ISCA Archive
http://www.isca-speech.org/archive

3rd International Conference on
Spoken Language Processing (ICSLP 94)
Yokohama, Japan
September 18-22, 1994

PERCEPTION, PRODUCTION AND TRAINING OF NEW CONSONANT CONTRASTS IN CHILDREN WITH ARTICULATION DISORDERS

Donald G. Jamieson and Susan Rvachew

HHCRU, University of Western Ontario, London, ON, Canada N6G 1H1 and
Speech Dept., Alberta Children's Hospital, Calgary, AB, Canada T2T 5C7

ABSTRACT

A substantial proportion of children who have normal oral and motor function, nevertheless demonstrate difficulty producing one or more sounds in their native language. Assessment and treatment of such children has traditionally focussed on speech production to the exclusion of speech perception. The possibility that speech production difficulties might be associated with atypical speech **perception** abilities has been discounted in traditional therapeutic approaches. In contradiction, we find that a subset of children who have a functional articulation disorder have a concurrent and correlated perceptual disorder and that in some children, a formal computer-based training technique produces striking improvements in speech perception abilities, and improvements in the quality of produced speech as well. A new computer-based system for IBM-PC compatible computers has been developed to implement this training strategy.

I. INTRODUCTION

While adults may experience great difficulty learning to identify non-native speech contrasts, young infants appear particularly able to discriminate both native and non-native speech sounds [21]. Over time, normally developing infants and young children demonstrate a growing facility with the speech sounds of their own language, coupled with a reduction in their ability to discriminate non-native speech sounds [20]. Speech production typically shows a parallel improvement, with both speech perception and speech production abilities reaching maturity at approximately eight years of age.

In contrast with this typical developmental pattern, children with impaired phonological skills misarticulate certain speech sounds within their native language, at ages when such misarticulations are not considered to be developmentally appropriate. Listening errors are correlated with these production errors: a substantial proportion of children who are identified as having a phonological disorder also have a perceptual difficulty, when tested using appropriate techniques [18]. The existence of this perceptual difficulty suggests that conventional speech therapy techniques which focus on **speech production** only may be highly suboptimal for at least a subgroup of children requiring treatment. Rather, the effectiveness of speech production training may be enhanced when combined with a structured program of perceptual assessment and training.

Unfortunately, the conventional view in North American Speech-Language Pathology has been that speech perception abilities have little relevance for the design of treatment programs for children. This view appears to be based on earlier findings with clinical tests requiring children to make "same/different" judgements about naturally-produced minimal pair words [4,16] or to identify a spoken word by pointing to one of several pictures. Unfortunately, these standardized assessment tools do not focus on the speech sound contrasts that are most relevant to children's typical speech errors [13]. For example, some tests require the child to discriminate /r/ vs. /l/, but not /w/ vs. /l/ — which is the more common substitution error for /l/, in children's speech. In fact, most common childhood speech sound errors are not targeted by any commonly-used test of speech perception ability.

In consideration of these facts, more than a decade ago, Locke [13] recommended certain nonstandardized procedures involving live-voice, therapist-produced imitations of the child's speech error [15]. However, this recommendation has had no apparent impact on contemporary Speech-Language Pathology practice. Moreover, it is now recognized that the procedures described by Locke can be further improved by increased control over stimulus characteristics. Using well-controlled stimuli which vary systematically with respect to the relevant acoustic cues permits the perceptual abilities be assessed in detail. Such studies have revealed speech perception to reflect a complex interaction between language experience and the acoustic details of the speech signal [12, 22].

Studies of the perceptual abilities of misarticulating children have **failed** to demonstrate a relationship between perceptual and productive ability in these children when testing involved unmodified natural speech stimuli [1]. On the other hand, studies employing synthetic speech stimuli have identified a clear relation between perceptual difficulties and production [3]. For example, some [r] misarticulating children have significant difficulty with the perception of the /r/-/w/ contrast, and these children appear to have inappropriate category boundaries for the F2 onset frequency cue.

In a study of fricative perception, Rvachew & Jamieson [19] found that normal speaking five year olds were able to identify sounds using the categories assigned by normal speaking adults. However, phonologically-delayed five year old children displayed abnormal functions, on average, reflecting widespread individual differences in identification abilities. Thus, the mean identification function obtained from the phonologically delayed group was not representative of the performance of the individual subjects. While some phonologically delayed children were able to identify these stimuli appropriately, more than half the subjects had considerable difficulty. For example, one child identified all of the stimuli as belonging to the "seat" category. His productions of both the /s/ and /ʃ/ categories were very inconsistent, with [s], [ʃ], and various distortions occurring for both sounds. Rvachew & Jamieson [19] found a similar pattern of results for [s] vs [θ]. Here **none** of the misarticulating children were able to identify synthetic tokens of the words "sick" and "thick", while normal children and adults had no difficulty. Studies such as these suggest that at least a subgroup of children who have difficulty producing sounds in their native language have significant difficulties identifying these sounds. Either they fail to perceive that the target and their error sound belong to two distinct categories, or they categorize the target and the error in nonstandard ways. Such findings suggest that speech perception training might facilitate phoneme production learning by these children

II. SPEECH PERCEPTION TRAINING

To date, there have been few studies of the effectiveness of speech perception training for children with speech problems. However a substantial body of empirical work demonstrating the value of speech perception training with adult second language learners. Moreover, useful principles have been identified which predict the degree of success experienced with such perceptual training [2,5,9]. These principles also provide guidance for the design of perceptual training procedures for children.

Flege [5] emphasizes that adult L2 learners have superior articulatory control and a more stable underlying phonetic system for producing speech, than do young children learning their first language. However, the processes underlying the language learning task are also similar for these two groups. Children who correctly articulate a given contrast also perceive the contrast correctly, although children who correctly perceive a contrast may misarticulate it [19]. Some children appear to attend to irrelevant acoustic variation when attempting to categorize stimuli which represent a glide-liquid contrast [6]. Finally, children's L1 productions are related to their categorizations of misarticulated sounds [19].

We believe, therefore, that there are three patterns of perceptual and productive error which are shared by second language learners and phonologically impaired children. (1) Neither member of the contrast is present in the underlying system; (2) Both members of the phoneme pair belong to a single category in the underlying system; and (3) Both members of the contrasting pair exist as separate categories in the underlying system but they are differentiated in terms of nonstandard cues, or nonstandard values of a standard cue. Details and examples in support of this view are available in Rvachew and Jamieson [19].

Two methods have been demonstrated to be effective in training listeners to hear new speech contrasts. Jamieson and Morosan [9] applied a modified "perceptual fading" technique with synthetic speech signals, to teach Canadian Francophone adults to hear the English voiced-voiceless "th" sounds. At the beginning of training, listeners were presented with extreme exemplars containing abnormal amounts of voiced or voiceless frication; this step was designed to draw the listener's attention to the cue most relevant to the speech contrast. Training began with the cue exaggerated in this way, with the cue gradually "fading" to more normal values, over time. Training the distinction in this way, led to a high level of identification accuracy, with few listener errors. Such training also generalized to new voices, and new words.

The second method which is demonstrably effective for training new speech contrasts involves the use of multiple tokens of natural speech tokens. For example, Logan, Lively and Pisoni [14] demonstrated that native speakers of Japanese could be trained to identify English /r/ and /l/ sounds, through a program involving words spoken by several different talkers, with accuracy feedback following each response. Such training generalizes to new talkers and to situations when the target sounds occur in word positions that were not encountered during training.

In summary, two techniques have been shown to be highly effective in training listeners to identify speech sounds. Both involve a structured program of identification training, using multiple signals with accuracy feedback provided to the listener, following each response. However, to date, neither approach has been shown to improve the **production** of foreign language speech sounds.

III. SPEECH PERCEPTION TRAINING WITH CHILDREN

As noted above, while it is likely that many children who display difficulties when acquiring their native language have a speech perception disorder [18], formal speech perception assessment and training has received little emphasis by clinicians and researchers. To determine whether an appropriately-designed program of perceptual training would be effective with children who showed both speech perception and speech production difficulties, Jamieson and Rvachew [11] applied a procedure similar to Jamieson and Morosan [9] with adult listeners. Using synthetic speech sounds within a fading technique with feedback, children aged 5 to 7 years were taught to identify English-language fricatives. For children who misarticulated /ʃ/, the training stimuli were four synthetic CV syllables that sounded like "sa" or "sha". A set of 4 "sa" and "tha" stimuli were created in a similar fashion for use with children who misarticulated /s/. During the first treatment session, each child was trained to identify the most extreme pair of stimuli from the synthetic continuum. Less extreme stimuli were then introduced one at a time, contingent upon the child's performance at each phase of training. A set of naturally produced stimuli contrasting the target with stop and glide consonants was used during control sessions. The child's task was to identify each stimulus by pointing to a happy face when the target was heard and to a picture of a sad face when a nontarget sound was heard. Feedback was provided for each response during both control and treatment sessions.

As expected, three children who demonstrated poor identification performance prior to training, showed improved identification performance following training. However, in addition to improved speech perception, speech production abilities also improved for these children, even though speech production was not specifically trained. For the fourth child, speech perception performance did not improve and (not surprisingly therefore) speech production performance also remained unchanged. These results suggest that at least some children who have both a perception and a production difficulty can benefit from a structured program of speech perception training, whether or not they receive speech production training.

In many instances it may be more practical to use multiple tokens of natural speech signals in training. Rvachew's [17] study was thus undertaken with natural speech signals, to study a larger sample of children. Procedural changes from the Jamieson and Rvachew [11] study included: (1) children were completely unstimulable for correct /ʃ/ production in any context; (2) natural speech stimuli were used so that the target sound could be contrasted with a broader range of error sounds; and (3) training occurred within the context of a video game, to provide more compelling reinforcement for correct responding.

Twenty-seven phonologically impaired children, aged from 42 to 66 months (M = 53) were assigned to one of three training groups. Each child received six weekly treatment sessions consisting of 60 perception training trials followed by 60 production training trials. During speech perception training, Group 1 children listened to multiple exemplars of correct and incorrect versions of the word "shoe" which were recorded from children and adults. The correct tokens varied considerably in quality. The incorrect tokens contained both distortion and substitution errors for the initial consonant combined with correct vowel production (e.g., [su], [tsu], [du] etc.). Group 2 children listened to the words "shoe" and "moo" while Group 3 children listened to the words "cat" and "Pete".

Traditional sound production training was used to teach the children correct articulation of [ʃ] in isolation, syllables, words, and sentences. Before and after training, children were required to identify correct and incorrect versions of the word "sheet", without feedback, and each child produced a single [ʃ] sound in isolation, and named five objects representing words containing a prevocalic [ʃ].

The effects of training were marked: children trained to identify correct and incorrect versions of the words "shoe" progressed further in production training, produced more correct [ʃ] sounds during a spontaneous naming task, and produced better quality [ʃ] sounds with more appropriate acoustic characteristics, in comparison to control children who learned to identify the words "cat" and "Pete". Group 1 and 2 production performance was significantly better than that of Group 3, but the production performance of children did not differ statistically for Groups 1 and 2. These results provide further evidence of the potential contribution of a structured program of perceptual training for at least some children who have difficulties producing sounds in their native language.

IV. TRANSFERRING TRAINING TECHNOLOGIES TO THE CLINIC

The research studies reviewed above make it clear that a structured program of speech perception training can be useful for both adult second language learners and when treating young children who have articulation disorders. **Unfortunately, such efficacy demonstrations are not sufficient to ensure that techniques are applied clinically.** In part, this is because work done within relatively well-funded, "state of the art" research laboratories, may ignore the clinical realities of cost and time restrictions. Adapting even a highly-refined experimental test for clinical use often requires the original application to be repackaged and recoded, to make it more acceptable to clients, easier to use, more time-efficient, and more consistent with existing clinical practice [8]. Fortunately, the hardware needed to apply such techniques in clinical practice is inexpensive and widely available, and the associated behavioral technologies are now being developed and disseminated [7].

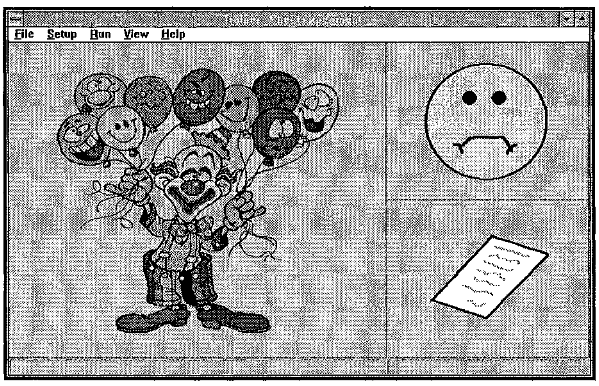

Figure 1. Sample screen from prototype speech perception trainer showing "feedback" cartoon and "response" images.

To facilitate clinical applications, we have prepared an expanded implementation of the system described by Rvachew [17], for use by speech-language pathologists. The new system contains an expanded and newly-recorded set of CD-quality speech stimuli together with colorful graphic images, within a new control program derived from our Computerized Speech Research Environment software [10]. Stimuli are presented using a high-quality, but inexpensive "multi-media" sound card, while the computer display screen provides motivating feedback images and is used to record the child's responses (see Figure 1). Feedback is provided by altering the display in the left portion of the screen for instance, a balloon may be added each time the child responds correctly. We have found such displays to be highly motivating for most children leading to favorable clinical impact.

ACKNOWLEDGEMENTS

*Work was supported by NSERC, MRC and the MSI Foundation. We are grateful to K Ramji, T Schneider and L Kieffer who assisted with various aspects of this project. Correspondence to Dr. DG Jamieson, Hearing Health Care Research Unit, The University of Western Ontario, London, ON N6G 1H1 (jamieson@uwovax.-uwo.ca).

REFERENCES

[1] Aungst LF & Frick JV "Auditory discrimination ability and consistency of articulation of /r/", *J of Speech & Hearing Disorders, 29*(1):76-85, 1964.

[2] Best CT, McRoberts GW & Sithole NM "Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants", *J of Experimental Psychology: Human Perception and Performance, 14* (3):345-60, 1988.

[3] Broen P, Strange W, Doyle S & Heller JH "Perception and production of approximant consonants by normal and articulation delayed preschool children", *J of Speech & Hearing Research, 26*(4):601-8, 1983.

[4] Cohen JH & Diehl CF "Relation of speech-sound discrimination ability to articulation-type speech defects", *J of Speech & Hearing Disorders, 28*(2): 187-90, 1963.

[5] Flege JE "Speech learning in a second language", in: Ferguson CA, Menn L & Stoel-Gammon C (eds) *Phonological development: Models, research, implications*, Timonium, MD: York Press, 1992.

[6] Hoffman PR, Stager S & Daniloff RG "Perception and production of misarticulated /r/", *J of Speech & Hearing Disorders, 48*(2):210-15, 1983.

[7] Jamieson DG "The impact of emerging technologies on Audiology and Speech-Language Pathology", *J of Speech Language Pathology & Audiology, 14*:19-26, 1990.

[8] Jamieson DG, Cheesman MF & Corbin H "Clinical evaluation of an automated technique for estimating speech reception thresholds", *J of Speech-Language Pathology & Audiology, 18*:23-8, 1994.

[9] Jamieson DG & Morosan DE "Training non-native speech contrasts in adults: Acquisition of the English /ð/-/θ/ contrast by francophones", *Perception & Psychophysics, 40*(4):205-15, 1986.

[10] Jamieson DG, Ramji K, Kheirallah I & Nearey TM "CSRE: A speech research environment", in: Ohala JJ, Nearey TM, Derwing BL, Hodge MM & Wiebe GE. *Proceedings, Second International Conference on Spoken Language Processing*, Edmonton: U. of Alberta (Personal Publishing Ltd.), 1127-30, 1992.

[11] Jamieson DG & Rvachew S "Remediating speech production errors with sound identification training", *J of Speech-Language Pathology & Audiology, 16*(3):201-10, 1992.

[12] Lisker L & Abramson AS "The voicing dimension: Some experiments in comparative phonetics", *Proceedings, Sixth International Congress of Phonetic Sciences*, Prague, 1967. Prague: Academia.

[13] Locke JL "The inference of speech perception in the phonologically disordered child. Part I: A rationale, some criteria, the conventional tests", *J of Speech & Hearing Disorders, 45*(4):431-44, 1980.

[14] Logan JS, Lively SE & Pisoni DB "Training Japanese listeners to identify English /r/ and /l/: A first report", *J of the Acoustical Society of America, 89*(2):874-86, 1991.

[15] McGregor KK & Schwartz RG "Converging evidence for underlying phonological representation in a child who misarticulates", *J of Speech & Hearing Research, 35*:596-603, 1992.

[16] McReynolds LV, Kohn J & Williams GC "Articulatory-defective children's discrimination of their production errors", *J of Speech & Hearing Disorders, 40*(3):327-37, 1975.

[17] Rvachew S "Speech perception training can facilitate sound production learning", *J of Speech & Hearing Research, 37*:347-57, 1994.

[18] Rvachew S & Jamieson DG "Perception of voiceless fricatives by children with a functional articulation disorder", *J of Speech & Hearing Disorders, 54*(2):193-208, 1989.

[19] Rvachew S & Jamieson DG "Learning new speech contrasts: Evidence from adults learning a second language and children with speech disorders", in: Strange W (ed) *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-Language Speech Research.* Timonium, MD: York Press (in press).

[20] Werker JF & Polka L "Developmental changes in speech perception: New challenges and new directions", *J of Phonetics, 21*:83-101, 1993.

[21] Werker JF & Tees RC "Cross-language speech perception: Evidence for perceptual reorganization during the first year of life", *Infant Behavior and Development, 7*(1):49-63, 1984.

[22] Yamada RA & Tohkura Y "The effects of experimental variables on the perception of American English /r/ and /l/ by Japanese listeners", *Perception & Psychophysics, 52*:376-92, 1992.