RESEARCH ON SPOKEN LANGUAGE PROCESSING
Progress Report No. 19 (1993-1994)
Indiana University

Talker Normalization in Normally Articulating and Phonologically Delayed Children: Methodological Considerations[1]

Karen Forrest,[2] Steven B. Chin,
David B. Pisoni, and Nancy Nelson Barlow[2]

Speech Research Laboratory
Department of Psychology
Indiana University
Bloomington, Indiana 47405

[1] This work was supported in part by NIH NIDCD Training Grant DC00012-15 and NIH Research Grants DC00111-17 and DC00260 to Indiana University at Bloomington. We are grateful to Margot Mason for assistance in gathering the data for this study.

[2] Department of Speech and Hearing Sciences, Indiana University at Bloomington.

Abstract

Although speech perception difficulties are suspected of contributing to phonological delays in children, research into the relation between perception and phonological impairments has been equivocal. The theoretical perspectives driving earlier investigations of speech perception in phonologically delayed children may help explain the ambiguity of the resulting data. Using a different perspective, it is hypothesized that variation in the acoustic signal associated with contextual changes and talker characteristics, which adversely affect adult listeners, are detrimental to the child's development of phonological models needed for correct speech production. In a series of experiments, word identification was tested in 4- and 5-year-old children with normal and delayed phonological development. Using the *Word Intelligibility by Picture Identification* (WIPI) test, each child was presented with a six-picture array and, after hearing a single-word token via headphones, responded by pointing to the picture of the stimulus item. All words were presented at 70 dB SPL and signal-to-noise-ratio was varied from quiet to +8 dB across the three experiments. Two lists of test words were presented from digital audio recorded tape; one list was spoken by a single talker and the second list had each of the 25 words spoken by a different talker. Significant differences were found on the single- versus multiple-talker lists for both the normal and phonologically delayed children. On average, the normally articulating children scored 12% higher on the single talker list, compared to performance in the multiple talker condition. Similar differences in word identification as a function of talker condition were found for the phonologically delayed children. Although there was no significant difference in performance between the phonologically delayed and normally articulating subjects, the phonologically delayed children did have reduced accuracy of word identification in both talker conditions compared to their normally articulating peers.

Talker Normalization in Normally Articulating and Phonologically Delayed Children: Methodological Considerations

Introduction

The causal bases of functional articulation impairments in children remain elusive, although a variety of factors have been suggested to underlie the disorder. For example, theorists have proposed that cognitive-linguistic limitations, neuromotor disabilities, psychosocial factors, and/or perceptual deficits (see Bernthal & Bankson, 1993 for a review) may be responsible for the development of disordered phonological systems. Repeated investigations have provided conflicting data about the contribution of these factors to speech articulation problems. For example, some children have been shown to have incorrect phonological representations of target sounds that they do not produce (Dinnsen, 1984; Gierut, Elbert & Dinnsen, 1987), whereas other researchers fail to establish such a link (Ingram, 1990). Children have been diagnosed to have developmental apraxia because of multiple, and often inconsistent, speech sound errors (Hall & Robin, 1993; Thoonen, Maassen, Gabreels, and Schreuder, 1994), whereas other analyses demonstrate little correspondence between movement deficits and articulation errors (Guyette & Diedrich, 1981). Speech perception problems, while implicitly related to articulation problems (e.g., speech in hearing impaired children), do not have a consistent relation to phonological delays. Studies that investigate the relation between speech perception and articulation errors find some correlation between these variables (Rvachew and Jamieson, 1989), but the relation usually holds only for a subset of articulation impaired children.

It is unlikely that a singular cause will account for all articulation disorders. However, the lack of strong theoretical and empirical motivations for investigations of factors underlying functional articulation impairments may account for our limited understanding of the causes. By way of example, speech perception deficits have been suggested to have a negative impact on phonological development (Travis & Rasmus, 1931; Winitz, 1984). Investigations have established a positive correlation between articulatory proficiency and speech discrimination (Sherman & Geith, 1967; Winitz, 1984); however, discrimination and articulation errors often do not coincide; misarticulated sounds often are perceived correctly and correctly produced sounds may be misperceived (Eilers & Oller, 1976).

Lack of correspondence between perceptual and productive errors in general discrimination tests has led to the development of phoneme-specific tests of production-perception links (Locke, 1980). Phoneme-specific tests of perception require a misarticulating child to distinguish a clinician's correct versus incorrect production of a word. These types of investigations generally indicate good discriminative abilities, even in children with severe articulation disorders. These types of studies have strong theoretical appeal; i.e., children develop articulation errors because they do not perceive speech sounds correctly. However, the empirical validity of these studies must be questioned. Functionally misarticulating children, even those with severe disorders, generally respond appropriately to verbal input. Whether playing with other children or following the dictates of a parent, phonologically delayed children appear to understand the messages relayed to them.

By contrast, empirically-based hypotheses about perception-production links during speech acquisition posit that phonological delays are linked to self-monitoring deficits. Citing evidence that children can make accurate judgements about sounds that are misarticulated, investigations have attempted to relate a child's own production of an error sound to his/her perception of that production. Specifically, did the child recognize his/her incorrect production of a sound? Results of these investigations (Lapko & Bankson, 1975; Shelton, Johnson & Arndt, 1977; Wolfe & Irwin, 1973) show only modest correlations between these measures of production and perception. Further, these correlations may relate to third-variable factors such as type of discrimination task employed and/or consistency of production errors (Shelton, et al., 1977).

Although a relationship between speech perception deficits and articulation impairment is logically compelling, the research data do not support a strong association. Demonstration of a link between phonological development and perceptual competence may depend on the adoption of a different theoretical approach, one that can be empirically supported. The present research is motivated by such a theoretical and empirical convergence. Based on principles of cognitive and motor development (e.g. Thelen & Smith, 1994), theoretical tenets of speech perception (Joos, 1948; Studdert-Kennedy, 1974, 1976) and results of empirical studies of word recognition (Creelman, 1957; Goldinger, Pisoni, & Logan, 1991; Martin, Mullennix, Pisoni, & Summers, 1987; Mullennix, Pisoni & Martin, 1989), it is suggested that the relation between articulation impairment and perceptual deficits should focus on talker normalization. Specifically, it is hypothesized that articulatory disorders may result from children's limited ability to extract linguistically relevant information from the linguistically irrelevant component of the speech signal.

A survey of the literature on speech acoustics and perception reveals the search for the invariant properties of the speech signal (Blumstein & Stevens, 1979; Joos, 1948; Kewley-Port, Pisoni, & Studdert-Kennedy, 1983; Stevens & Blumstein, 1978; Studdert-Kennedy, 1974, 1976). It is clear that the adult listener can understand messages produced by talkers of different dialects, different voice characteristics, and different vocal tract dimensions. Although the acoustic signal carries the variations that are unique to talkers, the adult listener can comprehend the spoken message. This remarkable ability of the human listener suggests that there is some constant within the speech signal that the perceiver is able to extract to understand the linguistic code. The nature of this constancy remains unclear; however, it has long been proposed that the adult listener must normalize the acoustic signal to extract the invariant properties of the linguistic system (Joos, 1948). This normalization process is costly and incomplete; research has shown decrements in word identification as a function of normalization demands (Goldinger, et al., 1991; Martin, et al., 1989; Mullennix, et al., 1989).

In a series of experiments, Pisoni and his colleagues (Goldinger, et al., 1991; Martin, et al., 1987; Mullennix, et al., 1989) have demonstrated that the adult listeners have an 8-10% decrement in word identification when talker normalization is required. Presented with a list of words that is spoken by a single talker and at a constant articulatory rate, adult listeners can identify more than 90% of the words correctly. However, performance falls to about 80% correct word identification if the list is comprised of successive tokens produced by different talkers. Similar performance decrements occur when successive words are presented at different speaking rates. The conclusion drawn from these experiments is that talkers do not discard information in the speech signal that is not linguistically important. Rather, this information about talker characteristics impacts perceptual processing in a way that degrades lexical identification. Although some process aids in phone recognition so that words can be identified and messages deciphered, normalization during speech perception is not complete.

Little is known about children's ability to normalize speech. Surely children are capable of normalization in that they can understand talkers who vary considerably in vocal tract and phonatory characteristics (e.g., men versus women). However, the sensitivity of children to variations in the speech signal has received little attention. The studies that have been undertaken (Oliver, 1989, 1990) show that 3-year-old children, like adults, have higher accuracy on word list identification when all words in the list are spoken by a single talker (single-talker condition), compared to a condition in which each word is produced by a different talker (multiple-talker condition). Four- and five-year-old children did not perform differently on the two talker conditions, mainly because of ceiling effects for both talker conditions. In an attempt to overcome these ceiling effects, Oliver (1990) conducted a second experiment in which word lists were presented to 3- to 5-year-old children in the presence of noise. Using a 0 dB signal-to-noise ratio, Oliver (1990) found that all age groups identified significantly more words correctly in the single talker condition, with the greatest effect of talker condition found in the 3-year-old subjects. From these data Oliver (1990) concludes that "not only is there a cost for talker normalization for pre-school children, but that it appears it has a higher cost [in] the younger children" (p. 380). This suggests that young children are less attuned to the linguistically relevant information in a signal and, perhaps, more distracted by talker-relevant information in speech than older children or adults.

As an extension of this hypothesis, it is suggested that phonologically delayed (PD) children who have phonemic repertoires that correspond to chronologically younger children, may have reduced abilities than their age-matched peers to normalize the speech signal. This failure to normalize, extract linguistic information and thereby develop categorical representations of speech sound, may be a relevant factor in the development of disordered phonological systems. This hypothesis was tested on children with severe phonological delays (i.e., more than six sounds in error across three manner categories on the *Goldman-Fristoe Test of Articulation*). Procedures and materials developed by Oliver (1990) were used in this experiment (see Methods, below) so that the words were presented to young PD children at 0 dB signal-to-noise ratio (SNR). Children were required to point to the picture of the word that was heard, from a display of six alternatives. Each child received two lists of words, one with a single talker for all words and a second list with each successive word produced by a different talker. The experimenter recorded the children's responses and determined the percent of responses that were correct. Data collected by Oliver (1990) represented the normal control group.

This experiment was short-lived. The first six phonologically delayed children that were investigated were unable to perform the task. When words were presented at 0 dB SNR, the phonologically delayed children correctly identified between 16 and 64% of the words on the single talker list and between 4 and 52% of the words on the multiple-talker list. The level of performance was at or below chance for half of the subjects tested. Before continuing with the examination of talker normalization in PD children, then, it seemed essential to verify that normally articulating children could perform the protocol in our laboratory. Initial results from normally articulating children were not encouraging when a signal-to-noise ratio of 0 dB was used. Therefore, a different protocol had to be developed. The remainder of this report describes the modifications made to this protocol, considerations that must be weighed when using children as subjects in speech perception experiments, and finally, some data from normal and phonologically delayed children in the performance of a task that was within their capabilities.

Experiment 1

As mentioned previously, presentation levels with a 0 dB signal-to-noise ratio appeared to render the perception task on the *Word Intelligibility by Picture Identification* test too difficult for the children being examined. Therefore, in Experiment 1, presentation of stimulus words on this task was in quiet. In addition, three screening examinations were employed to insure that the children being studied fell within normal limits, either as phonologically delayed children or as normally articulating children.

Methods

Subjects.

Subjects for Experiment 1 were 12 phonologically delayed children, ages 3 to 6 years, and 28 normally articulating (NA) children, ages 3 to 6 years. The NA children were recruited for the study through an advertisement in the local newspaper, and phonologically delayed children were identified by referrals from the Speech and Hearing Clinic at Indiana University. The newspaper call for NA children was a request for children between 3 and 6 years of age who were native speakers of English to participate in listening experiments in the Speech and Hearing Department at Indiana University. Subjects were offered free speech and hearing examinations, prizes, as well as monetary reimbursements to the parent for transportation.

Materials.

Both the normally articulating children and the phonologically delayed children were evaluated for their appropriateness for the study:

(1) Hearing sensitivity: A standard pure-tone audiometric screening at octave frequencies between 500 and 8000 Hz, with a threshold of 25 dB HL (re: ANSI, 1970)

(2) Articulation Development: The *Goldman-Fristoe Test of Articulation* (GFTA; Goldman & Fristoe, 1986). This screening procedure could determine assignment to either the normally articulating or phonologically delayed group.

3) Receptive Vocabulary: The *Peabody Picture Vocabulary Test-Revised* (PPVT; Dunn & Dunn, 1981). This was administered to determine if the children had age-appropriate vocabularies.

The experimental testing instrument employed in this study was the *Word Intelligibility by Picture Identification* (WIPI) test (Ross & Lerman, 1970, 1971), a 25-trial, 6-alternative forced-choice task. On each trial of this test, subjects were shown a card with six color drawings (three horizontally by two vertically); the test item and some of the foils shared a similar syllable nucleus or rhyme. Stimuli were presented through circumaural headphones, and subjects pointed to the picture corresponding to the word they heard. This closed-set forced-choice task was chosen for this study to avoid the potential problem of unintelligible or ambiguous verbal responses from children of the age tested, especially among those with phonological delay. The use of the picture-pointing task obviated the need and potential danger for subjective judgments on the part of the examiner.

Stimulus materials consisted of digital audio tape (DAT) recordings of two lists of 25 words spoken either by a single adult talker or multiple adult talkers. Single-talker stimulus sets contained a single male voice, and multiple-talker stimulus sets contained both male and female voices. Each of the two word-lists were spoken by both single and multiple talkers, so that the words appeared equally in each condition across the listeners.

Experimental Procedure.

The three screening tests and the WIPI test were all administered in a single session lasting approximately one hour. The order of test administration was as follows: (1) Audiometric screening, (2) WIPI, (3) GFTA, (4) PPVT. The GFTA was administered to any child who exhibited a hearing threshold greater than 25 dB SPL and then the child was released from further testing. If the child did pass the hearing screening, all three subsequent tests were administered. The experimental test, the WIPI, was administered immediately after the hearing screening in order to avoid possible attention effects (see Cantwell & Baker, 1985; Cantwell, Baker, & Mattison, 1981). Similarly, the GFTA was administered prior to the PPVT, because the former test required a more active response (i.e., saying a word aloud) than the latter (i.e., pointing to a picture).

All protocols of the study were conducted in the Speech and Hearing Sciences Department at Indiana University. Testing was administered in a sound-attenuated booth, and two examiners accompanied the child: the first presented stimulus materials, and the second recorded responses. Subjects were required to be accompanied by a parent, guardian, or other responsible adult, and such adults and other accompanying persons were permitted to view the test administration from an adjoining sound-attenuated observation booth separated from the experimental booth by sound-attenuating double-paned, transparent acrylic glass.

For administration of the WIPI, stimulus materials were played from a DAT recorder (Technics, SV-DA10) through an audiometer (Grason-Stadler, 1704) to insure a uniform presentation amplitude level of 70 dB SPL and into headphones (Telephonics, TDH50P) worn by the child. In Experiment 1, all presentations were in quiet. One examiner wore headphones and heard the stimuli to insure correct alignment of stimulus presentation with response set presentation. Additionally, the examiner stopped the tape between trials, and the subject was given an opportunity to respond. There were no temporal constraints on the response interval, although subjects were encouraged to respond quickly. Both a single-talker stimulus set and a multiple-talker stimulus set were presented to each child in counterbalanced order across subjects. Responses were recorded on score sheets by the second examiner. Children were presented with a sticker as a reward after completion of each of the two lists.

Subsequent to administration of the WIPI, the *Goldman-Fristoe Test of Articulation* (GFTA) was administered to assess subjects' articulatory development. The GFTA consists of 35 stimulus pictures containing all English consonants in a variety of word positions. Subjects were instructed, upon presentation of the stimulus picture by the first examiner, to say aloud the name of the object or action depicted. Responses in this test were recorded as phonetic transcriptions by the second examiner, a trained transcriber. In addition, all responses were recorded on audio tape for subsequent examination if questions arose regarding the on-line transcription.

Finally, the *Peabody Picture Vocabulary Test- Revised* (PPVT) was administered, in order to assess development of receptive vocabulary. The PPVT is a four-alternative forced-choice picture-pointing task, consisting of 175 response items. Targets are arranged in order of increasing semantic sophistication, and testing terminates when a specified criterion error pattern is reached. In Experiment 1, PPVT stimulus words were presented by live-voice in a sound field, and subjects were required to point to the picture corresponding to that word.

Results

All scores were converted to percent correct, and means were calculated across subjects in each age group. A two-way ANOVA indicated no significant main effects for age group (F=2.84; p=0.09) or talker condition (F=0.14; p=0.72). This is in contrast to the data presented by Oliver (1989) in which the 3-year-old children did significantly better on word identification in the single talker condition. As can be seen in Figure 1, 3-year-old PD children did not show differences in identification accuracy as a function of talker condition; approximately 60% (15/25) of the words were identified correctly in both the single- and multiple-talker conditions. The accuracy of the 3-year-old PD children on the multiple talker lists was comparable to the data reported by Oliver (1989); however, this youngest group of PD subjects scored about 20% below the level children in Oliver's study (1989) on the single-talker condition. Level of performance for the 4- and 5-year-old PD subjects in the present study were comparable to those reported by Oliver (1989) for normally articulating children of this age.

Insert Figure 1 about here

Normally articulating children between the ages of 3 and 6 years also participated in this study. Comparison of Figures 1 and 2 reveals comparable performance for the PD and NA children in each age group. As with the PD children, the NA children showed no effect of talker condition (F=0.04; p=0.84); the 3-year-old NA children identified about 60% of the words correctly in both talker conditions, and performance levels of the older NA children ranged from 70 to 92% correct word identification, a difference that was significant (F=3.67; p=0.02).

Insert Figure 2 about here

In an attempt to understand the poor word identification of the NA 3-year-old subjects, we surveyed their performance on the standardized articulation and receptive vocabulary tests. These data, presented in Table 1, indicate that the 3-year-old children in the NA group were in the lower half of the normal distribution for articulation. Recall, that these children were recruited by newspaper advertisements. Although all children were within the normal distribution, they did not demonstrate high articulatory proficiency. Because Oliver (1989) did not report any standardized measures of articulatory performance of her subjects, comparisons on this variable cannot be made. It is possible that the NA subjects reported here had less advanced articulatory skills than the subjects studied by Oliver. Alternatively, the task may be too difficult for 3-year-old children.

Insert Table 1 about here

Experiment 2

The results from Experiment 1 established that children between 4 and 6 years could perform the task well in quiet, but as with adults responding in closed-set tasks, there was no differential performance on a multiple-talker task as compared to a single-talker condition. Experiment 2, therefore, addressed the challenge of developing a task that would bear out the differences between single-talker and multiple-talker tasks and still be within the capabilities of children. Results from the attempted replication of Oliver's studies and from Experiment 1 indicated that presentation with a 0 dB SNR was too difficult and that presentation in quiet was too simple. Previous research indicated that public school classrooms exposed kidergarten children with SNRs of +1 dB to +5 dB (Sanders, 1965). In Experiment 2, a slightly more advantageous SNR was used. Subjects in this experiment heard words presented in noise, with an SNR of +6 dB.

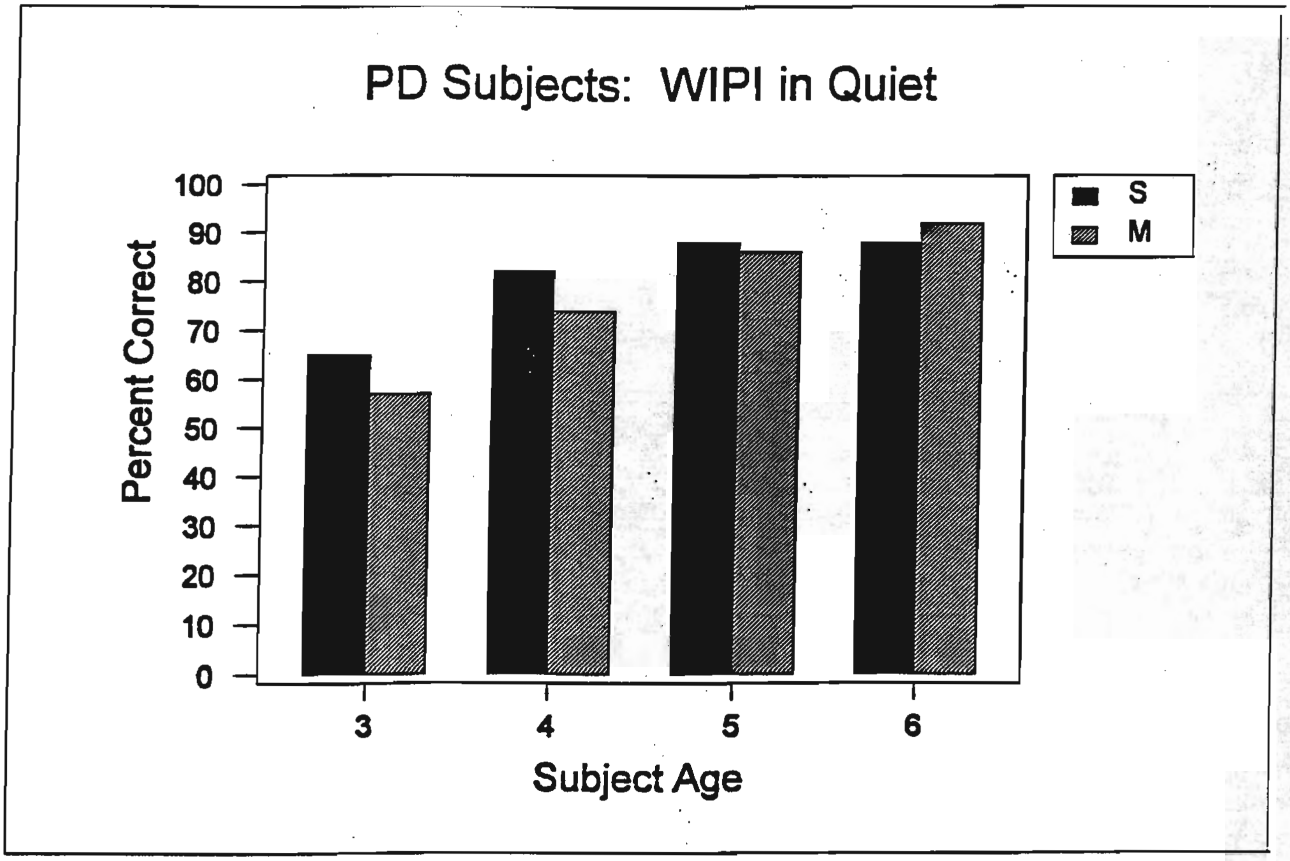

Figure 1: Percent correct word identification on the WIPI when words were presented in quiet. Subjects were phonologically delayed (PD) children between the ages of 3 and 6 years. Each subject was tested on both a single-talker (S) and multiple-talker (M) list.

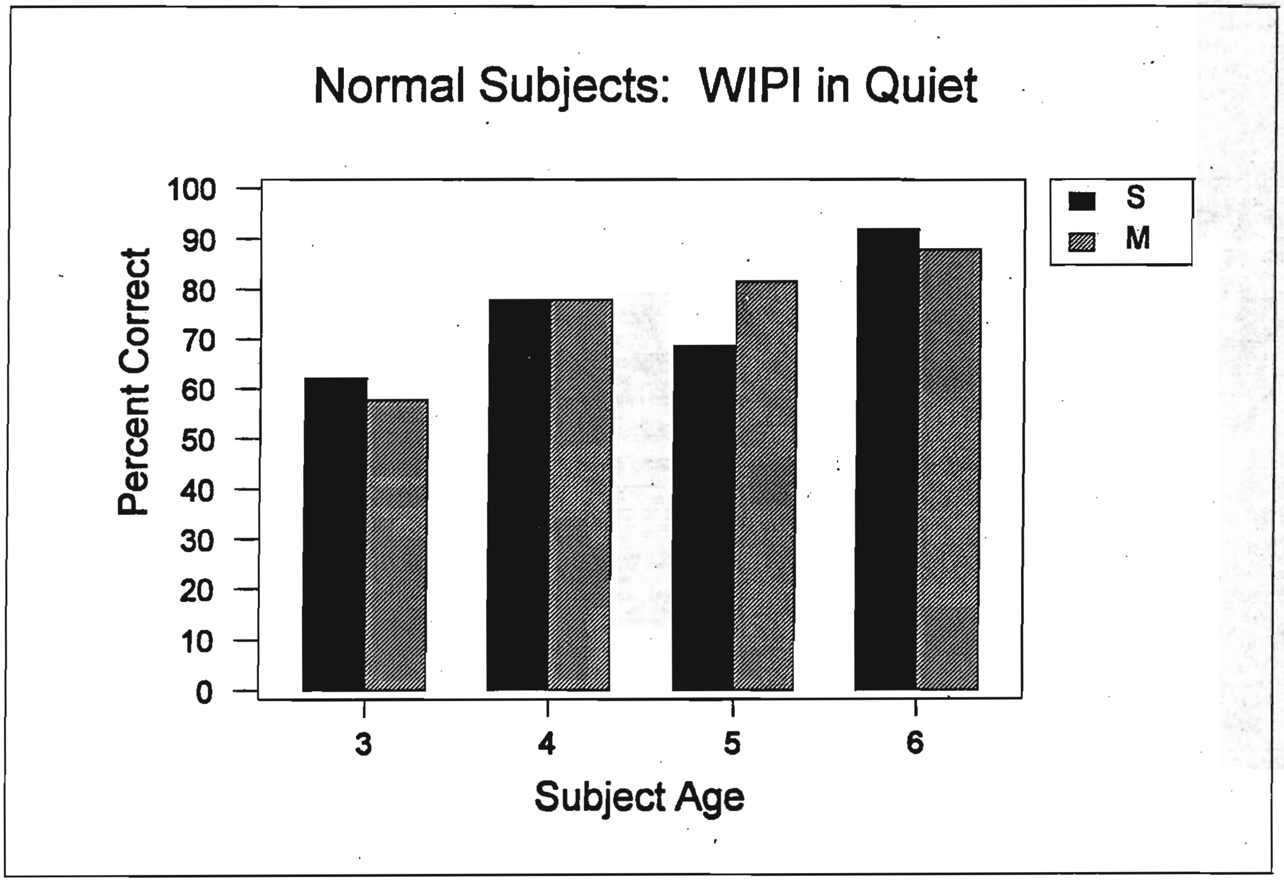

Figure 2: Percent correct word identification on the WIPI test when words were presented in quiet to normally articulating children. Subjects heard both a single-talker (S) and a multiple-talker (M) list.

Table 1

Standardized test results for groups of normally articulating children in Experiment 1. The Goldman Fristoe Test of Articulation (GFTA) was used to determined articulatory proficiency relative to age-normalized standard scores and the Peabody Picture Vocabulary Test (PPVT) screened the receptive vocabularies of the children in this investigation.

Subject Group	**Mean Age (SD) (years;months)**	**Average GFTA Percentile (SD)**	**PPVT Standard Score (SD)**
3 year old	3;4 (.25)	30.9 (16.0)	104.3 (13.2)
4 year old	4;0 (.50)	64.5 (18.4)	106.1 (15.8)
5 year old	5;0 (.29)	59.5 (32.0)	96.0 (11.3)
6 year old	6.2 (.41)	58.0 (23.6)	97.0 (17.6)

Method

Subjects.

Subjects for Experiment 2 were 16 normally articulating children between the ages of 4;0 and 5;5. These subjects were recruited in the same manner as in Experiment 1.

Materials.

Materials for both screening and experimental testing for Experiment 2 were identical to those used in Experiment 1, i.e., screening tests for hearing acuity, articulatory development, and receptive vocabulary development; and experimental testing using the WIPI.

Experimental Procedure.

Procedures for Experiment 2 were identical to those for Experiment 1, with the following exception. Whereas in Experiment 1, WIPI stimuli were presented over headphones in quiet, in Experiment 2, stimuli were presented over headphones embedded in background noise. As the signal passed through the audiometer, speech noise was added to achieve an SNR of +6 dB.

Results

A two-way ANOVA revealed significant main effects of talker condition (F=4.45; p=0.04), but no effect of NA subject age (F=0.13; p=0.72) or any interaction between talker condition and listener age (F=0.94; p=0.35). As seen in Figure 3, NA children between the ages of 4 and 5 years could only identify about 50% of the WIPI words correctly, on average, when the SNR was +6 dB. Although previous research suggested that school-aged children could understand speech presented at 6 dB SNR, the present data indicate that preschool children have difficulty with this task. Identification of words at the 50% level, though above chance performance, still is rather poor for normally articulating children. If a difference does exist between NA and PD children in their ability to normalize speech, this difference might be obscured by the low scores of the NA children. That is, the poor performance of the NA children did not yield a sufficient margin for the PD children if they had more difficulty with talker normalization. Recall that all children in these investigations were required to have normal hearing. The word identification scores of the NA children in this study may be too close to a lower limit of speech discrimination for normal hearing children. In an effort to improve the performance of the NA children, the experiment was repeated with some additional modifications to the protocol.

Insert Figure 3 about here

Experiment 3

Two additional changes in protocol were implemented in Experiment 3. First, it was not clear that all of the stimulus words were present in the children's lexicons. This might occur if the stimulus word had a synonym and it was only the synonym that was present in the child's lexicon. For example, one of the stimulus words was "pail," but in the dialect area from which these subjects were drawn, "bucket" is the more common lexical item. It seemed plausible that if children did not know the target word, then they could not choose the correct representation. Further, some of the pictures were artistically obscure representations of the target words. To overcome these problems, the test protocol was changed so that all six pictures in each trial were identified for the child before each stimulus word was presented. In addition

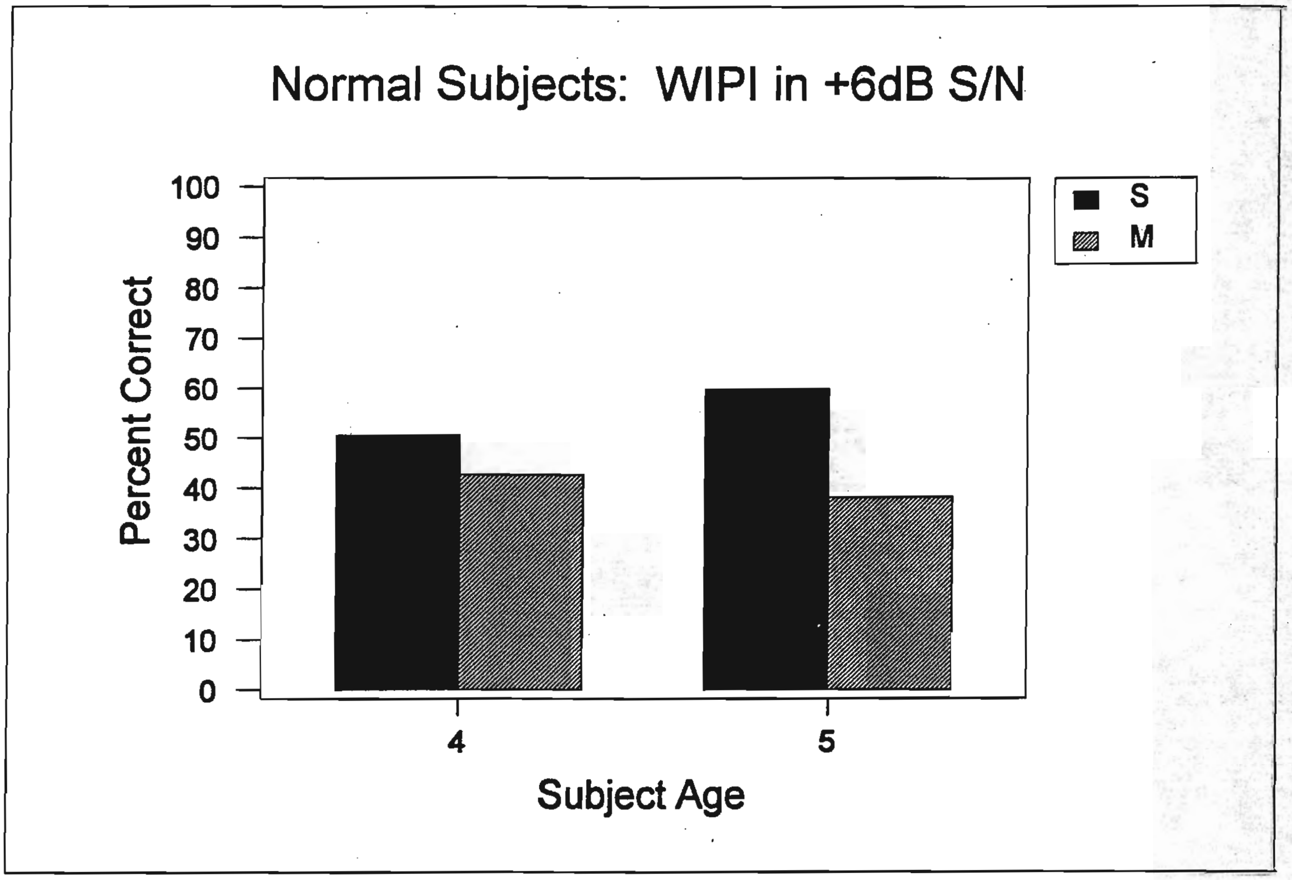

Figure 3: Percent correct word identification by normally articulating children. Words from the WIPI test were presented via headphones at +6 dB signal-to-noise ratio. Both single-talker (S) and multiple-talker (M) lists were used for all children.

to this change in protocol, the SNR was raised from +6 dB, as in Experiment 2, to +8 dB. Finally, since all test words, as well as foils, were presented to the children in conjunction with the corresponding pictorial representations, the PPVT was not administered in Experiment 3.

Methods

Subjects.

Subjects for Experiment 3 were 20 normally articulating children, ages 4 to 5, and 15 phonologically disordered children, ages 4 to 6.

Materials.

Materials for Experiment 3 were identical to those for Experiments 1 and 2 with one exception. Because of the change in the procedure for administration of the WIPI, as described below, the PPVT was no longer administered. Thus, materials for Experiment 3 consisted of the test of hearing acuity, the GFTA, and the WIPI, albeit with a change in administration procedure.

Experimental Procedure.

Procedures for administration of the hearing screening and the GFTA remained in Experiment 3 as in Experiments 1 and 2. Two changes in the WIPI protocol were introduced in Experiment 3. First, the SNR was raised to +8 dB. Second, in order to obviate possible gaps in the children's lexicons that might affect their responses on the WIPI, each trial was preceded by a review of the vocabulary items corresponding to the six response pictures. First, the six-picture plate was shown to the child, and the examiner would point to each picture in turn while saying aloud the word corresponding to that picture. When all six pictures had been identified, the examiner then pointed to each picture in turn and asked the child to name each picture. If the child could not remember the word corresponding to a picture, the examiner repeated the word until the child could correctly associate the words with the pictures. When this familiarization stage was completed, the stimulus word was presented at a +8 dB SNR. This procedure was followed on all twenty-five trials on both the single-talker and multiple-talker lists.

Results

Comparison of Figures 3 and 4 reveals that NA children's performance on the WIPI improved when the SNR was increased to +8 dB and subjects were familiarized with the stimulus words. In fact, the NA children in this study performed comparably to the children in Oliver's investigation (1990) in which children heard the words of the WIPI test at 0 dB SNR and no introduction to the stimulus words. A two-way ANOVA indicated a significant main effect for talker condition (F=7.02; p=0.014) but no effect of children's age (F=1.34; p=0.26) nor any interaction between these variables (F=0.53; p=0.48)

Insert Figure 4 about here

The success of this protocol with NA children led to the examination of word identification in noise (+8 dB SNR) by PD children. As indicated in Figure 5, PD children had levels of performance that were similar to those of the NA children. In the single-talker condition the PD children averaged 73% correct word identification, compared to 63% correct identification on the multiple-talker list, a difference that was significant (F=13.97; p=0.001). Figure 6 shows that NA children had slightly higher levels of performance on the single-talker (78% correct) and multiple-talker conditions (66% correct word identification) than

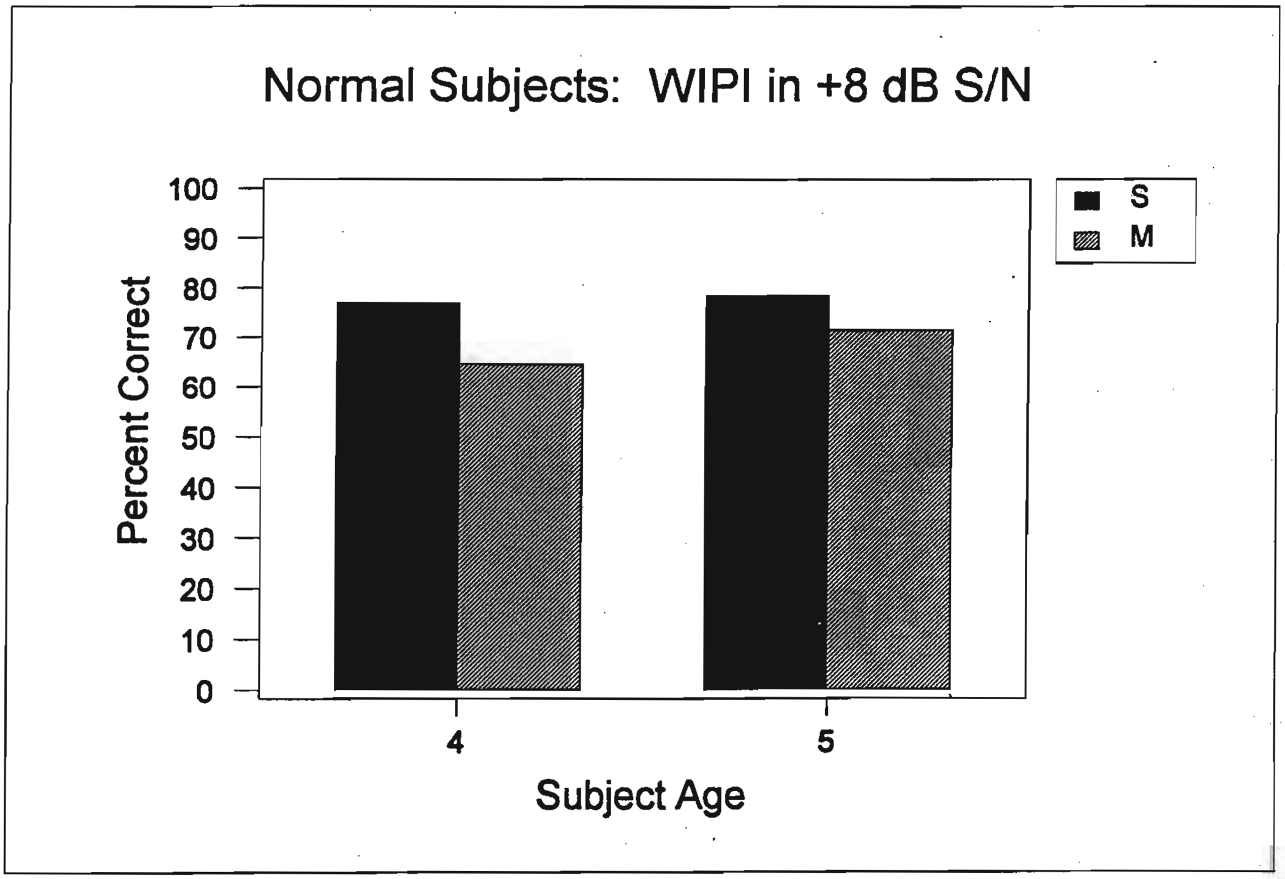

Figure 4: Percent correct word identification for normally articulating children on the WIPI when words were presented at +8 dB signal-to-noise ratio. All children heard both a single-talker (S) and a multiple-talker (M) list.

were found in the PD subjects; however, the differences between subject groups were not significant (F=0.87).

Insert Figures 5 and 6 about here

DISCUSSION

Within this series of experiments, attempts were made to develop a protocol to test talker normalization abilities of young child with normal and delayed speech. Part of this goal was met in that a satisfactory procedure was obtained to examine talker normalization in normally-articulating children. The data from these studies indicate that, like adults, speech perception in young children is negatively impacted by talker variation. Compared to their performance in the single-talker condition, the NA children scored approximately 12% worse when successive words in a list were spoken by different talkers. Although this difference in word identification accuracy as a function of talker condition is greater for the NA children than for adults (Mullennix, et al., 1989), differences in the protocol used to test adults and children make direct comparisons of effects somewhat difficult. The magnitude of difference in word identification for single- versus multiple-talker lists is similar, if not greater, for young NA children compared to adults (Mullennix, et al., 1989).

A number of shortcomings to the test procedure for young children were identified in the current study. Experiment 1, in which WIPI test words were presented in quiet, revealed that three-year-old children do not perform well on this task. Compared to previous studies in which 3-year-old children scored between 78 and 92% correct on the WIPI (Oliver, 1989; Sanderson-Leepa & Rintelmann, 1976), 3-year-old subjects in the present study averaged 63% correct identification of the words spoken by a single talker. Although it is difficult to account for the differences between the present data and those reported previously, some factors that may have influenced the results should be noted. Compared to its use in audiometric environments in which words are presented through a loudspeaker (Sanderson-Leepa & Rintelmann, 1976) or monitored live voice (Ross & Lerman, 1970), subjects in the present study listened to stimuli presented via headphones. It is possible that the discomfort and/or novelty of headphones was distracting to the youngest subjects. Headphone presentation of stimuli was chosen to be consistent with the procedures used by Oliver (1989, 1990). However, there was one change implemented in the present investigation; whereas Oliver allowed parents to accompany their children into the test booth, parents of the children in this study were permitted to observe the experiment only from an adjoining booth. This procedure was adopted to ensure that parents did not provide any cues to their children about the correct response. Although parents could be seen through a plexiglass window, it is possible that the youngest children in this study felt ill-at-ease in the test booth.

Other factors may have influenced the relatively poor results of the 3-year-old subjects in this study. Through the course of Experiments 1 and 2, as well as through communication with practicing audiologists (Iler-Kirk, personal communication), it became clear that many of the stimulus words were unfamiliar to our subjects. For example, one response set included the words "spring" with a bed spring as the picture and "string," depicted by a ball of yarn. These items are not common to young children. Further, some pictures could be identified with different words; the target "pail" was called "bucket" by the children in this investigation. Novelty of a word in the children's lexicons may have influenced response choices.

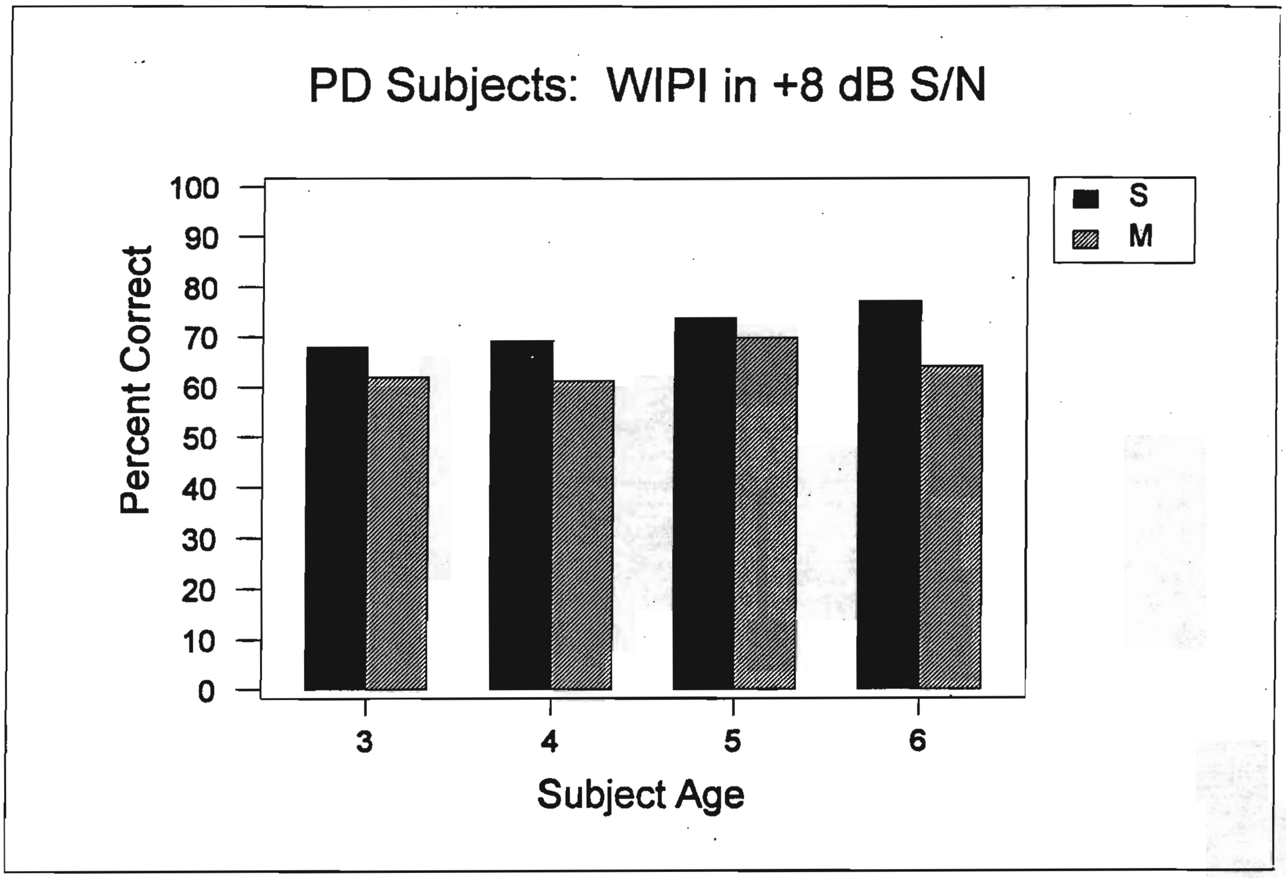

Figure 5: Percent correct word identification on the WIPI for phonologically delayed (PD) children. Words were presented at +8 dB signal-to-noise ratio and all children heard both a single-talker (S) and a multiple-talker (M) list.

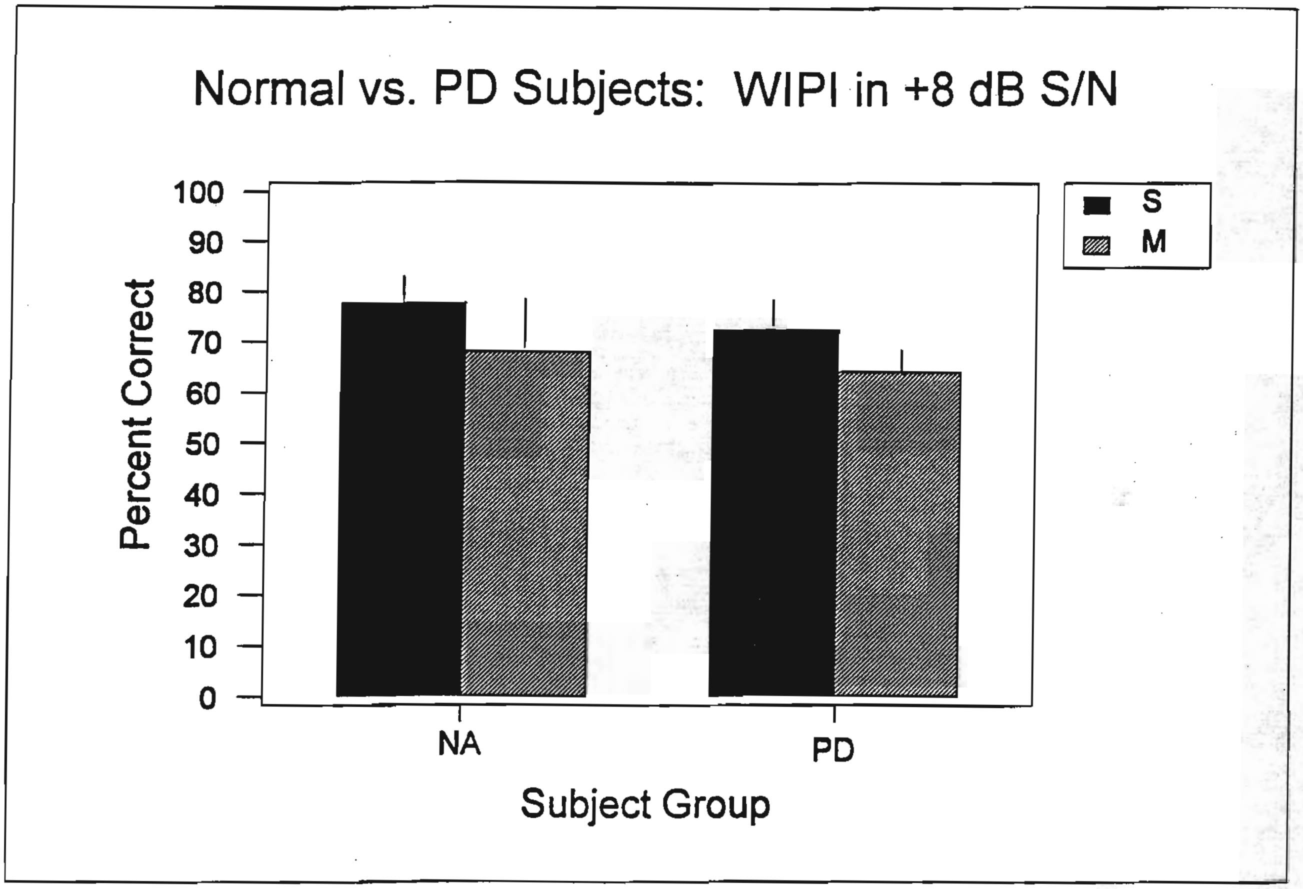

Figure 6: Comparison of percent correct word identification on the WIPI for phonologically delayed (PD) and normally articulating (NA) children. Data are a composite of figures 4 and 5. Words were presented at +8 dB signal-to-noise ratio and all children heard lists spoken by a single talker (S) and multiple talkers (M).

Two final factors should be considered when interpreting the present results. First, the children in the NA 3-year-old group had articulatory skills that were at the low end of the normal distribution (Goldman & Fristoe, 1986). Although there is no percentile that demarcates delay, the youngest subjects in Experiment 1 had articulatory proficiencies that were below the average of their peers. The low performance of these subjects on the WIPI may be reflective of a generalized delay in speech sound development. If these 3-year-old children had articulatory delays, the WIPI may have underestimated the speech perception abilities of normally articulating 3-year-old subjects. As a second factor, the six alternative responses for each stimulus word on the WIPI are distinguished by variations in the consonants, only. For example, one response set includes the words "pail," "mail," "sail," "nail," "jail," and "tail," i.e., words that rhyme with one another. Recent research shows that phonologically delayed children have difficulty with rhyme schemes (Wells, Stackhouse & Vance, 1994). Children who are acquiring speech normally may understand the relation between the words in a WIPI response set, but phonologically delayed children may not have this recognition. The complexity of the response matrix in the WIPI test, including the rhyme schemes of the words, may present added distractions to children with reduced articulatory proficiency.

Talker normalization research with phonologically delayed children presents some unique challenges, some related to the subject population and others related to the protocol needed to show effects of talker variation on speech perception. On the one hand, studies that demonstrate that talker normalization is incomplete, thereby reducing the accuracy of word identification by adults, only show this effect for open-set response formats (Humes, personal communication). With a limited set of response alternatives, word identification by adult listeners is not impacted by stimulus variation. If there are no limitation on the response possibilities, as is characteristic of an open-set format, adult listeners' word identification decreases with stimulus variation.

Unfortunately, an open-set response protocol cannot provide valid data from phonologically delayed children. Because children have a variety of substitution patterns, and because tests are targeting children's perception of misarticulated sounds, it may be difficult to identify the word produced by the PD child without some visual cue. For example, a common error of PD children is the production of [t] for /k/ thereby making homonyms of words like "tot" and "cot." In a test of word identification in which stimulus items are differentiated by a single phone, it would be unreasonable to have the child respond orally since many words produced by the PD child would sound the same to the experimenter. A closed response set is mandatory to test talker normalization in the misarticulating child, but the format of the test must by adapted to avoid "ceiling" effects known to accompany closed-set formats.

The most obvious adaptation is the introduction of noise into the signal. However, differences between preschool children and adults must be considered in choosing the intensity of the noise relative to the signal. Results of Experiment 2 demonstrate that preschool children are more affected than their school age counterparts in identifying words in the presence of noise. Whereas school age children can understand speech at a SNR of +6 dB (Ross & Giolas, 1974), preschool children do not score well on tests of word identification when this SNR is used. In fact, previous results demonstrate that preschool children make more errors of speech identification than their school-aged counterparts, even at more advantageous signal-to-noise ratios (Schwartz & Goldman, 1974). In the present series of experiments, it was found that preschool children between the ages of four and five years can identify words presented in noise if the SNR is +8 dB.

A final caveat offered by the present results is that children must be aware of the intended word for the stimulus, as well as the picture that corresponds to that word. This may seem obvious; however, standardized tests of speech discrimination and articulation were developed decades ago. These tests may target words that are no longer in common usage, particularly by children. Further, linguistic diversity may influence the words in children's lexicons, even for the same target item. Although an adult may be able to offer various synonyms for a picture, it is unlikely that a young child will display this competence. Using the example cited earlier, one of the stimulus words is "pail." As noted by Edward Eggleston (1871), "They have no pails in Indiana." In the local lexicon, children use the word "bucket" to identify a large vessel designed to carry water. Telling the subjects that the word "pail" may not result in the selection of the correct picture when that stimulus is presented. Care must be taken in the construction of speech perception tests to ensure that words are familiar to the target subject population. Further, the pictures must present clear representations of the words. Ambiguity about the identify of a picture may cause the child to avoid it as a response.

The initial goal of the present study was to investigate talker normalization in phonologically delayed children. Although PD children generally had poorer word identification than their NA peers, an accurate assessment of talker normalization in PD children may not have been revealed in this study because of difficulties in the development of an appropriate procedure. In addition to the suggestions offered above, it seems critical that speech perception tests designed for PD children target sounds that are commonly misarticulated. The test that was used in the present study (i.e. the WIPI) was designed to test word identification in hearing impaired children. The considerations of the developers of that test (Ross & Lerman, 1970) were consistent with the needs of that target group. It is imperative that a more appropriate instrument be developed to assess speech perception and talker normalization in phonologically delayed children.

References

American National Standards Institute. (1970). *Specifications for audiometers* (ANSI S3.6-1969). New York: ANSI.

Bernthal, J.E. and Bankson, N.W. (1993). *Articulation and phonological disorders.* Englewood Cliffs, NJ: Prentice Hall.

Blumstein, S.E. and Stevens, K.N. (1979). Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants. *Journal of the Acoustical Society of America*, **66**, 1001-1017.

Cantwell, D. P. & Baker, L. (1985). Psychiatric and learning disorders in children with speech and language disorders: A descriptive analysis. *Advances in Learning and Behavioral Disabilities*, **4**, 29-47.

Cantwell, D. P., Baker, L., & Mattison, R. (1981). Prevalence, type, and correlates of psychiatric diagnoses in 200 children with communication disorder. *Developmental and Behavioral Pediatrics,* **2**, 131-136.

Creelman, C.D. (1957). Case of the unknown talker. *Journal of the Acoustical Society of America*, **29**, 655.

Dinnsen, D.A. (1984). Methods and empirical issues in analyzing functional misarticulations. In M. Elbert, D.A. Dinnsen, and G. Weismer (Eds.), *Phonological theory and the misarticulating child* (ASHA Monograph, 22) (pp. 59-68). Rockville, MD: American Speech-Language-Hearing Association.

Dunn, L.M. and Dunn, L.M. (1981). *Peabody picture vocabulary test-revised.* Circle Pines, MN: American Guidance Service.

Eggleston, E. (1871). *The Hoosier schoolmaster: A story of backwoods life in Indiana.* New York, NY: Orange Judd and Company.

Eilers, R.E. and Oller, D.K. (1976). The role of speech discrimination in developmental sound substitutions. *Journal of Child Language*, **3**, 319-329.

Gierut, J.A., Elbert, M., and Dinnsen, D.A. (1987). A functional analysis of phonological knowledge and generalization learning in misarticulating children. *Journal of Speech and Hearing Research*, **30**, 462-479.

Goldinger, S.D., Pisoni, D.B. and Logan, J.S. (1991). On the nature of talker variability effects on recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **17**, 152-162.

Goldman, R. and Fristoe, M. (1986). *Goldman-Fristoe test of articulation.* Circle Pines, MN: American Guidance Service.

Guyette, R. and Diedrich, W., (1981) A critical review of developmental apraxia of speech. In N. Lass (Ed.), *Speech and language: Advances in basic research and practice, Volume 5,* New York, NY: Academic Press.

Hall, P., Jordan, L., and Robin, D. (1993). *Developmental apraxia of speech: Theory and clinical practice*. Austin, TX: Pro-Ed.

Ingram, D. (1990). *Phonological disability in children.* London: Cole and Whurr Limited.

Joos, M.A. (1948). Acoustic phonetics. *Language*, **24**, Suppl. 2, 1-136.

Kewley-Port, D., Pisoni, D.B. and Studdert-Kennedy, M. (1983). Perception of static and dynamic acoustic cues to place of articulation in initial stop consonants. *Journal of the Acoustical Society of America*, **73**, 1779-1793.

Lapko, L. and Bankson, N. (1975). Relationship between auditory discrimination, articulation stimulability and consistency of misarticulation. *Perceptual and Motor Skills*, **40**, 171-177.

Locke, J.L. (1980). The inference of speech perception in the phonologically disordered child, Part I: A rationale, some criteria, the conventional tests. *Journal of Speech and Hearing Disorders*, **45**, 431-444.

Martin, C.S., Mullennix, J.W., Pisoni, D.B., and Summers, W.V. (1989). Effects of talker variability on recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **15**, 676-684.

Mullennix, J.W., Pisoni, D.B. and Martin, C.S. (1989). Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America*, **85**, 365-378.

Oliver, B.R. (1989). Talker variability and word recognition: A developmental study. *Research on Speech Perception Progress Report No. 15* (pp. 471-485). Bloomington, IN: Speech Research Laboratory.

Oliver, B.R. (1990). Talker normalization and word recognition in preschool children. *Research on Speech Perception Progress Report No. 16* (pp. 379-389). Bloomington, IN: Speech Research Laboratory.

Ross, M. and Giolas, T. (1974). Effect of three classroom listening conditions on speech intelligibility. *American Annals of the Deaf,* **116**, 580-584.

Ross, M. and Lerman, J. (1970). A picture identification test for hearing impaired children. *Journal of Speech and Hearing Research*, **13**, 44-53.

Ross, M. & Lerman, L. (1971). *Word intelligibility by picture identification*. Pittsburgh: Stanwix House, Inc.

Rvachew, S. and Jamieson, D.G. (1989). Perception of voiceless fricatives by children with a functional articulation disorder. *Journal of Speech and Hearing Disorders*, **54**, 193-208.

Sanders, D. (1965). Noise conditions in normal school classrooms. *Exceptional Children*, **31**, 344-353.

Sanderson-Leepa, M.E. and Rintelmann, W.F. (1976). Articulation functions and test-retest performance of normal-hearing children on three speech discrimination tests: WIPI, PBK-50, and NU Auditory Test No. 6. *Journal of Speech and Hearing Disorders*, **41**, 503-519.

Schwartz, A. and Goldman, R. (1974). Variables influencing performance on speech sound discrimination tests. *Journal of Speech and Hearing Research*, **17**, 25-32.

Shelton, R., Johnson, A., and Arndt, W. (1977). Delayed judgment speech sound discrimination and /r/ or /s/ articulation status and improvement. *Journal of Speech and Hearing Research*, **20**, 704-717.

Sherman, D. and Geith, A. (1967). Speech sound discrimination and articulation skill. *Journal of Speech and Hearing Disorders*, **10**, 277-280.

Stevens, K.N. and Blumstein, S.E. (1978). Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America*, **64**, 1358-1368.

Studdert-Kennedy, M. (1974). The perception of speech. In T.A. Sebeok (Ed.), *Current trends in linguistics, Volume 12* (pp. 2349-2385). The Hague: Mouton.

Studdert-Kennedy, M. (1976). Speech perception. In N. Lass (Ed.), *Contemporary issues in experimental phonetics* (pp. 243-293). New York: Academic Press.

Thelen, E. and Smith, L.B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.

Thoonen, G., Maassen, B., Gabreels, F. and Schreuder, R. (1994). Feature analysis of singleton consonant error in developmental verbal dyspraxia (DVD). *Journal of Speech and Hearing Research*, **37**, 1424-1440.

Travis, L. and Rasmus, B. (1931). The speech sound discrimination of cases with functional disorders of articulation. *Quarterly Journal of Speech*, **17**, 217-226.

Wells, B., Stackhouse, J., and Vance, M. (1994). Rhyme and representation in phonological development and disorders. Paper presented at the Fourth Symposium of the International Clinical Phonetics and Linguistics Association, New Orleans.

Winitz, H. (1984). Auditory considerations in articulation training. In H. Winitz (Ed.), *Testing and treating articulation disorders: For clinicians by clinicians*. Baltimore, MD: University Park Press.

Wolfe, V. and Irwin, R. (1973). Sound discrimination ability of children with misarticulation of the /r/ sound. *Perceptual and Motor Skills*, **37**, 415-420.