The Effect of Speaker’s Gender and Number of Syllables on the Perception of Words by Young Children: A Developmental Study

Limor Adi-Bensaid and Tova Most
Tel-Aviv University

Abstract
The present study examined the effect of a speaker’s gender and the number of syllables per word on the perception of words by children. The participants included 39 children across four age groups: 3, 4, 5 and 6 years. Words differing in length were presented to the children by male and female speakers. Percent correct perception was calculated for each child. The results demonstrated lower perception of words by the 3-year-olds than by the 5- and 6-year-olds. Perception of words produced by the female speaker was better than perception of words produced by the male speaker, and perception of the polysyllabic words was better than perception of the monosyllabic words.

Keywords
development, speaker’s gender, speech perception, word length

Speech is a complex acoustic signal that contains both linguistic and paralinguistic information. Linguistic information includes segmental and suprasegmental as well as lexical, grammatical, and semantic features. Paralinguistic information contains extralinguistic cues that serve to identify speaker characteristics such as age, gender, voice quality, emotional state, and physical state (Abercrombie, 1967). Both sources of information (linguistic and paralinguistic) play a critical role in the perception of speech by listeners. The present study addressed the effect of extralinguistic information (speaker’s gender) and prosodic information (word length) on the perception of speech by children.

EFFECT OF AGE ON SPEECH PERCEPTION
Many previous studies have examined the effect of age on various aspects of speech perception (e.g., acoustical cues, phonemes, words). Many of these studies compared the performance of young children to that of adults and reported that children are less sensitive to perceptual cues than adults (Elliott, 1986; Elliott, Longinotti, Meyer, Raz, & Zucker, 1981; Elliott & Hammer, 1988; Sussman & Carney, 1989; Sussman, 1993; Morrongiello, Robson, Best, & Clifton, 1984; Nittrouer & Studdert-Kennedy, 1987). For example, Ryalls and Pisoni (1997) investigated the effect of talker variability on word recognition. They found that children between the ages of 3 to 5 years were less accurate than adults at identifying words produced by multiple speakers than those spoken by a single speaker, regardless of whether the words were produced in quiet or in noise. Another set of these studies compared speech perception of children across different ages and adults. For example, Hazan & Markhan (2004) compared the performance of 7- to 8-year-olds with 11- to 12-year-olds and with a group of adults. They found that the 7- to 8-year-olds made significantly more errors than the 11- to 12-year-olds and the adults in a word recognition task with background noise. Drager, Clark-Serpentine, Johnson, and Roeser (2006) investigated the perception of words and sentences in background noise by children aged 3 to 5 years. They reported that the 3-year-olds performed more poorly than the 4- and 5-year-olds. These researchers used synthesized words and sentences that were digitized using the speech of an 11-year-old female speaker.

Some researchers claim that the poorer performance of children in comparison to adults is due to the children's immature sensory processing in either the peripheral or central auditory systems for both speech and nonspeech auditory stimuli (Elliott, 1986; Elliott et al., 1981; Hall & Grose, 1991; Sussman & Carney, 1989). Other investigators claim that it is the inability of younger children to attend selectively to the task at hand that limits their performance (Allen, Wightman, Kistler, & Dolan, 1989; Morrongiello et al., 1984; Wightman, Allen, Dolan, Kistler, & Jamleson, 1989). Allen et al. (1989) concluded that processing efficiency (i.e., the ability to filter interfering noise), frequency resolution and listening performance are abilities that improve with age. Thus, better listening performance is due to maturation of the central nervous system. In other words, increasing age improves the ability to allocate the attentional mechanism.

Some studies have suggested that the type of task used will determine the age at which children will demonstrate adult-like speech perception performance. For example, when the task requires discrimination of temporal and frequency cues, performance does not become adult-like until the age of 10 to 11 years (Allen & Wightman, 1992; Sussman & Carney, 1989). However, when the task is an identification one, adult-like performance for speech sounds occurs earlier, (i.e., at about 6 years of age) (Sussman & Carney, 1989; Walley & Carrell, 1983).

EFFECT OF GENDER ON SPEECH INTELLIGIBILITY AND SPEECH PERCEPTION
Male and female glottal characteristics differ considerably (Hanson, 1995; Klatt & Klatt, 1990), and listeners are generally able to distinguish male from female voices quite easily (Nygard, Som-
mer, & Pisoni, 1994; Tielen, 1992). However, gender differences are also noted in patterns of speech production. Based on an analysis of 1,680 phonetically-transcribed utterances produced by 168 U.S. English speakers in the TIMIT (Texas Instruments in Conjunction with the Massachusetts Institute of Technology) database, Byrd (1994) reported that male speech is characterized by a greater prevalence of phonological reduction phenomena than speech produced by females. These phenomena include, for example, vowel centralization, alveolar flapping, a reduced frequency of stop releases. Thus, there is some evidence that gender is a salient characteristic that could affect overall intelligibility. In fact, Bardlow, Torretta, & Pisoni, (1996), as well as Markham and Hazan (2004) found that female talkers received a significantly higher overall intelligibility score than male talkers. The intelligibility scores were awarded by a group of listeners who were asked to identify sentences that they heard. These results raised the question of what specific acoustic-phonetic characteristics led to this gender-based intelligibility difference.

Fundamental frequency. Fundamental frequency ($F_0$) is a global speaker characteristic that typically differs markedly across males and females. The mean $F_0$ of a speaker is related to the average pitch of a person’s voice. However, it is not clear that it is an acoustic attribute that directly affects speech intelligibility. Bradlow et al. (1996) found a significantly greater $F_0$ range for a group of female talkers (mean = 175 Hz) than for a corresponding group of male talkers (mean = 103 Hz). They claimed, however, that the higher $F_0$ mean of women might be only one of the female speech characteristics that contribute to the generally higher intelligibility of female speech relative to male speech. Other characteristics to consider are vowel production, speaking rate, and loudness.

Vowel production. Adult female vocal tracts tend to be shorter than those of adult males. Furthermore the pharynx takes up a greater proportion of overall vocal tract length in adult males than in adult females. Consequently, female formants (i.e., amplitude peaks in the frequency spectrum) tend to be higher in frequency.

Additionally, women typically produce more distinct vowels than men (Labov, 1972). This gender-based difference in the production of vowels has been demonstrated for English, Swedish, French, and Dutch speakers (Henton, 1995), as well as for Korean speakers (Yang, 1996). Furthermore, women lead vowel change by producing longer and clearer vowel variants (Jacewicz, Fox, & Salmons, 2006). Men, however, lead sound changes that further reduce the distance between vowels (Heffernan, 2007). Relation between vowel space and speech intelligibility reveals that talkers with larger vowel spaces were generally more intelligible than talkers with reduced spaces (Bradlow et al., 1996). The findings of Diehl, Lindblom, Hoemeke, and Fahey (1996) also support the assumption that higher $F_0$ may interfere with vowel identification. Thus females, who have higher $F_0$s automatically compensate by depressing vowels more than men.

Speaking rate. Speaking rate is known as one of the most salient global talker-specific characteristics (Picheny, Durlach & Braida, 1989; Krause & Braida, 1995). In fact, many phonological reduction phenomena are directly related to changes in speaking rate. The effect of speaking rate may be related to greater articulatory precision in slower speech (Krause & Braida, 2002; 2004). However, the relation between speech rate and speech intelligibility is unclear. In her analyses of the TIMIT data, Byrd (1994) found that males had significantly faster speaking rates than females on the two calibration sentences that were read by all talkers, and this was true across all dialects of English. However, Byrd’s study also found an interaction of gender and dialect region such that the slowest speaking region for the male speakers (the South Midland) was only the fourth slowest for the female speakers. Krause and Braida (1995) reported that trained talkers were considered as having intelligible speech at even faster speaking rates. Thus, they concluded that it is possible to produce fast intelligible speech. Bradlow et al. (1996) did not find a clear relationship between speech rate and overall speech intelligibility scores. There was no correlation between speaking rate and speech intelligibility across all 20 talkers, and there was no significant difference in the means between the male and female speaking rates.

Loudness. The voice intensity of men is generally louder than that of women (Borden, Harris, & Raphael, 1994). However, it is unclear whether there is a relation between voice intensity and speech intelligibility. Gordon-Salant (1986) and Hazan and Simpson (1998) both found that Consonant-Vowel (CV) syllables with artificially enhanced CV amplitude ratios were more intelligible than syllables with smaller differences in amplitude between the consonant and vowel, as did Montgomery and Edge (1988) when testing listeners with hearing-impairment. However, speech from different talkers has not typically been examined for the effects of naturally occurring CV amplitude differences on perception.

EFFECT OF GENDER ON CHILDREN’S SPEECH PERCEPTION

Studies have shown that infants as young as two months can discriminate between different speakers (DeCasper & Fifer, 1980; Jusczyk, Pisoni, & Mullenix, 1992) and can generalize speech sounds across different speakers (Kuhl, 1979; 1983). However, the effect of gender on the speech perception of children is not clear. Hazan and Markham (2004) investigated acoustic-phonetic correlates of intelligibility for adults and children. Word intelligibility was measured for 45 talkers (women, men and children) by sampling three groups of listeners including adults (29.9 years), older children (11 to 12 years), and younger children (7 to 8 years). Their findings showed that women were perceived to be more intelligible than the other groups, and the perception scores of the younger children were lower than those of the older children and adults. However, there was no significant talker group by listener group interaction.

Petrini and Tagliafietta (2008) investigated the perceptual criteria used by children at different ages (5 to 6 years; 8 to 9 years; 10 to 11 years) and by adults (25 to 30 years) to judge talker similarity by changing two acoustic properties of the voice: pitch (i.e., perceived $F_0$) and speech rate (measured in syllable per second). The youngest children (5 to 6 years) were found to focus on pitch information in their judgments. However, 8- to 9- year-old children were also able to consider variations in speech rate. The 10- to 11-year-old children and adult groups considered both the rate and pitch information.
EFFECT OF WORD LENGTH ON SPEECH PERCEPTION

The most relevant sublexical unit in both the recognition and production of a word is the syllable (Aichert & Ziegler, 2005; Hofmann, Stenneken, Conrad, & Jacobs, 2007). The experimental manipulation of the number of syllables within words has repeatedly resulted in longer response latencies for stimuli with increasing numbers of syllables. This has been reported in production tasks such as picture naming (Klapp, Anderson, & Berrian, 1973; Santiago, MacKay, Palma, & Rho, 2000), word naming (Ferrand & New, 2003, Lee, 2001; Stenneken, Conrad, & Jacobs, 2007), two-digit number naming (Spoehr & Smith, 1973), and in making same-different judgments during perception tasks (Klapp, 1971). However, the results are not always consistent. For example, the typical syllable-number effect has been found to be nonexistent in some picture- and symbol-naming studies (Bachoud-Levy, Dupoux, Cohen, & Mehler, 1998; Forster & Chambers, 1973). The influence of syllable number has seldom been investigated in lexical decision. However, one recent study reported a syllable-number effect for lexical decision on low-frequency words in French (Ferrand & New, 2003). Interpretations of the results are generally constrained by the fact that the number of syllables is typically correlated with word length.

Few studies have investigated word-length effects on speech perception. In fact, a remarkably large portion of the word recognition literature has been entirely based on monosyllables, the shortest possible words. This is especially true for visual word recognition (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001), but it is also the case for auditory spoken word recognition (Gaskell & Marslen-Wilson, 1997; 1999; Goldinger, Luce, & Pisoni, 1989; Kessler & Treiman, 2002; Marslen-Wilson & Warren, 1994; Vitevitch & Luce, 1999). For example, Frauenfelder and Peeters (1990), conducted four different studies to compare the perception of the final fricatives /s/ and /sh/ in monosyllabic versus trisyllabic words. Results showed that peak (asymptotic) activation is positively associated with the number of phonemes in a word. The longer the word, the more highly it is activated. In other words, they claimed that long words should produce stronger lexical activation than short words. Two reasons were provided to explain this. First, long words provide more bottom-up evidence than short words. Secondly, short words are subject to greater inhibition due to the existence of more similar words. For example, there are more phonemes to support the lexical activation of the word ‘dangerous’ than of the word ‘bet’. There should also be more lexical competition for ‘bet’ than for ‘dangerous’, because there are many words that are phonetically similar to ‘bet’ (e.g., ‘bit, ‘met, ‘bed’) but none that are equally close to ‘dangerous’. For future studies, the authors recommended using words of varying length in order to explore the impact of lexical processes more systematically.

EFFECT OF WORD LENGTH CHILDREN’S SPEECH PERCEPTION

No previous reports have addressed the effect of word length on the perception of speech by typically developing children. Kirk, Hay-McCutcheon, Sehgal, & Miyamoto (2000) examined the effect of lexical difficulty (lexically easy and difficult words), talker variability (single male speaker and multiple speakers), and word length (monosyllabic and multisyllabic words) on spoken word recognition by children (5 years and older) with cochlear implants. Results demonstrated that multisyllabic words were identified with significantly greater accuracy than monosyllabic words.

PURPOSE

The purpose of the present study was to examine the effect of the speaker’s gender and the number of syllables on the perception of words by young normally-developing children from 3 to 6 years of age. We hypothesized that perception would improve as the children got older. We further hypothesized that more words spoken by a female would be perceived accurately than those spoken by a male and that longer words would be perceived more accurately than monosyllabic words.

Table 1. The word list with English translations in parentheses

<table>
<thead>
<tr>
<th>1-Syllable Words</th>
<th>2-Syllable Words</th>
<th>3-Syllable Words</th>
<th>4-Syllable Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>pil (elephant)</td>
<td>bamba (snack)</td>
<td>otobus (bus)</td>
<td>lexayaim (cheeks)</td>
</tr>
<tr>
<td>tof (drum)</td>
<td>balon (balloon)</td>
<td>ambatyia (bath)</td>
<td>mixnasaim (trousers)</td>
</tr>
<tr>
<td>dag (fish)</td>
<td>adom (red)</td>
<td>yakevet (train)</td>
<td>ofanaim (bike)</td>
</tr>
<tr>
<td>tut (strawberry)</td>
<td>kelev (dog)</td>
<td>oznaim (ears)</td>
<td>kaduţregel (football)</td>
</tr>
<tr>
<td>lev (heart)</td>
<td>baiet (house)</td>
<td>bęaox (chick)</td>
<td>ipopotam (hippopotamus)</td>
</tr>
<tr>
<td></td>
<td>mastik (chewing)</td>
<td>telefon (phone)</td>
<td>nakniyia (hot-dog)</td>
</tr>
<tr>
<td></td>
<td>kova (hat)</td>
<td>iparıyon (pencil)</td>
<td>mısşanaim (scissors)</td>
</tr>
<tr>
<td></td>
<td>juvan (table)</td>
<td>agala (cart)</td>
<td>mjikfaim (glasses)</td>
</tr>
<tr>
<td></td>
<td>yaldia (girl)</td>
<td>kaduţyrım (balls)</td>
<td>taynegolet (hen)</td>
</tr>
<tr>
<td></td>
<td>bakbuk (bottle)</td>
<td>sukańya (candy)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>banana (banana)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>mityiya (umbrella)</td>
<td></td>
</tr>
</tbody>
</table>
## METHOD

### PARTICIPANTS

Thirty-nine children (13 boys, 26 girls) participated in this study. These children were divided into four groups defined by the chronological ages of 3, 4, 5, and 6 years. The group of 3-year-olds included 9 children (3 boys and 6 girls). The three remaining groups each included 10 children. Specifically, the 4-year-olds included 3 boys and 7 girls. The 5-year-olds included 3 boys and 7 girls; and the 6-year-olds included 4 boys and 6 girls. All children were native Hebrew speakers. All had hearing within the normal range (see the procedure section) and all were free of speech and language difficulties.

### MATERIALS

Thirty-six words representing four syllable lengths were chosen from a large pool of words that had been created by 10 speech-language pathologists (SLPs) working with young children. All words on the list were judged by at least 7 of the 10 SLPs to be in the vocabulary of most 2-year-olds. All words were common and familiar to young children, and they represented different semantic categories (e.g., animals, toys, body parts, food). The list of words is presented in Table 1 below.

Table 1 shows that the list of words included five 1-syllable words (e.g., pil [elephant], dag [fish]), ten 2-syllable words (e.g., kelev [dog], balon [balloon]), twelve 3-syllable words (e.g., banana [banana], mityija [umbrella]) and nine 4-syllable words (e.g., taynegol [hen], ipopotam [hippopotamus]).

All the speech materials were first recorded in an audiometric test booth by a native Hebrew-speaking male and female with clear voice and articulation using Sony's Sound Forge 7 software. The speakers' voice and speech naturalness and typicality were evaluated by 2 independent SLPs. The mean F₀ values of the male (100 Hz) and the female (250 Hz) talkers represented the F₀ of their gender (Borden et al., 1994). The talkers were instructed to read the words aloud in normal voice and speaking rate.

The recordings were normalized so that a consistent intensity was maintained throughout the recordings. The order of the male recordings and female recordings was mixed, and then the words were recorded on a CD in the following way: After every 3 different words presented by the male there were 3 other words presented by the female and so on. After every 3 words there was a change of speaker so that the listeners would not get adjusted to the speaker. The words by the male and female were presented in a counterbalanced order. There were a total of 72 recordings (36 words X 2 speakers [a male and a female]).

### PROCEDURE

Each child was tested individually in a quiet room. There were four stages in the procedure:

#### Stage 1: Hearing Screening.
To confirm normal hearing, each child's hearing thresholds were screened at 500 Hz, 1 KHz, 2 KHz and 4 KHz in each ear through the use of a portable audiometer (type Amplaide A177) and earphones (TDH 49).

#### Stage 2: Familiarization.
Procedures were used to ensure that each child was familiar with the test materials and to assess his or her production of the words that were later presented in the perception test (see Stage 3 below).

To this end, each child was presented with a list of 72 words, 36 of which were later included in the perception test. A female examiner (other than the female speaker for the perception test) presented each of the 72 words once, using live voice. The child was asked to imitate each of the presented words (e.g., “say a 'doll’”). Two other examiners transcribed the child's productions. The transcriptions were later compared to assess inter-transcriber reliability. The latter was determined by dividing the number of agreements by the total number of productions, yielding an inter-transcriber reliability score of 90%. Since all the words were common and familiar to the children, they performed the task with no difficulties during this stage.

#### Stage 3: Perception Assessment.
Using the CD described earlier, all children were presented with 36 words. Each word was presented twice, once by the male and once by the female speaker. Each presentation was delivered at 35 dB SL to the better ear. When both ears were similar, the words were presented to the right ear. Additionally, the words were presented in the presence of speech noise generated by the audiometer in a 0 signal to noise ratio (SNR). Speech-noise was used in order to avoid a ceiling effect (i.e. perception around 100% accuracy), since the children had normal hearing. The SNR level was chosen based on data collected in a pilot study with 6-year-old children whose perception scores were 80% correct at this level. At the beginning of this perception assessment, each child was instructed to repeat what he or she heard. In order to avoid age related difficulties such as lack of motivation or poor concentration during the word presentations, the children were given non-contingent breaks intermittently.

#### Stage 4: Scoring Responses.
Each child received a percent correct perception score by dividing the number of correct responses by 72 (i.e., the total number of response opportunities) during the perception test. The responses of each child were transcribed by two examiners. In order to achieve inter-transcriber reliability the same procedure that was used in the familiarization stage was also used here. The inter-transcriber reliability score was 90%. Responses that were identical to the recorded model were counted as correct. Responses that differed from the recorded model were counted as correct if they matched the child's production of the same word during the familiarization task. For example, if a child produced [akevet] for /yakevet/ ('train') during the perception task (i.e., deleting /y/ in the initial syllable), and if this same production was observed during the familiarization task, then the response was scored as a correct. However, if an incorrect response produced during the perception task did not match the child's production of that word during the familiarization task, then the child did not get credit for that response.

### RESULTS

Table 2 presents the means and the standard deviation (SD) of the correct perception of words presented proportion of words perceived correctly when spoken by a male and a female speaker as well as the total correct perception by the children in the four age groups.

Table 3 presents the means and the SD of the correct percep-
tion of words proportion of words perceived correctly as a function of their syllable length (1-syllable, 2-syllable, 3-syllable, 4-syllable) by the children in the four age groups.

A MANOVA with repeated measures was performed, with age as the independent between variable and speaker's sex and word length as the within variables. The correct perception of the words was the dependent variable. The analysis revealed a statistically significant main effect of age F(3,31) = 6.12, p < 0.002, a significant main effect of speaker sex F(1,31) = 63.53, p < 0.0001 and a significant main effect of word length F(3,29) = 10.06, p < 0.0001. There were no significant 2-way and 3-way interactions. Regarding age group, multiple contrast analysis (Ryan Einot Gabriel Welsch test) revealed that the 3-year-olds performed significantly differently than the 5- and 6-year-olds (p < 0.05). The other groups were not significantly different from one another (p > 0.05). With regard to speaker's gender, the perception of the female speaker was significantly better than the perception of the male speaker. With regard to word length, successive contrast analysis revealed that the perception of the 1-syllable words was significantly lower than the perception of the 2-syllable words F(1,31) = 31.37, p < 0.0001. However, the perception of the 2-syllable words was not significantly different than the perception of the 3-syllable words, and the perception of 3-syllable words was not significantly different than the perception of the 4-syllable words (p > 0.05).

**DISCUSSION**

The purpose of the present study was to examine the effect of two variables (a speaker's gender, word length) on the perception of words by young children ranging in the age from 3 to 6 years.

Our first hypothesis was that speech perception in general would improve with age. Our findings confirmed this hypothesis. The perception scores of the 3-year-olds were significant-ly lower than the perception scores of the 5- and 6-year-olds. These results are consistent with the findings of Drager, Clark-Serpentine, Johnson, and Roeser (2006), who investigated the perception of words and sentences in noise by children ranging in ages from 3 to 5 years. They reported that the performance of the 3-year-olds was poorer than that of the 4- and 5- year olds. Similar to the Drager et al. study, the present study examined the performance of young children in the age range between 3 to 6 years. However, the present study differed from Drager et al. in the stimuli used for the study. In the present study, the stimuli consisted of natural speech produced by adults. The Drager et al. study used synthesized words and sentences that were digitized using the speech of an 11-year old female. The results of the present study showed an improvement in the perception of natural speech across age groups. The older children (i.e., 5- and 6-year-olds) performed significantly better than the youngest children (i.e., the 3-year-olds).

The present findings support the position that children's speech perception skills do continue to develop through early childhood (Elliott, 1986; Elliott et al., 1986; Elliott et al., 1981; Sussman & Carney, 1989). Since speech perception abilities of children have been characterized as ‘developing’ (Elliott, 1986; Elliott et al., 1986; Elliott et al., 1981; Sussman & Carney, 1989), further research with more groups of children and a group of adults will contribute to the understanding of this developmental process.

As mentioned previously, target words in the present study were presented to the children in the presence of noise in order to avoid a ceiling effect. Although the background noise was present for all of the age groups, the 3-year-olds performed more poorly than the older groups. Several explanations may account for these results. Markham and Hazan (2004) noted that the poorer perception of speech by children in degraded conditions

---

**Figure 1. Correct recognition (%) of words spoken by male and female as a function of age**

**Table 2.** Means and SD (%) of the proportion of words perceived correctly when spoken presented by a male and a female speaker by the children in the four age groups.

<table>
<thead>
<tr>
<th>Age Group</th>
<th>N</th>
<th>Mean (%)</th>
<th>SD</th>
<th>Mean (%)</th>
<th>SD</th>
<th>Mean (%)</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-year-olds</td>
<td>9</td>
<td>50.89</td>
<td>10.58</td>
<td>64.78</td>
<td>7.61</td>
<td>58</td>
<td>6.81</td>
</tr>
<tr>
<td>4-year-olds</td>
<td>10</td>
<td>57.1</td>
<td>13.77</td>
<td>72.6</td>
<td>14.16</td>
<td>65</td>
<td>12.47</td>
</tr>
<tr>
<td>5-year-olds</td>
<td>10</td>
<td>66.7</td>
<td>5.81</td>
<td>84.9</td>
<td>5.96</td>
<td>75.5</td>
<td>3.67</td>
</tr>
<tr>
<td>6-year-olds</td>
<td>10</td>
<td>71</td>
<td>10.14</td>
<td>84.3</td>
<td>6.36</td>
<td>77.5</td>
<td>6.99</td>
</tr>
</tbody>
</table>

**Table 3.** Means and SD of the proportion of words perceived correctly percent of correct perception of words as a function of their length by children in the four age groups.

<table>
<thead>
<tr>
<th>Age Group</th>
<th>1-Syllable Words</th>
<th>2-Syllable Words</th>
<th>3-Syllable Words</th>
<th>4-Syllable Words</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (%)</td>
<td>SD</td>
<td>Mean (%)</td>
<td>SD</td>
</tr>
<tr>
<td>3-year-olds</td>
<td>43%</td>
<td>15.81</td>
<td>63%</td>
<td>14.14</td>
</tr>
<tr>
<td>4-year-olds</td>
<td>51%</td>
<td>15.23</td>
<td>71%</td>
<td>16.33</td>
</tr>
<tr>
<td>5-year-olds</td>
<td>52%</td>
<td>16.86</td>
<td>81%</td>
<td>9.94</td>
</tr>
<tr>
<td>6-year-olds</td>
<td>62%</td>
<td>25.73</td>
<td>83%</td>
<td>11.59</td>
</tr>
</tbody>
</table>
is partly due to their poorer use of linguistic and contextual information (Eisenberg et al., 2000), as well as their poorer use of sensory information (Nittrouer & Boothroyd, 1990). Also, Fal- lon, Trehub, & Schneider (2000) reported that children required a greater SNR to achieve a level of performance equivalent to that of adults. Other investigators have claimed that it is the inability of younger children to attend selectively to the task at hand that limits their performance (Allen, Wightman, Kistler & Dolan, 1989; Morrongiello et al., 1984; Wightman, Allen, Dolan, Kistler, & Jamieson, 1989). Allen at al. (1989) claimed that processing efficiency (i.e., the ability to filter interfering noise), frequency resolution and listening performance are abilities that improved with age. They concluded that better listening performance was due to maturation of the central nervous system. In other words, increasing age improved the ability to allocate the attentional mechanism.

Our second hypothesis was that words spoken by a female would be perceived better than those spoken by a male. This hypothesis was also confirmed. The perception of words produced by the female speaker was significantly better than the perception of words produced by the male speaker by all four groups of children. Previous studies indicated that males and females differ considerably in their glottal and vocal tract characteristics. Some of the characteristics include differences in the fundamental frequency (F0), vowel formants frequencies, vowel spaces and durations, speaking rate, and loudness. These differences lead to better perception of speech produced by females. Bardlow et al. (1996), for example, examined whether the talker’s gender would be a correlate of variability in intelligibility. They found that the group of 10 female talkers had a significantly higher overall intelligibility score than the group of 10 male talkers. Moreover, the 4 talkers with the highest overall intelligibility scores were female and the 4 talkers with the lowest overall intelligibility scores were males. Markham and Hazan (2004) found the same tendencies in their study. Women talkers as a group appeared to be slightly more intelligible than other talkers. However, high and low intelligibility talkers were present in each of the talkers groups (i.e., men, women and 13-year-old children). Future research should continue to examine what specific acoustic-phonetic characteristics lead to this gender-based intelligibility difference.

Our third hypothesis was that longer words would be perceived accurately more often than monosyllabic words. The findings of the study revealed that the perception of monosyllabic words was significantly lower than the perception of disyllabic words. However, there was no significant difference between the perception of disyllabic words and trisyllabic words or between the perception of trisyllabic words and quadrisyllabic words. To date, there has been very little investigation of word-length effects on speech perception. Most studies used monosyllabic stimuli both with visual word recognition tasks (Coltheart et al., 2001) and spoken word recognition tasks (Gaskell & Marslen-Wilson, 1997; 1999; Goldinger, et al., 1989; Kessler & Treiman, 2002; Marslen-Wilson & Warren, 1994; Vitevitch & Luce, 1999). Frauenfelder and Peeters (1990) proposed that longer words are easier to perceive because they have more acoustic and lexical redundancy and provide more bottom-up information than short words. Moreover, the category of short words (e.g., monosyllabic words) includes more similar words that are candidates for confusion in word perception. This claim can be illustrated by the monosyllabic word tof (drum) used in our study. There are a few other words (e.g., kof [monkey], tof [good], of [chicken], sof [end]), which are phonetically similar to the target word tof which might have confused the children. On the other hand, the trisyllabic word balonim (bal-loons) and the quadrisyllabic word ipopotam (hippopotamus) for example, have fewer or even no phonetically similar candidates. We believe, however, that other factors may influence word recognition, especially of polysyllabic words. Some of these factors may include prosodic aspects, such as words with different stress patterns (i.e., iambic, trochee), words with different syllable structures (i.e., with or without onset or coda; words with medial coda vs. final coda) and words with different segmental structures (i.e., various consonants and vowels constitutes). Other factors may include lexical aspects such as frequent versus infrequent words and non-words (nonsense) versus real words. Future research should continue the examination how these factors may influence perception. However, since there are many lexical and phonological factors that may influence performance, future research may use a different kind of approach to assess the influence of these individual influences. For example, training studies could be conducted in which children are taught novel words for which these factors are controlled.

The present research included only a small corpus of mono-syllabic words. This may have impacted the results. Therefore, we recommend that a future study be conducted with an equal number of items in each of the word length categories. This would make the statistical comparison easier.

In conclusion, the present research revealed that the perception of words in the presence of noise by young children improved as the children got older and when the words were produced by a female rather than by a male speaker. Moreover, longer (polysyllabic) words were perceived more accurately than short (monosyllabic) words. The findings of this study have implications for the clinical assessment and treatment of special populations (e.g., children with hearing impairments). While evaluating speech perception abilities of typically developing children, it would be valuable to consider these factors and present them hierarchically, starting with the easier materials (i.e. polysyllabic words presented by a female speaker) followed by the more difficult ones.

REFERENCES


**AUTHOR INFORMATION**

**LIMOR ADI-BENSAID, PHD**

Department of Communication Disorders
Tel-Aviv University, Israel 52621
Email: limorbensaid@gmail.com

**TOVA MOST, PHD**

School of Education
Tel Aviv University
Israel 69978.
Phone: 972-3-640-8472
Fax: 972-3-640-9477
Email: tovam@post.tau.ac.il

---

1 Current affiliation:

Department of Communication Sciences and Disorders
Faculty of Health Profession,
Ono Academic College, Israel