SPEAKER-LISTENER
FAMILIARITY: PARENTS AS
JUDGES OF DELAYED SPEECH
INTELLIGIBILITY

PETER FLIPSEN JR.
School District #81 Fort Nelson, B.C. Canada

It has been frequently proposed that one source of bias in intelligibility judgments is whether the judge knows the subject personally. This study attempted to describe this bias using 4 speech-delayed children (aged 4 and 7 years) with normal hearing and no significant language delay. They were audiotaped at four points during therapy over a 12-18 month period using imitation of single words derived from the Assessment of the Intelligibility of Dysarthric Speech. Parents and a group of unfamiliar adult couples served as judges. Results indicated that mothers were significantly better than all other listeners (including fathers) at identifying the words being spoken. An interaction between listener type and subject age class was also observed with the maternal advantage being much less apparent with the 2 older subjects. Results are discussed relative to the possible source of the mothers’ superior performance.

INTRODUCTION

In working with children who have significant speech delays, the ultimate goal is to improve the child’s ability to get their message across to the listener (i.e. improve intelligibility). Formal measures of intelligibility are not routinely made by most clinicians working with children. There are at least three reasons for this. First, a variety of metrics have been proposed (e.g. interval scaling, direct magnitude estimations, forced-choice measures and transcription tasks). There may be some confusion as to the most appropriate technique to use. A second reason may relate to reading skills.
Preparing samples of sentence or text materials for intelligibility judgments is usually done by having the subject read the material onto tape. Young children often do not have reading abilities that permit this. Judgments then tend to be limited to single word tasks. A third possible reason for the avoidance of formal measures of intelligibility with children may be the large number of factors often cited as influencing intelligibility judgments. Among these are "...the listener’s familiarity with the speaker, contextual cues provided to the listener, the nature of the speech sample and familiarity with the speech sample" (Beukelman & Yorkston, 1980; p. 34). The extent of the specific influence of each of these factors has been studied in some depth but is not fully understood.

Clinicians wanting to formally measure the intelligibility of children are thus often faced with conducting a single-word measurement without a clear sense of either the validity of their chosen technique or how it can be reliably carried out.

This study was an attempt to add to the knowledge base in this area by assessing whether high familiarity with a speaker (i.e. knowing them personally) creates a bias in intelligibility judgments. Specifically, the question was examined relative to the intelligibility of children with delayed speech.

Such a bias (or intelligibility advantage) would have considerable clinical significance. For example, if familiarity results in elevated intelligibility scores, clinical judgments may be biased as therapy progresses. During the course of therapy, particularly with highly unintelligible clients who spend long periods in treatment, clinicians may effectively become "familiar" listeners. This is what has been termed the “accommodation” effect (Weiss, Gordon & Lillywhite, 1987; p. 125). If this occurs, the use of repeated measures by a clinician on the same subject might not be a reliable metric of change (Brodkey, 1972; Doyle, Swift & Haaf, 1989). Yorkston and Beukelman (1983) noted that "it is reasonable to expect that such [clinician] familiarity would elevate intelligibility scores". (p. 161).

A second clinical impact of a familiarity bias could be that if parents are making the pre-referral judgments, the result may be a delay in the referral. This may reduce the chances for success through early intervention.

Should this bias exist, at least three possible reasons may account for it. It has been suggested that "...it would seem entirely possible that a familiar partner may perceive and respond to a child’s behavioral cues that may be 'missed' by an experimental observer" (Wilcox, Kouri & Caswell, 1990, p. 681). Second the familiar listener may simply have a greater contextual knowledge base. He or she may, for example, know the usual topics of conversation, the names of friends and relatives and the individual's food and activity preferences. And third, there has been speculation that the highly familiar listener may be learning to "translate" the atypical speech patterns themselves (Boothroyd, 1985; Goehl & Martin, 1987; Weist & Krupey, 1977).

The Question of Familiarity

Of the factors influencing judgments of intelligibility, familiarity (also called judge experience) has been examined in a number of ways. Researchers have examined familiarity with disordered speech in general (e.g. Beukelman & Yorkston, 1980; Doyle et al., 1989; Ellis & Fucci, 1991; Goehl & Martin, 1987). They have studied exposure to the speech patterns of particular disordered populations (e.g. Boothroyd, 1985; Brodkey, 1972; Doyle et al., 1989; McGarr, 1981; Mencke, Ochsner, & Testut, 1983; Monsen, 1978; Monsen, 1983) and have assessed familiarity with the content of the material being spoken (e.g. Beukelman & Yorkston, 1980; Cullinan, Brown & Blalock, 1986; Garret & Saint-Pierre, 1980; Mullenix, Pisoni & Martin, 1989; Yorkston & Beukelman, 1980a).

Relative to the current study, a number of authors have suggested that a judge who is personally familiar with the speaker may have an advantage in understanding what is being said (e.g. Beukelman & Yorkston, 1980; Boothroyd, 1985; Elbert & Geirut, 1986; Ellis & Fucci, 1991; McGarr, 1983; Monsen, 1981; Yorkston & Beukelman, 1980a).

A few studies have examined this question empirically. Yorkston and Beukelman, (1983) for example, attempted to control the degree of familiarity by studying a group of dysarthric adults who were initially unknown to 2 groups of judges. One group of judges was allowed to become more familiar with the speakers by listening to recorded samples of sentence material three times. They were then asked to make judgments on new material from the same speakers. Results indicated no significant difference in intelligibility scores when compared to judges who were not permitted this “familiarization” phase.

Brodkey (1972) looked at the question more directly by studying teachers and their ESL students and found a distinct familiarity advantage. Using a discourse transcription task, teachers judging their own students generated intelligibility scores of 1.5–2 standard deviations above the mean obtained by other (unfamiliar) teachers on the same students. Those same teachers performed "... well within 1 SD of the mean for ESL teachers" (p. 215) when listening to students that they were not personally teaching.

This question has also been examined with reference to very young normally developing children (Weist & Krupey, 1977). Using videotape extracts of the spontaneous sentences of children aged 25–42 months, results indicated that mothers were better at determining what was said than any other listeners including fathers. Adults were better than children with siblings being superior to non-siblings.

Goehl and Martin (1987) studied 3 speech-delayed children producing sentences containing words chosen to elicit each child’s specific segmental error patterns. The children’s mothers, 4 experienced speech-language pathol-
ologists (SLPs), and 4 inexperienced SLP students transcribed the tapes. Results indicated that the mothers performed consistently better than both of the other groups. No statistical test results were reported however.

In sum, the research to date appears to substantiate at least some comprehension (or intelligibility) advantage for those who are personally familiar with the speaker. Intelligibility scores may be elevated as a result of this familiarity.

The Current Question

This study was an additional examination of the question of whether being personally familiar with the speaker yields an intelligibility advantage (i.e. elevates intelligibility scores) when judging children with delayed speech. Parents were chosen as the focus of the study because they potentially possess the greatest degree of familiarity with their young children. Parents were, in fact, initially chosen as judges because being the most familiar listeners, their judgments may be the most socially valid (Campbell & Dollaghan, 1992). Finally, parents of children with delayed speech frequently report that they have greater success at understanding their children than do unfamiliar listeners.

METHOD

Subjects

Subjects were drawn from the clinical population being served by the author. Clients are referred to the Speech and Language Clinic by a variety of sources including teachers, parents, physicians and public health nurses. Those clients with significantly reduced intelligibility routinely undergo formal intelligibility testing. For children, parents often act as the judges with testing typically repeated at regular intervals as one measure of progress in therapy.

The criteria for inclusion in the current study were reduced intelligibility (below 75% in spontaneous speech), significant speech delay, normal hearing, and no significant language delay. Subjects also had to be living with both natural parents. During the period of the study, four children attending the clinic met these criteria. Their parents were approached and all were willing to participate in the study.

Significant speech delay was defined as either scoring below the 10th percentile on the Goldman-Fristoe Test of Articulation (G-F; Goldman & Fristoe, 1986) or achieving a deviancy score of at least 40 (without age penalty) on the Assessment of Phonological Processes-Revised (APP-R; Hodson, 1986). Age penalty was removed from the APP-R scores to allow for comparison of the subjects based on the frequency of process use only.

As can be seen from Table 1, neither test was used with all the subjects. This was owing to the partially retrospective nature of the current study. The use of unfamiliar listeners to supplement parental judgments was not initiated until 3–6 months after all the subjects had begun therapy.

All subjects were judged to have normal hearing acuity. Subjects 2 and 4 passed a standard hearing screening (25 DB HTL at 500, 1000, 2000 & 4000 Hz bilaterally). Subject 1 failed the screening but was subsequently found to have normal hearing acuity as determined by a clinical audiologist. Subject 3 had been seen by a clinical audiologist prior to being referred for speech services and had also been found to have normal hearing acuity.

During the course of the study, all 4 subjects were treated for speech-related difficulties only (i.e., no language intervention goals beyond phonology). Subject 4 was deemed to have no significant language deficiencies during initial assessment as evidenced by analysis of spontaneous language samples. At the beginning of the study, expressive language skills of the other subjects could not be reliably assessed beyond the single-word level, owing to the highly reduced intelligibility of their spontaneous speech (see Table 1). Following 6–14 months in therapy, analysis of spontaneous language samples had revealed normal expressive language skills for subjects 1 and 2 but a mild expressive language delay for subject 3 (for which treatment was initiated after the end of the current study).

In addition, all subjects achieved standard scores of at least 80 (i.e., within 1.33 SD of their age-group mean) on standardized tests of both expressive and receptive language (see Table 2). These tests were administered by the author during the initial assessment of the children prior to the start of therapy.

Expressive vocabulary skills were assessed using the Expressive One-Word Picture Vocabulary Test-Revised (EOWPVT-R; Gardner, 1990). Re-
Table 2. Language Test Standard Scores

<table>
<thead>
<tr>
<th>Subject</th>
<th>ROWPVT&lt;sup&gt;a&lt;/sup&gt;</th>
<th>EOWPVT&lt;sup&gt;b&lt;/sup&gt;-R&lt;sup&gt;c&lt;/sup&gt;</th>
<th>TACL-R&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>106</td>
<td>125</td>
<td>97</td>
</tr>
<tr>
<td>2</td>
<td>108</td>
<td>103</td>
<td>103</td>
</tr>
<tr>
<td>3</td>
<td>96</td>
<td>97</td>
<td>90</td>
</tr>
<tr>
<td>4</td>
<td>86&lt;sup&gt;e&lt;/sup&gt;</td>
<td>86</td>
<td>82</td>
</tr>
</tbody>
</table>

<sup>a</sup>Receptive One-Word Picture Vocabulary Test.<br>
<sup>b</sup>Expressive One-Word Picture Vocabulary Test Revised.<br>
<sup>c</sup>Test for Auditory Comprehension of Language-Revised; Total test score.<br>
<sup>d</sup>Revised Peabody Picture Vocabulary Test-Revised; Form M.<br>

 receptive vocabulary was assessed for subjects 1–3 with the Receptive One-Word Picture Vocabulary Test (ROWPVT; Gardner, 1985) and for Subject 4 with the Peabody Picture Vocabulary Test-Revised (PPVT-R, Form M; Dunn & Dunn, 1981). In addition all subjects were tested using the Test for Auditory Comprehension of Language-Revised (TACL-R; Carrow-Woolfolk, 1985).

Recordings

Subjects were audiotaped four times each during the 12–18 month course of therapy. There was a 3–6 month interval between recordings for each child, again because these measurements were originally intended to monitor progress in therapy. For subjects 1, 3 and 4, the initial tape was made during the first month of therapy. For Subject 2, the first tape used for this study was made 4 months after the start of therapy, owing to some technical problems with the first tape. In total, there were sixteen (16) separate speaking events (four for each of the 4 subjects).

During each speaking event, the subject imitated the author speaking 50 single-words derived from the Assessment of the Intelligibility of Dysarthric Speech (Yorkston & Beukelman, 1984). Only the item number and the subject’s productions were recorded on the tape.

The test “...consisted of 50 sets of 12 similar sounding words...” (Yorkston & Beukelman, 1980b, p. 18). It was chosen for several reasons. It has a relatively large stimulus pool, stimuli are chosen randomly from this pool for each administration, and the authors cite a high level of test-retest reliability. These all make it useful as a clinical tool when conducting multiple administrations over time.

A specific advantage of this test for the current study was the possibility of using randomly chosen single words on audiocassette. This format would remove both speaking context and behavioral cues as possible sources of advantage in understanding what was said. Any advantage obtained would then likely be the result of the listener learning to “translate” the atypical speech patterns (Boothroyd, 1985; Goel & Martin, 1987; Weist & Kruppe, 1977).

The initiative task was initially chosen also for clinical reasons. It allowed for a relatively rapid administration of the procedure with even very young children. As well, for the purposes of the current study, it negated the problem of limited or nonexistent reading skills in all of the subjects. Finally, the test was originally designed for use with adults and, thus, many of the stimuli were not easily picturable.

For each speaking event, the author selected the 50 words by using a table of random numbers (1–12 provided with the test) to select one of the 12 words from each set. Each event thus contained its own unique collection of 50 words.

All events were recorded on normal bias audio cassette tape on a Panasonic Model RX-CS750 portable cassette stereo which included a five band graphic equalizer. Input was via an Audiotechnica Model AT9500 omnidirectional condenser microphone held 2–3 inches from the subject’s mouth by the author who conducted all of the recording sessions.

Judges

In addition to parents, a total of 20 unfamiliar couples judged the tapes. These couples were divided into two groups. Since the parents had judged their own child’s speech four times, a group of 4 couples (multi-event listeners) judged all events for 1 subject each in the same order as the subject’s parents. This was intended to control for the possibility that parents might become familiar with the judging procedure and the test word pool (Doyle et al., 1989). The second group of 16 couples (single-event listeners) judged only one event each. For each of the 16 speaking events there were then 6 judges: parents; unfamiliar multi-event couple; unfamiliar single-event couple. This yielded a total of 96 event judgments.

All judges (parents included) were married couples in good health between the age of 23 and 50 years of age with no history of significant hearing loss. All passed a pure tone hearing screening at 30 dB HTL at 500, 1000, 2000 and 4000 Hz bilaterally.

The 40 unfamiliar judges were certified classroom teachers and their spouses. None of the teacher-judges had any special training or significant experience in working with speech-delayed children. The 4 multi-event couples were themselves parents. Of the single-event couples, 10/16 were parents. Only 1/40 unfamiliar judges (a single-event male) did not have significant experience with children (being neither a teacher nor a parent).

In all cases, each listener judged the same event(s) as his/her spouse. In the majority of cases, the couples made their judgments simultaneously. No more than 3 judges were present at any one judging session. For multi-event
listeners there was a minimum of 8 weeks between judgment sessions in order to limit the possibility that they might be "accommodating" or learning the speech patterns of the subjects. Parental judgments were closely timed to actual tape preparation and thus their judgment sessions were all separated by at least 3 months.

Judging Conditions and Listening Task

Judges listened to each event played back on the same tape recorder used to make the recordings. Judgements were made in soundfield in quite rooms while sitting at a table 2–3 feet from the tape player speakers. Volume and frequency output levels were kept constant for each event using the tape recorder controls. All judgment sessions were conducted by the author.

The judges' task was to try to identify the 50 words spoken by the child. Judges listened to the words once only (as per the test manual) and selected the words they thought the subject had attempted to say using the multiple-choice form included with the test. The form displays the 12 possible choices for each word. The tape was paused after each word to allow the judges time to make their selection.

Judges were instructed not to preview the choices before hearing the words nor change selections once having progressed to subsequent items. In the event of uncertainty, they were instructed to guess. The number of correctly identified items was calculated and this value multiplied by 2 to derive a percent intelligibility score.

Intra-Judge Reliability

Following completion of data collection, intra-judge reliability was assessed by having 25% (24/96) of the judgments repeated using representatives of all the listener groups. Parents and multi-event couples repeated judgments on the last event judged. If the judges had been "translating" the atypical speech patterns, they would likely be doing so most effectively with the most recent version of those patterns (recall that the subjects were participating in therapy and thus modifying their phonological systems). Rejudgments made on older events might result in poorer reliability values that actually reflect changes in comprehension skill rather than poor reliability. Four of the 16 single-event couples were chosen quasi-randomly to repeat the judgment they had made (one randomly chosen from the four representing each subject). For all of the rejudgments, there were at least 8 weeks between the initial and the rejudgment sessions.

A Pearson Product Moment correlation coefficient was then calculated on the rejudgments and found to be 0.90 overall. Values ranged from 0.85–0.94 across listener groups. This is quite similar to values reported in other intelligibility studies (e.g. Cullinan, Brown & Blalock, 1986; Sheard, Adams & Davis, 1991). It also matched the test-retest reliability coefficient of 0.90 reported in the test manual for the multiple-choice format (Yorkston & Beukelman, 1984).

RESULTS

A Four-Factor Analysis of Variance (ANOVA) was conducted on the data (see Table 3). The four factors were: the six listener groups; the two age classes (4 and 7 years); the two sexes; the four time periods.

Significant differences were indicated amongst the listener groups (critical $p < 0.05$). To ascertain which groups differed from each other, a multiple comparison of the listener groups was carried out using Scheffe's method. The 6 listener groups were each compared two at a time (15 comparisons) as were the 3 couple types (three comparisons). A family confidence coefficient of 0.95 was obtained indicating a 95% probability that all 18 decisions about significant differences were correct.

Comparisons between the groups were based on the average difference in intelligibility score across all 16 events. Parents achieved intelligibility scores that averaged 4.4% higher than multi-event couples and 5.8% higher than single-event couples. Multi-event couples scored an average of 1.5% higher than single-event couples. Comparisons amongst the gender-based listener groups are shown in Table 4.

Results of the comparisons indicated that scores obtained by mothers were significantly different (i.e. higher) than all other listener groups including fathers. Parental scores as well were significantly higher than scores of either the unfamiliar couple type. However, since the differences between fathers and the unfamiliar listener groups were not significant, the better parental scores appeared to be due entirely to the superior performance of mothers. None of the other comparisons showed significant differences.

To determine whether the superior performance of mothers was linked to speaker age class, sex or the time period, the interaction between the factors

Table 3. Partial ANOVA Table

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Listener group</td>
<td>5</td>
<td>1108.71</td>
<td>221.74</td>
<td>6.63</td>
<td>0.000</td>
</tr>
<tr>
<td>Group x sex</td>
<td>5</td>
<td>254.21</td>
<td>50.84</td>
<td>1.52</td>
<td>0.199</td>
</tr>
<tr>
<td>Group x age class</td>
<td>5</td>
<td>447.21</td>
<td>89.44</td>
<td>2.67</td>
<td>0.032</td>
</tr>
<tr>
<td>Group x time</td>
<td>15</td>
<td>260.13</td>
<td>17.34</td>
<td>0.52</td>
<td>0.919</td>
</tr>
</tbody>
</table>
Table 4. Listener Group Comparisons: Mean Percentage Differences Across 16 Events

<table>
<thead>
<tr>
<th>Parents</th>
<th>Multi-event</th>
<th>Single-event</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mother</td>
<td>Father</td>
</tr>
<tr>
<td>Parents</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother</td>
<td>7.5&quot;</td>
<td>7.6&quot;</td>
</tr>
<tr>
<td>Father</td>
<td>0.1</td>
<td>1.3</td>
</tr>
<tr>
<td>Multi-event</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>1.1</td>
<td>1.1</td>
</tr>
<tr>
<td>Male</td>
<td>-0.2</td>
<td>1.9</td>
</tr>
<tr>
<td>Single-event</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>2.1</td>
<td></td>
</tr>
</tbody>
</table>

*Significant difference (p < .05).

was then examined (see again Table 3). The ANOVA indicated an interaction of listener group with age class but not with sex or time.

To examine the nature of the interaction of listener group with age class, the percent intelligibility scores were plotted for the 4 subjects and are shown in Figure 1.

As can be seen from Figure 1, scores obtained by mothers were equal or superior to those of other listener groups in every case for subjects 2 and 3 (the 4 year-olds). The results for the mothers of subjects 1 and 4 (the 7 year-olds) were much less dramatic. They had the highest scores for only 2/8 events. They did however have at least the second best score for 7/8 events.

DISCUSSION

The overall superior performance of mothers over other listener groups is similar to that found by Goehl and Martin (1987) as well as that of Weist and Kruppe (1977). It also seemed to confirm the previously noted anecdotal reports made by parents of speech-delayed children who cite superior skill at understanding their children when compared to unfamiliar listeners. Some of these reports specifically note a maternal advantage.

These conclusions are of course, less definitive when one considers that the superiority was clearly discernible for only 2/4 subjects. The interaction between listener group and age class suggests that the maternal advantage may be lost as the child gets older.

Of course, it could be argued that intelligibility level may be more critical than age. At the beginning of the study, the 4 year-old subjects also happened to be considerably less intelligible, at least at the single-word level (50% and 54% versus 76% and 80% for the 7 year-olds).

Mencke et al. (1983) however, suggest that an extremely reduced intelligibility level may actually negate any familiarity advantage. They were however referring to values below 50%. As well, they were looking at the speech of the deaf in a study of listener familiarity with that population in general and not personal familiarity with the speaker. The suggestion does however lend some weight to the notion that age is a more critical factor.

Another argument against intelligibility level as a critical factor is the obtained pattern (younger being less intelligible) is reminiscent of that expected from children developing speech normally (Weiss et al., 1987). Because (for the younger subjects) both of the multi-event couples and 5/8 single-event couples were parents of preschool children, they would have been fairly familiar with less intelligible speech. As well, the study by Weist and Kruppe (1977) showed a distinct maternal advantage in intelligibility judgments with younger normal children (albeit using a very different judging task; spontaneous sentences on video vs. imitated single-words on audio).
The maternal advantage for the younger subjects may also have been the result of a third factor, namely the amount of contact time between mother and child (actual time spent with the child around the time of the recordings/judgments). The two older subjects (no consistent gap between mothers and other listener groups) were both of school age for the duration of the study and thus were spending less time at home. Subject 3 (with the largest gap between the mother and the other groups) did not begin school until after the end of the study period. Subject 2 entered school after the second speaking event.

During the study period, the mothers of subjects 2 and 3 only worked part-time outside of the home and thus were in a position to consistently spend more time with the child than the father. The mothers of subjects 1 and 4 had no such advantage since their children were in school for the duration of the study period. No formal measures of contact time were attempted. It would prove interesting to repeat the current study while documenting such contact time and compare families where fathers and mothers were in a position to consistently spend more time with the child than their respective spouses.

In addition to the above, several other factors must be cited to qualify any conclusions being drawn from the current study. Given the small sample size, the listener group by age class interaction may have been, at least in part, the result of individual variability amongst the mothers (cf. Wilcox et al., 1990). The fact that the small group was spread across 2 different age classes only served to exacerbate this possibility.

The results may also have been confounded by the participation of all the subjects in therapy. Because the subjects were modifying their phonological systems, both parents and the multi-event listeners would have needed to adjust their listening for those modifications. If contact time was the crucial factor accounting for the maternal advantage, mothers of the older subjects may not have had sufficient time to adjust their listening to keep pace with the changes.

At the same time, parents of all 4 subjects were participating in therapy by conducting “carryover” activities at home. While formally documented, informal reports suggested that mothers were the ones doing the majority of this homework. At least some of the overall maternal advantage may have come from the very focussed contact time that these activities would have created.

Future investigators may wish to use larger sample sizes to explore the relationship between the amount of contact time (and/or the content of that time) and a parent’s ability to understand their child’s delayed speech. This might involve a similar paradigm to the current one but with samples limited to a particular point in therapy. Since parents are reporting their superiority at the outset of therapy, that might be a useful point to begin.

The extension of the current results to broader speaking contexts (i.e., phrases, sentences, narrative, conversation) must also only be made with considerable caution. Several authors have questioned the validity of extending single-word results to contextual speech (e.g., McGarr, 1981; Morrison & Shriberg, 1992; Tikofsky & Tikofsky, 1964; Yorkston, Beukelman & Bell, 1988). Schmidt-Nelson (1983) noted that “different aspects of the acoustic information are used in connected speech than in carefully pronounced isolated words” (p. 737).

This caution is especially important given the fact that the current study used imitation rather than spontaneous productions. A number of studies in the 1950s and 1960s demonstrated that imitated productions are produced more accurately by speech-delayed children (see Ingram, 1989, pp. 81-83). It is interesting to note that, despite this fact, fully 23% of studies on articulation and phonological disorders in the 1970s and 1980s continued the practice of using imitated words as a measurement procedure (Sommers, Logsdon & Wright, 1992).

Two recent studies in fact cast some doubt on any broad-based maternal advantage in interpreting their children’s communications. Kwiatkowski and Shriberg (1992) for example, studied a group of 15 caregivers. They asked them to prepare a “gloss” or transcription of connected speech from an interaction between their child and a clinician (which they had just observed and heard). A reference gloss was prepared by combining the caregiver gloss with one made by the child’s clinician and supplementing these with a gloss by one of the authors. While the caregivers glossed 78% of utterances and 81% of words, they were only correct on 58% of utterances and 73% of words. Thus the caregivers understood significantly less than they thought they did. That study did not include unfamiliar judges.

A second relevant study used 24 developmentally delayed children aged 16-36 months (Wilcox et al., 1990). It looked at mothers’ ability to identify early communication behaviors from 5 minute videotape samples. Results indicated that as a group, mothers were not significantly different from SLPs or teachers at identifying those behaviors in their own children. The authors did note a great deal of individual variability however stating that “...for a given child some observing partners recognize large amounts of communication and others recognize little” (p. 690).

In spite of the above noted cautions, a statistically significant maternal advantage was observed in this study. Because it was found in the highly decontextualized setting of single words on audiotape, the mothers may well be learning to “translate” the atypical speech patterns (Boothroyd, 1985; Goehl & Martin, 1987). Another follow-up to this study might involve determining whether this ability extends to their understanding the speech of other children with similar error patterns.

In summary, this study attempted to examine whether parents of speech-delayed children have an advantage in understanding what they say. The results obtained suggest that mothers in particular may have some advantage.
at least in so far as imitated single words on audiotape are concerned. The exact source of the maternal advantage is however, unclear.

The implication of this advantage is that one must be cautious in using a parent's ability to understand their child as an index of the child's communicative skill. During screening interviews, clinicians would be well advised to include inquiries about how well unfamiliar listeners understand the child. A final recommendation for further research would be to examine whether clinicians have an advantage in judgments of intelligibility on their own clients as a result of personal familiarity.

The author would very much like to thank the parents and their 4 children for participating in this study. As well, the assistance of the unfamiliar judges is greatly appreciated. The insightful comments of two anonymous reviewers also proved very helpful. Finally, this study could not have been completed without the patient statistical analysis of Douglas Kelker of the Statistical Services Centre at the University of Alberta's Department of Statistics and Applied Probability.

REFERENCES


Manuscript received November 12, 1993; Revised May 31, 1994.