable, not checked. Again, the relative frequency in each speech error category and training stage was calculated.

The number of training words, speech errors, short-term corrections, and habitual corrections varied across students. For the purpose of comparison we aggregated data and formed three measures: (1) students' speech performance before training session: each participant's number of speech errors for each error type was converted to percent (based on the total number of trained words); (2) students' speech performance after training session: each participant's number of corrections for each correction type was converted to percent (based on the total number of speech errors made); and (3) students' stability of speech performance after two consecutive sessions: each participant's number of habitual corrections for each stability type was converted to percent (based on the total number of corrections made). All three measures were computed both as an overall measure as well as separately for the nine completed training stages.

Results

Teachers' Evaluation of the Program (Teacher Questionnaire)

The majority of teachers (93%) found the use of the device easy without studying the handbook intensively. Most of the teachers had no problems in segmenting the patterns after instruction. Recognizing the invariant patterns and labeling the spectrograms were more difficult to manage. The number of training stages undertaken with students varied strongly across schools and classes. These conditions had not been sufficiently controlled by the researchers. Only half of the teachers worked strictly in the way the training program suggested.

Preferences suggested by the teachers related to the further use of the visual speech aid in lessons were two teachers per class (50%), training to be conducted individually (31.3%), the visual device to be available in the classroom (37.5%), an emphasis on the improvement of articulation (93.8%), and speech error assessment (31.3%). As the majority of teachers said, it was possible to combine the visual speech aid with different training approaches of speech and language instruction as well as manual methods of communication. From their point of view, students' acceptance and motivation was high. It was possible to work with the device with nearly all students. There were not always articulation improvements, but teachers' judgments about the use of the device were generally very positive. They

Table 6 Proportions (%) of teachers' free and fixed responses to the questionnaire items depending on students' grade level: Grades K–4 and Grades 5–10

Item / frequency %	Grades K–4 (8 teachers)	Grades 5–10 (8 teachers)
Speech instruction		
Exercises following the program[a]		
Basic exercises	87.5	100
Word exercises	75.0	87.5
Different phoneme order	50.0	12.5
Individual exercises	87.5	75.0
Individual exercises adapted to or based on[b]		
Students' individual problems	37.5	0
Students' vocabulary	62.5	25.0
Materials from lessons	25.0	25.0
Review of the device		
Improvement of articulation[a]	65.3	86.7
Other variables that might have effects on the learning outcomes[b]		
Degree of hearing loss	37.5	0
Speech performance	75.0	25.0
Motivation	12.5	0

[a]Based on fixed responses.

[b]Based on free responses.

concluded that the device is easy to use, a good motivator, and an effective supplement to conventional speech instruction methods.

For some questionnaire items, however, age differences could be found (see Table 6). The training program consists of 11 training stages in which all German sounds are successively introduced. The teachers of the younger children changed the suggested phoneme order because some articulation patterns had not yet been acquired by their children. The teachers of the younger students also modified the training program by taking more into account the students' individual needs and abilities, particularly concerning vocabulary and sounds. The teachers of the older students worked in the suggested manner.

Articulation improvements were more often reported with the older students than with the younger ones. Variables like teachers' oberservations of children's intelligence and degree of hearing loss seemed to have a stronger effect on the results of speech training of the preschool children and students of the lower grades. On the other hand, in our pilot study Schildhammer (1996) found a significant correlation score (r = .59) between intelligence (measured by Raven's Standard Progressive Matrices) and improvement of speech intelligibility (after speech training) only for the older students of her sample (12,6 to 15,3 years). For both age groups no relationships between degree of hearing loss and speech production skills (before and after speech training) are reported.

Table 7 Mean and standard deviation of speech errors in the five categories

Error category	Mean	Standard deviation
Substitution	25.7	21.1
Omission	3.7	4.4
Addition	11.8	9.8
Phonation	4.1	5.0
No classification	2.4	3.5

Assessment of Students' Achievement

The following results are based on the subsample of 17 students (see Table 2). As Table 7 shows, two of the most frequent errors during the training belong to the categories substitution and addition (obtained from phonetic transcriptions by the teachers). Correcting these error types seems to be an important task of speech training with visible speech.

Table 8 provides insight into individual achieve-

Table 8 Error profiles "articulatory deficits" and "speech deficits" for each student based on the frequencies of errors (in %) in 5 error categories

Student	Error profile	Substitution	Omission	Addition	Phonation	NC
K1	1	20.3	2.9	10.1	0	0
K2	1	17.6	1.8	12.3	3.5	3.5
K3	1	7.0	0	5.6	0	0
K4	1	24.2	0	9.1	1.5	0
K5	1	14.1	6.3	3.1	1.6	0
K6	1	5.6	0	22.5	1.4	0
K7	1	6.7	0	1.7	0	0
K8	2	49.1	5.5	12.7	5.5	1.8
K9	2	30.8	11.5	40.4	13.5	0
K10	2	71.2	1.9	17.3	1.9	7.7
K11	2	69.2	3.9	15.4	17.3	1.9
K12	2	37.5	8.3	18.8	2.1	12.5
K13	2	38.8	15.3	16.5	10.6	2.4
K14	1	3.5	1.2	3.5	3.5	0
K15	1	18.6	2.3	4.7	4.7	3.5
K16	1	14.9	1.1	4.3	2.1	6.4
K17	1	8.5	1.1	2.1	1.1	1.1

NC = no classification.

ments in five error categories. There appear to be two different error profiles. Subjects who fit the "articulatory deficits" profile typically substituted (category 1) and added (category 3) phonemes. They showed mainly articulatory errors; characteristics of the voice were not strongly affected. Students who fit the "speech deficits" profile, on the other hand, showed deviations in all categories. This means they made articulatory errors, and they showed numerous phonatory deficits or nonclassifiable speech errors.

For analysis of students' speech performance after training, we computed a MANOVA with a within-subject design. The number of error corrections served as the dependent variable and the type of correction (not corrected [$M = 5.0$, $SD = 8.3$], partly corrected [$M = 28.6$, $SD = 14.2$], and corrected word [$M = 66.4$, $SD = 20.1$]) served as within-subject factor. There was a significant effect, $F(2,32) = 48.33$; $p < .01$, for type of speech correction. This means that training words were more likely to be corrected completely than partly or not corrected at all.

For the analysis of the stability of speech performance a MANOVA with within-subject design was computed. The number of habitual corrections served as the dependent variable. The type of habitual corrections (nonstable [$M = 3.0$, $SD = 5.7$], partly stable [$M = 22.0$, $SD = 13.0$], stable [$M = 30.5$, $SD = 23.6$]) served as the within-subject factor. A significant effect of the type of stability could be found, $F(2,32) = 12.85$; $p < .01$. Training words were more likely to be stably corrected than partly stable or nonstable, which means that former speech errors occurred less frequently.

Table 9 Correlations between the number of corrected or not corrected words or the number of stable or not stable corrected words and the degree of hearing losses at 500 Hz, 1 kHz, 2 kHz with hearing aid ($n = 17$)

Training words	500 Hz	1 kHz	2 kHz
Stable corrected	−.12	.53*	.27
Corrected	−.46	−.08	−.32
Not stable corrected	.26	.03	.20
Not corrected	.54*	.09	.29

*$p < .05$.

Correlational analysis (Table 9) indicated that the frequency of uncorrected speech errors correlated significantly with the degree of hearing loss at 500 Hz (with amplification) ($r = .54$). That means an effect of residual hearing (at 500 Hz) on learning achievement could be found. On the other hand, there was a positive correlation between the degree of hearing loss at 1000 Hz (with amplification) and the number of stable corrections ($r = .53$). Habitual speech correction (measured over a period of two training sessions) seemed

Table 10 Mean and standard deviation of speech correction in 9 subsequent training stages (aggregated into 3 levels)

Training stage	Mean	Standard deviation
L1–L3	61.9	24.5
L4–L6	61.5	33.6
L7–L9	78.9	40.8

Table 11 Number of subjects with "no speech errors, complete correction, partial correction, and no correction of speech errors" at the successive training stages (L1 to L9)

Training stages	No speech errors	Complete correction	Partial correction	No speech correction
L1	0	7	7	3
L2	1	6	8	2
L3	2	7	4	4
L4	4	2	10	1
L5	3	7	4	3
L6	2	8	4	2
L7	5	6	5	1
L8	3	9	2	3
L9	0	7	0	3

not to be positively related to residual hearing abilities. No significant correlations could be found with hearing loss measured at 2 kHz. We therefore concluded that there is no substantial influence of residual hearing abilities on learning achievements.

No significant differences could be found between the two error profiles "articulatory deficits" ($M = 51.2$, $SD = 14.9$) and "speech deficits" ($M = 54.3$, $SD = 17.8$) with respect to the number of speech correction. Independent of the given error type, speech corrections occurred with the help of the visual speech aid.

Furthermore, we examined the error correction depending on the training stage using a MANOVA. In nine succeeding training stages, students learned words with a successively increasing phoneme repertoire. The visual discrimination and the articulatory difficulty of training words increased step by step. The repeated measure factor "training stage" was aggregated to three levels. This was necessary to ensure that there were enough cases for each level. The number of complete word corrections served as the dependent variable. No effect could be found. There was no significant influence of the within-subject factor training stage on speech correction, although the articulatory and visual difficulty of the training words increased with training stage. There was even a small training effect. Articulatory errors concerning the more difficult words of training stages 7 to 9 were more often completely corrected than errors during the first training stages where easier training material was used (Table 10).

The numbers of subjects who had no speech errors, complete corrections, partial corrections, and no corrections of speech errors was counted at each training stage. The results are presented in Table 11. There were training stages in which no speech errors occurred among some students. The majority of students tended to make more complete word corrections than partial word correction or no correction with increasing training stage. Detailed analyses of error patterns indicated that /h/ as initial sound and the substitution of nasals with stops was difficult to correct. More complete corrections occurred with additions of sounds, voicing deficits, and impaired loudness control.

Discussion

The aim of the present study was to examine the use of spectrographic feedback for the improvement of speech production skills of deaf children and adolescents. Teachers and students of five German schools for the deaf participated. The investigation focused on two major issues: teachers' acceptance and implementation of the program and the individual achievements of the students evaluated by the teachers.

Our experiences with instructing 16 teachers for the deaf showed mostly positive results. After a short instruction period the teachers were able to make use of the visual speech aid for the speech training of deaf students of different grade levels. It was not difficult for the teachers to use the device technically. They were specially instructed and supervised to segment and label visible speech patterns (see Ruoss & Schildhammer, 1993). Although they had problems in segmenting and labeling the spectrograms, as well as interpreting speech errors during the initial stages, their proficiency increased rapidly with practice.

The most preferred condition for using the device in their lessons was teaching with two teachers. One teacher could concentrate on speech training with a single student, while the other teacher carried on teaching the rest of the class. Nearly all teachers thought the visual speech aid was good for articulation improvement (detailed analysis of learning achievements are given later) but less useful for the initial acquisition of sounds. One third of the teachers saw it as an effective instrument for the assessment of speech deficits. The majority of teachers thought that the device is compatible with different training concepts and could be combined with other techniques and manual communication methods. This means that the visual speech device could be used as a supplementary aid within a multisensory approach to speech instruction. Teachers judged that the students' acceptance and motivation was positive. They concluded that the visual speech aid was a convenient tool with a strong motivating effect in speech instruction.

Taking into account the different age groups, the results of the teachers' questionnaires revealed that the training program should have focused more strongly on individual speech performance of preschool and elementary school children, particularly in terms of different phonemes and words that they are able to use. The teachers suggested that the training program should be kept flexible to take the different needs of the children into account. With the older students there were no problems in training them in accordance with the overall program. The systematic and structured material was judged to be very helpful for speech instruction with students of higher grades.

A detailed analysis of learning achievements in individual training sessions was conducted with the older students only. Only with them had it been possible to keep the learning conditions constant. Learning achievements were measured according to the number and type of speech errors and corrections. Furthermore, the degree of hearing loss as well as the frequency of training sessions and the training stage were taken into account as factors influencing the learning process.

Substitutions and additions of phonemes were the most frequent error categories. It was possible to categorize two error profiles based on an analysis of the frequency of different error types for each student. Students with the profile "articulatory deficits" mostly made errors in the area of articulation. Students with the profile "speech deficits" showed phonatory or nonclassifiable errors in addition to errors in articulation. Insofar as the training program focuses particularly on articulatory features, we expected that students with both articulatory and phonatory deficits would profit less from our training program than those with articulatory deficits only.

The type of error profile had unexpectedly no effect on correction and constancy of speech production. Training words were more likely to be completely corrected by the visual speech aid than partially or not corrected at all. Speech correction was very effective in most cases and students did not repeat the same errors. They had learned to correct them within one lesson. Correlational analysis between degree of hearing loss, speech corrections (in one training session), and habitual corrections (over two successive training sessions) indicated that the visual speech aid could be used successfully by individuals having different amounts of residual hearing. Similar results were found in the clinical study of Schildhammer (1996). There was no significant correlation between the degree of hearing loss and intelligibility after speech training with the visual speech aid.

There was also no significant effect of the training stage reached on speech correction. It was even possible to find a small training effect. More corrections of speech errors occurred during the advanced than during the initial training stages despite the more difficult speech material used. Analyses of speech errors revealed that some specific errors were less frequently corrected (e.g., articulatory problems with /h/ as the initial sound, substitutions of nasals with stops) than others (like additions of sounds, voicing deficits or impaired loudness control). Therefore, it seems to be important to focus on individual error profiles in the training program. Detailed analyses of individual functional deficits are the prerequisites for goal-oriented therapeutical actions. Further research, therefore, should address individual error profiles that could serve as a basis for a single-subject-oriented speech training. One interesting approach in this regard was proposed by Ziegler and Hartmann (1993) for the assessment of dysarthria. They developed a computerized intelligibility test that allows a functional analysis of speech errors by specific phonetic structured items. The instrument can also be used for assessment of the speech production of deaf children.

In conclusion, results of this study indicate that a visual speech device with accompanying training program can be an effective supplementary tool for speech training and assessment. The particular training program examined here could be improved by being tailored to individual and age-dependent speech performance and also by teaching the symbolic meanings of words with the help of signs and written words. Such a program could be readily integrated into the speech lessons used in programs for deaf children.

References

Arends, N., Povel, D.-J., van Os, E., Michielsen, S., Claasen, J. & Feiter, I. (1991). An evaluation of the Visual Speech Apparatus. *Speech Communication, 10,* 404–414.

Becker, R. (1993). *Sprechfertigkeit und Verständlichkeit gehörloser Sprecher: Möglichkeiten der Fehlerdarstellung über eine Lautsprachvisualisierung* [Speech production skills and intelligibility of deaf speakers. Display of speech errors with visible speech]. Frankfurt/M: Lang.

Becker, R., & Druba, M. (1994). Visualisierung von fehlerhafter Aussprache bei gehörlosen Sprechern [Visualization of speech errors of deaf speakers]. In: G. Kegel, T. Arnhold, K. Dahlmeier, G. Schmid, & B. Tischer (Eds.), *Sprechwissenschaft & Psycholinguistik,* Vol. 6, (pp. 67–89). Opladen: Westdeutscher Verlag.

Becker, R., Schildhammer, A. & Ruoss, M. (1994). *Sprachliche Leistungen von gehörlosen Kindern und Jugendlichen beim mündlichen und schriftlichen Erzählen einer Bildergeschichte* [Spoken and written language skills of deaf children and adolescents]. Zeitschrift für *Experimentelle und Angewandte Psychologie, 41,* 349–377.

Bernstein, L. E., Goldstein, M. H., & Mahshie, J. J. (1988). Speech training aids for hearing-impaired individuals: I. Overview and aims. *Journal of Rehabilitation Research and Development, 25,* 53–62.

Breiner, H. L. (1990). Untersuchungen zur vibrotaktilen Wahrnehmung von Sprachlauten an der Körperoberfläche eines Sprechers [Studies on the vibrotactile perception of speech sounds at the body surface of a speaker]. *Sprache-Stimme-Gehör, 14,* 18–25.

Cole, R. A., Rudnicky, A. I., Zue, V. W., & Reddy, R. D. (1980). Speech as patterns on paper. In R. A. Cole (Ed.), *Perception and production of fluent speech* (pp. 3–49). Hillsdale, NJ: Lawrence Erlbaum Associates.

Conrad, R. (1979). *The deaf schoolchild.* London: Harper and Row.

Esser, G., Nolte, P., & Printzen, R. (1983). Verbesserung der Sprachentwicklung und Artikulation durch visuelle Übermittlung von Sprache (Sprach-Farbbild-Transformation) [Improvement of speech development and articulation through visualization of speech]. *Pädaudiologie aktuell* (pp. 185–193), Mainz: Krach.

Hobohm, K. (1993). *Verfahren der Spektralanalyse und Mustergenerierung für die Realzeitvisualisierung gesprochener Sprache* [Speech analysis and speech visualization in real-time]. PhD Dissertation, Technical University Berlin.

Klajman, S., Huber, W., Neumann, H., Wein, B., & Böckler, R. (1988). Ultrasonographische Unterstützung der Artikulationsanbahnung bei gehörlosen Kindern [Ultrasonography in speech training of deaf children]. *Sprache-Stimme-Gehör, 12,* 117–120.

Kloster-Jensen, M., & Jussen, H. (1974). Lautbildung bei Hörgeschädigten: Abriß einer Phonetik. [Acquisition of speech sound gestures by hearing-impaired individuals: a phonetic review]. In: H. Jussen (Hrsg.), *Schriften zur Hörgeschädigtenpädagogik* (Vol. 3), (pp. 1–216). Berlin: Marhold.

Levelt, W. J. M. (1989). *Speaking. From intention to articulation.* Cambridge: MIT Press.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1968). Why are speech spectrograms hard to read? *Amererican Annals of the Deaf, 113,* 127–133.

Lindner, G. (1984). *Entwicklung von Sprechfertigkeiten* [Acquisition of speech production skills]. Berlin: Volk und Wissen.

Lippmann, R. P. (1982). A review of research on speech training aids of the deaf. In: N. J. Lass (Ed.), *Speech and language: Advances in basic research and practice* (Vol. 16), (pp. 105–133). New York: Academic Press.

Mahshie, J. J., Vari-Alquist, D., Waddy-Smith, B., & Bernstein, L. E. (1988). Speech training aids for the hearing-impaired individuals: III. Preliminary observations in the clinic and childrens' homes. *Journal of Rehabilitation Research and Development, 25,* 69–82.

Markides, A. (1970). The speech of deaf and partially-hearing children with special reference to factors affecting intelligibility. *British Journal of Disorders of Communication, 5,* 126–140.

McAnnaly, P. L., Rose, S., & Quigley, S. P. (1987). *Language learning practice with deaf children.* Boston: Little, Brown and Company.

Neppert, J. M. H. (1990). Der Computer als Lernhilfe in der Hör-, Abseh- und Sprecherziehung. [The computer as training tool for hearing, lip reading and speech production]. In: H. Jussen & H. Claußen (Eds.), *Chancen für Hörgeschädigte: Hilfen aus internationaler Perspektive* (pp. 346–352). München: E. Reinhardt.

Pratt, S. R., Heintzelmann, A. T., & Deming, S. E. (1993). The efficacy of using the IBM SpeechViewer vowel accuracy module to treat young children with hearing impairment. *Journal of Speech and Hearing Research, 36,* 1063–1074.

Prillwitz, S., & Wudtke, H. (1988). *Gebärden in der vorschulischen Erziehung gehörloser Kinder.* [The use of signs in preschool education of deaf children]. Hamburg: Verlag hörgeschädigte Kinder.

Ruoss, M. (1992). Verständlichkeitsverbesserungen bei Gehörlosen [Improvement of intelligibility of the deaf]. *Hörgeschädigtenpädagogik, 46,* 362–378.

Ruoss, M., & Schildhammer, A. (1993) Sprachwissen und Lernen von Mustern visualisierter Lautsprache bei gehörlosen Schülern und Experten [Language competence and learning of visible speech patterns. Comparison of deaf students and experts]. *Sprache & Kognition, 12,* 115–121.

Schildhammer, A. (1996). *Einfluß eines Trainings mit Lautsprachvisualisierung auf die Sprechverständlichkeit gehörloser Kinder* [Influence of a speech training with visible speech on the intelligibility of deaf children]. Frankfurt/M: Peter Lang.

Schildhammer, A., & Becker, R. (1993). *Trainingskonzept für den Einsatz der Lautsprachvisualisierung* [Training concept for the application of visible speech]. Unpublished manuscript, Technical University Berlin.

Schönle, P. W., Gräbe, K., Wenig, P., Höhne, J., Schrader, J., & Conrad, B. (1987). Electromagnetic articulography: Use of alternating magnetic fields for tracking movements of multiple points inside and outside vocal tract. *Brain and Language, 31,* 26–35.

Stewart, L. C., Larkin, W. D., & Houde, R. A. (1976). *A real time sound spectrograph with implication for the speech training for the deaf.* IEEE International Conference on Acoustics, Speech and Signal Processing, Philadelphia, 590–593.

Uden, A. M. J. van (1977). *A world of language for deaf children.* Amsterdam: Swets & Zeitlinger.

Watson, C. S., & Kewley-Port, D. (1989). Advances in computer-based speechtraining: Aids for the profoundly hearing impaired. *Volta-Review, 91,* 29–45.

Watson, C. S., Reed, D. J., Kewley-Port, D., & Maki, D. (1989). The Indiana Speech Training Aid (ISTRA) I: Comparison between human and computer-based evaluation of speech quality. *Journal of Speech and Hearing Research, 32,* 245–251.

Weisenberger, J. M. (1898). Evaluation of the Siemens Minifonator vibrotactile aid. *Journal of Speech and Hearing Research, 32,* 24–32.

Wisch, F. H. (1990). *Lautsprache und Gebärdensprache* [Spoken and sign language]. Hamburg: Signum.

Ziegler, W., & Hartmann, E. (1993). Das Münchner Verständlichkeits-Profil: Untersuchungen zur Validität und Reliabilität [The Munic-Intelligibility-Profile: Investigation into validity and reliability]. *Nervenarzt, 64,* 653–658.