Scholarship and Professional Work - Communication

College of Communication

2012

Effects of simultaneous speech and sign on infants' attention to spoken language

Jonathan Y. Ting

Tonya R. Bergeson
tbergeso@butler.edu

Richard T. Miyamoto

Effects of Simultaneous Speech and Sign on Infants' Attention to Spoken Language

Jonathan Y. Ting, M.D., **Tonya R. Bergeson, Ph.D.**, and **Richard T. Miyamoto, M.D.**
Babytalk Research Laboratory, Department of Otolaryngology-Head and Neck Surgery, Indiana University School of Medicine, Indianapolis, Indiana

Abstract

Objectives—To examine the hypothesis that infants receiving a degraded auditory signal have more difficulty segmenting words from fluent speech if familiarized with the words presented in both speech and sign compared to familiarization with the words presented in speech only.

Study Design—Experiment utilizing an infant-controlled visual preference procedure.

Methods—Twenty 8.5-month-old normal-hearing infants completed testing. Infants were familiarized with repetitions of words in either the speech + sign (n=10) or the speech only (n=10) condition. Infants were then presented with four six-sentence passages using an infant-controlled visual preference procedure. Every sentence in two of the passages contained the words presented in the familiarization phase, whereas none of the sentences in the other two passages contained familiar words.

Results—Infants exposed to the speech + sign condition looked at familiar word passages for 15.3 seconds and at non-familiar word passages for 15.6 seconds, t(9)=−.130, p=.45. Infants exposed to the speech only condition looked at familiar word passages for 20.9 seconds and to non-familiar word passages for 15.9 seconds. This difference was statistically significant, t(9) = 2.076, p =.03.

Conclusions—Infants' ability to segment words from degraded speech is negatively affected when these words are initially presented in simultaneous speech and sign. The current study suggests that a decreased ability to segment words from fluent speech may contribute towards the poorer performance of pediatric cochlear implant recipients in total communication settings on a wide range of spoken language outcome measures.

Introduction

The acquisition of language is a complex process for young infants. One of the earliest and most important tasks in language acquisition is the formation of a lexicon based on exposure to the native language [1]. Words enter the lexicon as meaningful units after being identified as distinct sound patterns, but this task is made infinitely more difficult by the fact that most speech directed to infants consists of words continuously strung together in sentences.[2]

Despite this challenge, Jusczyk and Aslin [1] demonstrated that 7.5-month-old infants can recognize words in fluent speech. In this study, half the infants were familiarized with the words cup and dog, whereas the other half were familiarized with the words bike and feet. During a subsequent test phase, infants heard four randomly ordered six-sentence passages that each contained one of the target words (bike, cup, feet, or dog) in every sentence. Infants familiarized with the words cup and dog listened significantly longer to passages containing these two words during the test phase than to passages containing the two novel words bike and feet, whereas infants familiarized with the words bike and feet listened significantly longer to passages containing these two words than to passages containing the two novel words cup and dog.

These results provided compelling evidence that infants as young as 7.5 months could extract segments of auditory information from fluent speech that adults treat as words even if these infants had not yet attached meaning to these segments. This ability to segment word forms from fluent speech is a necessary precursor to the mapping of sounds to meaning. [2]

This task is infinitely harder for hearing-impaired infants who receive consistently degraded auditory input. Every year, 12,000 babies are born with some level of permanent hearing loss. Although cochlear implantation prior to the age of 3 years has been associated with improved language outcomes in severely hearing-impaired (HI) children [3], there are several factors that affect speech and language outcomes independent of variables such as age at implantation. One source of variability that contributes to the success and benefit with a cochlear implant is the nature of the early sensory and linguistic environment that deaf children are immersed in after receiving a cochlear implant [4].

Children who use cochlear implants typically receive one of two predominant communication modes: oral communication (OC), in which children utilize auditory/oral skills and total communication (TC), in which children utilize simultaneous signed and spoken English. Oral communication methods can range from auditory-verbal therapy, in which auditory information is heavily emphasized and lipreading is discouraged, to cued speech, in which specific hand cues are used to supplement lip-reading information. Similarly, total communication can range from an emphasis on spoken English to equal emphasis on signed and spoken English, e.g. Signing Exact English, and finally to an emphasis on manual signs.

Numerous studies have reported that children with cochlear implants who are immersed in OC programs have better performance on a variety of spoken language measures than children who are enrolled in TC programs. Children who use OC have significantly better spoken word recognition scores and better expressive language and speech intelligibility skills than children who use TC [5–8]. Interestingly, among children who use TC, those who rely more on speech than sign also have better speech perception and speech intelligibility skills than those who rely more on sign [9].

One proposed explanation for this observation is that simultaneous communication methods used with TC children may result in competition between speech and manual communication for limited attention and processing resources in working memory, both of which are assumed to play major roles in language comprehension and word recognition tasks [10,11].

To examine competition between speech and manual communication for limited attention and processing resources, we hypothesized that infants receiving a degraded auditory signal would have more difficulty segmenting word forms from fluent speech if they were familiarized with the words presented in both speech and sign (as in a TC environment) compared to familiarization with the words presented in speech only (as in an OC

environment). To test this hypothesis, we familiarized normal-hearing infants with words presented in background noise, which has often been used to simulate the degraded auditory signal received by hearing-impaired infants. These words were presented in either *Speech only* or *Speech + Sign* . We then examined the ability of both groups to segment these words from fluent speech. We predicted that the infants who were familiarized with words presented in the *Speech + Sign* condition would have more difficulty segmenting these words from fluent speech due to competition for a limited amount of working memory capacity and information processing skills.

Materials and Methods

Subjects

8.5-month-old normal-hearing infants were enrolled from the greater Indianapolis area with a goal of 10 infants successfully completing each condition for a total of 20 infants. All infants passed newborn hearing screens, were full-term deliveries and had no known developmental problems. Infants who had more than 3 ear infections in the last year or whose caregivers had concerns regarding their hearing underwent an audiogram. Caregivers were reimbursed $10 for participating in the study.

Stimuli—We video-recorded a native female speaker of American English who was also fluent in Signing Exact English producing four passages in an infant-directed manner. The passages were all six sentences in length, and were identical to those used in Jusczyk and Aslin, 1995 (Table 1). Every sentence within each passage contained a common target word, i.e. cup, feet, dog or bike.

The speaker also produced versions of the isolated target words (cup, dog, feet and bike) to be used in the familiarization phase of the experiment; as in Jusczyk and Aslin (1995), the each target word was repeated with some variation 15 times in a row. In addition, the speaker was instructed to produce each list of words in two conditions: a *Speech + Sign* condition, in which the speaker simultaneously spoke and signed, and a *Speech Only* condition, in which the speaker spoke with her hands visibly placed across her chest (see Figure 1). The videos were then edited in Final Cut Pro (Apple Inc., Cupertino, CA) and background multi-talker babble was inserted into all 8 word list videos at a −10 dB signal-to-noise ratio to simulate the more impoverished auditory signal a hearing-impaired child receives.

We used two additional stimuli in this experiment - a silent video of a smiling infant was used as an attention getter before each trial, and a computer-graphic animation consisting of a looming geometric shape paired with a sequence of short, varying tones was used to gauge infants' general attention level before and after the experiment.

Apparatus—Infants were tested in a custom-designed double-walled Industrial Acoustics Company sound booth. Infants sat on their caregivers' laps approximately 5' in front of a 55" widescreen television monitor. To minimize bias, all caregivers wore earplugs and headphones over which a mixture of music and concatenated stimuli from other experiments was played. Visual stimuli were presented at the center of the television monitor at the approximate eye level of the infants. Auditory stimuli were presented through the left and right loudspeakers of the TV monitor between 60–70 dB SPL. A concealed video camera located above the television monitor operated by one experimenter allowed a second experimenter to monitor the infants' looking responses through a closed-circuit television from a separate control booth. This second experimenter, who was blinded to the stimuli being presented, controlled the experiment using the Habit software package [12] on a Macintosh® G5 desktop computer (see Figure 1).

Procedure—Infants' looking times to the videos were coded using the Habit software program[12] .The experiment consisted of a familiarization phase and an infant-controlled test phase. In the familiarization phase, infants received the target word lists in either the *Speech + Sign* condition or the *Speech Only* condition (see Figure 2). Half of the infants in each condition heard the words cup and dog, and the other half heard the words feet and bike. Each word list was presented three times in alternating order, for a total of six familiarization trials of 10 sec each. Between each trial, the attention getter was used to elicit the infant's visual attention before the next trial was initiated. In the test phase, all four passages were presented using an infant-controlled visual preference procedure. Order of presentation of passages was randomized across infants. Prior to each trial, infants' visual attention was elicited by presenting the attention getter. The trial was then initiated by the experimenter and continued until the infants looked away from the video display for longer than 1 sec or until a maximum trial length of 30 sec was reached. Infants' total looking times to the video display for each trial were calculated. Immediately before and after the experiment, infants' alertness was assessed with the looming computer-graphic animation described above.

Results

To ensure that infants had sufficient exposure to the words, any infant who did not look for a total of 55 seconds during the learning phase (out of a possible 60 seconds) was not included in the analysis. This criterion was used to ensure a similar duration of exposure to the words in the familiarization phase as the infants tested in Jusczyk and Aslin's original study (who received at least 30 seconds of familiarization to each word in a modified head-turn preference procedure).

A total of 10 infants in each condition successfully completed testing, for a total of 20 infants. Infants in the *Speech + Sign* condition had an average age of 8.5 months (range 7.0–9.5, S.D. = 1.00) and infants in the *Speech Only* condition had an average age of 8.4 months (range 7.1–9.5, S.D = .93).

Mean looking times for the familiar and non-familiar word passages were calculated for infants in the *Speech + Sign* condition and the *Speech Only* condition. One-tailed paired t-tests were performed to assess for significant differences in mean looking times for familiar and non-familiar word passages (Figure 3).

Infants exposed to the *Speech + Sign* condition looked at the familiar words for a total of 15.3 seconds and at the non-familiar words for a total of 15.6 seconds, $t(9) = -.130$, $p = .45$.

Infants exposed to the *Speech Only* condition looked at the familiar words for a total of 20.9 seconds and to the non-familiar words for a total of 15.9 seconds. This difference was statistically significant, $t(9) = 2.076$, $p = .03$.

Discussion

Numerous studies have reported that children with cochlear implants who are immersed in OC programs do much better on a range of oral language tasks than their counterparts enrolled in TC programs (see Introduction). Tests assessing speech feature discrimination, comprehension, spoken word recognition, receptive and expressive language, and speech intelligibility all consistently show that children in OC programs do much better than children in TC programs. One proposed mechanism for the poorer performance of TC children is competition between speech and manual communication for controlled attention

and limited processing resources in working memory, both of which are assumed to play major roles in language comprehension and spoken word recognition [10].

To examine this possibility, we assessed the ability of 8.5-month-old infants to segment words from fluent speech after exposure to these words in background noise. The current study demonstrates that 8.5-month-old infants familiarized with words in background noise are able to segment these words from fluent speech when they are familiarized in a speech-only condition, whereas they are unable to do so if they were familiarized with a simultaneous speech and sign condition.

There may be several reasons for the inability of infants familiarized with simultaneous speech and sign to segment these words from speech. Under simultaneous communication methods, speech and sign do not specify the same gestures and common underlying articulatory events of the talker. Information from the two sensory channels cannot be integrated quickly in the same way as it is when a listener simultaneously sees the talker's lips and hears speech under cross-modal, audiovisual presentation conditions. Moreover, because the child is looking at the talker's hands rather than his or her face, reliable lipreading cues may not be readily available to provide additional complementary phonetic information about the speech signal that is coded in the visual display of the talker's face. Thus, little facilitation or enhancement can be gained from the manual visual input; if anything, substantial competition and even inhibition effects resulting from two divergent input signals may occur. [11]

The findings in the current study suggest that a decreased ability to segment words from fluent speech may be a contributing factor towards the poorer performance of TC children on a wide range of outcome measures. However, any such extrapolation from the current study should be qualified by the fact that although an attempt was made to simulate a degraded auditory signal by presenting the words in background noise, only normal hearing infants were tested in this experiment. It is uncertain if similar findings would be replicated in children with hearing aids or cochlear implants, whose audiovisual speech perception skills are influenced by hearing impairment and for whom auditory experience plays a role in audiovisual speech perception. [13] Additionally, we were unable to assess whether infants' visual attention was divided between the speakers face and the hand gestures. Examining infant visual attention with an eyetracker system may provide further insight into differences in visual attention to speakers in speech only versus simultaneous speech and sign communication environments.

In addition, this study does not suggest that hearing-impaired infants cannot benefit from TC educational environments. Rather, this study suggests that one underlying factor for the observed difference in speech and language outcomes between TC and OC is a diminished ability to segment words from fluent speech in a degraded auditory environment.

Finally, the use of signing to support pre-speech language development has been popularized as a vehicle for parental support of pre-speech language development. The studies supporting the use of 'baby sign language' were performed in hearing children of hearing parents who were exposed to symbolic gestures, not as a complete language system but as a symbolic and linguistic auxiliary to oral language [14]. While the data supporting the use of 'baby sign language' are equivocal, the current study neither supports nor refutes its use [15]. Rather, the current study examines the possible competitive effects of sign on infant attention and learning in an impoverished auditory environment similar to that of a hearing-impaired infant.

Conclusion

In conclusion, we found that the ability of 8.5-month-old infants to segment words from fluent speech after prior exposure to these words presented in background noise was affected by the presence of simultaneous manual communication. We propose that the competition between speech and manual communication for limited processing resources as well as the inability to integrate cross-modal redundant facial cues for speech processing are the factors that explain these findings. Similarly, these factors may partially explain previously reported disparities between the performance of children with cochlear implants in TC and OC programs across a wide range of outcome measures. Future experiments in children with hearing impairment as well as eye-tracking studies may provide further insight on this issue.

References

1. Jusczyk PW, Aslin RN. Infants' detection of the sound patterns of words in fluent speech. Cognitive Psychology. 1995; 29:1–23. [PubMed: 7641524]

2. Jusczyk, PW. The Discovery of Spoken Language. MIT Press; 2000.

3. Nicholas JG, Geers AE. Effects of early auditory experience on the spoken language of deaf children at 3 years of age. Ear Hearing. 2006; 27:286–98.

4. Bergeson, TR.; Pisoni, DB. Audiovisual speech perception in deaf adults and children following cochlear implantation. In: Calvert, G.; Spence, C.; Stein, BE., editors. Handbook of Multisensory Processes. MIT Press; 2004. p. 749-72.

5. Cullington H, Hodges AV, Butts SL, Dolan-Ash S, Balkany TJ. Comparison of language ability in children with cochlear implants placed in oral and total communication education settings. Annals of Otology,Rhinology, and Laryngology. 2000; 185:121–3.

6. Kirk, K.; Pisoni, D.; Miyamoto, R. Lexical discrimination by children with cochlear implants: Effects of age at implantation and communication mode. In: Waltzman, SB.; Cohen, NL., editors. Cochlear implants. 2000.

7. Svirsky MA, Robbins AM, Kirk KI, Pisoni DB, Miyamoto RT. Language development in profoundly deaf children with cochlear implants. Psychol Sci. 2000; 11:153–8. [PubMed: 11273423]

8. Tobey EA, Geers AE, Douek BM, et al. Factors associated with speech intelligibility in children with cochlear implants. Ann Otol Rhinol Laryngol Suppl. 2000; 185:28–30. [PubMed: 11140993]

9. Geers A, Spehar B, Sedey A. Use of speech by children from total communication programs who wear cochlear implants. Am J Speech-Lang Pat. 2002; 11:50–8.

10. Baddeley A, Gathercole S, Papagno C. The phonological loop as a language learning device. Psychol Rev. 1998; 105:158–73. [PubMed: 9450375]

11. Doherty-Sneddon G, Bonner L, Bruce V. Cognitive demands of face monitoring: Evidence for visuospatial overload. Mem Cognition. 2001; 29:909–19.

12. Cohen, LB.; Atkinson, DJ.; Chaput, HH. Habit X: A new program for obtaining and organizing data in infant perception and cognition studies. 1.0. Austin: University of Texas; 2004.

13. Bergeson TR, Houston DM, Miyamoto RT. Effects of congenital hearing loss and cochlear implantation on audiovisual speech perception in infants and children. Restorative Neurology and Neuroscience. 2010; 28:157–65. [PubMed: 20404405]

14. Goodwyn SW, Acredolo LP, Brown CA. Impact of symbolic gesturing on early language development. Journal of Nonverbal Behavior. 2000; 24:81–103.

15. Johnston JC, Durieux-Smith A, Bloom K. Teaching gestural signs to infants to advance child development: A review of the evidence. First Language. 2005; 25:235–51.
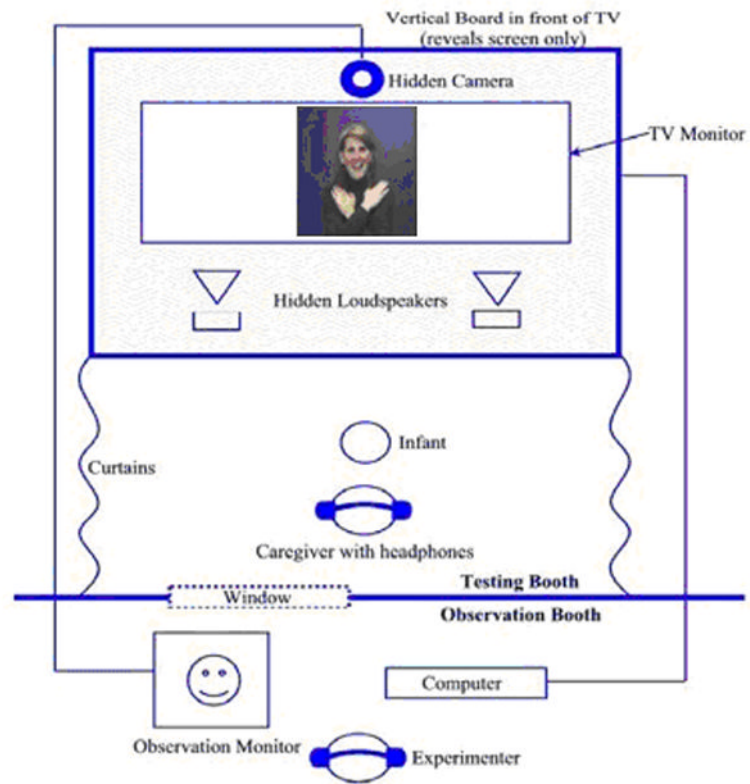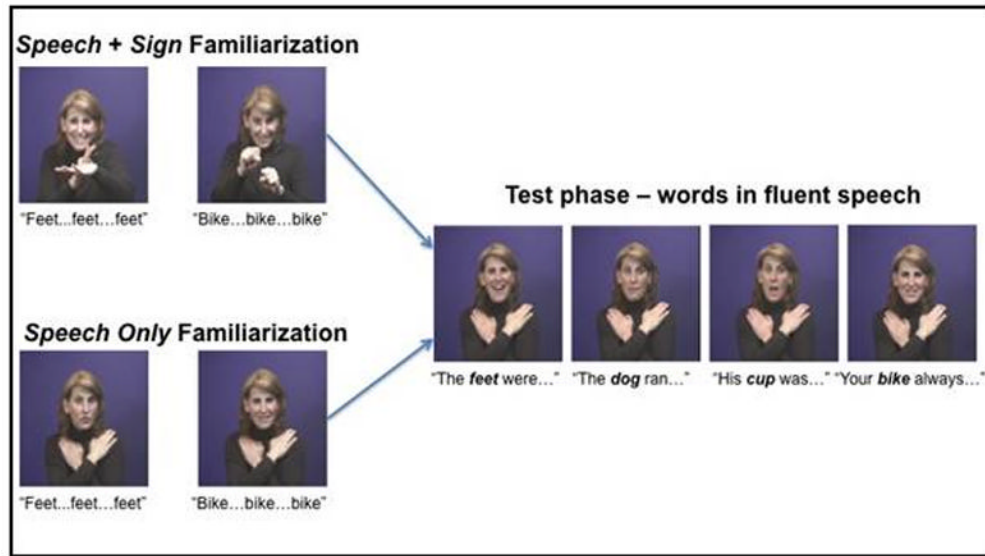
Figure 1.
Diagram of setup of apparatus.
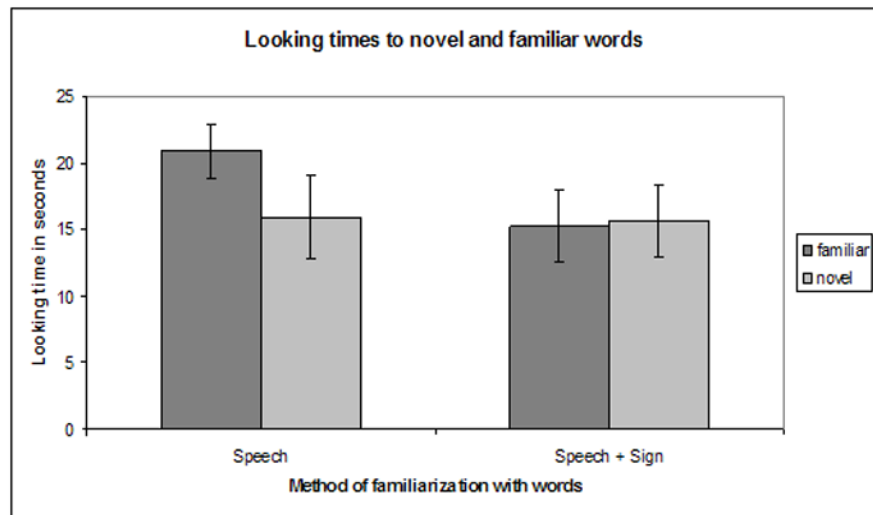
Figure 2.
Schema of experiment

Figure 3.
Mean looking times to familiar and non-familiar words during the test phase

Table I

Six-sentence passages used in experiment (see Jusczyk & Aslin, 1995)

Cup	The cup was bright and shiny.
	A clown drank from the red cup.
	The other one picked up the big cup.
	His cup was filled with milk.
	Med put her cup back on the table.
	Some milk from your cup spilled on the rug
Dog	The dog ran around the yard.
	The mailman called to the big dog.
	He patted his dog on the head.
	The happy red dog was very friendly.
	Her dog barked only at squirrels.
	The neighborhood kids played with your dog.
Feet	The feet were all different sizes.
	This girl has very big feet.
	Even the toes on her feet are large.
	The shoes gave the man red feet.
	His feet get sore from standing all day.
	The doctor wants your feet to be clean.
Bike	His bike had big black wheels.
	The girl rode her big bike.
	Her bike could go very fast.
	The bell on the bike was really loud.
	The boy had a new red bike.
	Your bike always stays in the garage.