INTERSPEECH 2023
20-24 August 2023, Dublin, Ireland

Did you see that? Exploring the role of vision in the development of consonant feature contrasts in children with cochlear implants

James Mahshie[1], *Michael Larsen*[2]

[1]George Washington University, Washington, DC
[2]St. Michaels College, Colchester, Vermont USA
`jmahshie@gwu.edu, mlarsen@smcvt.edu`

Abstract

This project aimed to explore the potential role of vision in speech contrast production and auditory perception development in children with cochlear implants (CWCI). Ten CWCI between 43 and 61 months of age, with at least 2 years of CI experience, served as participants. Employing an auditory imitation task, children's ability to auditorily perceive contrasts that are more or less visible was examined both at baseline and one year after the initial assessment. The children's ability to produce these contrasts was also examined through a picture-naming task. The CWCI tended to produce features in both visibility conditions with greater accuracy than they perceived, both at baseline and at 1 year. Production and perception accuracy increased after one year of CI usage, with the mean perceptual gain for the more visible contrasts exceeding that of the less visible contrasts. The implications of the role of vision in contrast development are discussed.

Index Terms: speech perception, speech production, segment visibility, speech feature contrast, children with cochlear implants

1. Introduction

Cochlear implants have had a profound impact on children's access to speech, which in turn has an important and positive impact on the development of both the ability to perceive speech and to produce it. While the timing and rate of development are often behind those of children with typical hearing, children with cochlear implants (CWCI) exhibit significant speech production and auditory perception development during the early years of implant usage (for example, [1]. Various factors have been shown to be significant predictors of the development of these aspects of spoken language, including the development of other aspects of spoken language, duration of implant usage, age at which the implant is activated, and the child's language environment.

Mahshie, et al. [2] reported on a group of 15 children with cochlear implants and their ability to perceive and produce words that reflect a range of speech contrasts, including vowel height and place, front and back places of production, consonant voicing, and consonant manner of production. Our findings revealed that CWCI were generally able to accurately produce utterances that reflect the acoustic distinctions associated with nearly all consonant contrasts. However, these children were less consistent in their ability to accurately perceive these features based on audition alone. More specifically, many of the children exhibited difficulty perceiving acoustic differences that reflected a less visible place of articulation contrasts, consonant voicing, and the difference between continuant and non-continuant consonants. Even when the children received visual and auditory information about the utterance reflecting these contrasts, they exhibited some difficulty perceiving a distinction [3].

Speech is largely a bimodal percept in which we rely on both auditory and visual (AV) information to decipher what is being said. For individuals with typical hearing, vision can play a particularly critical role when the auditory signal is low in intensity or there is a noisy listening environment.

Research suggests that CWCI perform better when presented with AV speech information than when speech is presented auditorily or visually alone [4]. These researchers also found that AV sentence comprehension skills were strongly correlated with measures of language and speech intelligibility.

While AV perception appears to be important for speech perception by CWCI, not all children are equally able to integrate visual and auditory information. Schorr et al. [5] reported that auditory visual fusion in speech perception in CWCI declined with age at implant, suggesting that there is a critical period for bimodal integration for speech perception.

While the lack of visibility and limited audibility of these features could explain the difficulty exhibited by some children with cochlear implants, the role that maturation and experience with cochlear implants might play in the ability to perceive and produce these feature distinctions has not been studied. The goal of the present study was to examine how auditory perception and production of speech features that are more or less visible change over the course of one year for children with cochlear implants.

2. Method

The data for this study are a subset of children with cochlear implants who participated in a study examining the relationship between speech production and auditory perception. The 10 children in this study represent those for whom we were able to collect additional data one year after the initial data collection session.

2.1 Participants

Participants were recruited through advertisements and flyers distributed to clinics, preschools, and parent groups in the metropolitan DC area. Prior to the beginning of the study, the George Washington University's institutional review board reviewed and approved the project, along with a parent written consent form and child assent protocol.

10.21437/Interspeech.2023-1933

Table 1: *Participant characteristics. All ages and durations are in years;months.*

Participant	Age at initial testing	Onset of HL	Age of Activation R	Age of activation L	CI type	Duration of CI Use at initial Testing R	Duration of CI use at initial Testing L	Use of HA before CI
Ci007	3;8	Pre-lingual	1;4	hearing aid	Internal Freedom/External 5	2;4	hearing aid	Yes
Ci009	3;7	Pre-lingual	1;1	1;6	dvance Bionics 90K Body wor	2;6	2;3	Yes
Ci011	3;7	Pre-lingual	2;1	1;1	Cochlear Nucleus Freedom	1;5	2;5	Yes
Ci019	4;0	at 1 yr	2;1	2;3	Cochlear Freedom Nucleus 5	2;0	1;9	Yes
Ci020	3;9	Pre-lingual	hearing aid	1;2	Cochlear Freedom	hearing aid	2;7	Yes
Ci021	4;5	Pre-lingual	0;10	0;10	Cochlear Nucleus Freedom	3;7	3;7	Yes
CI023	5;1	1;10	hearing aid	2;6	Cochlear Nucleus Freedom	hearing aid	2;7	Yes
Ci025	3;2	Pre-lingual	1;7	1;7	Cochlear Nuclear 5	1;7	1;7	Yes
Ci026	3;10	Pre-lingual	1;4	1;7	Advanced Bionics Harmony	2;6	2;3	Yes
Ci028	3;8	Pre-lingual	1;4	1;10	Cochlear Nucleus 5	2;3	1;6	Yes

Ten children between 43 and 61 months of age at the time of the exam served as participants. All children received their implants prior to 30 months of age and had no reported additional disability that could impact speech and language abilities. The children had at least 2 years of experience with their implant prior to their initial assessment. The participant characteristics are shown in Table 1.

2.2 Procedure

To obtain roughly comparable production and auditory perception data, we employed the OLIMSPAC test to examine speech feature perception and developed a set of picturable utterances that were used to examine speech feature production. This process is described below.

Trained research assistants collected data during one or two sessions for each time point; the number of sessions was dependent upon the ability of the child to attend to the tasks. Data were collected either at the child's school or at home in a relatively quiet environment free from distractions. To ensure that each child's cochlear implants worked properly, the Ling six-sound test [6] was administered at the beginning of each session through live voice, in which the children repeated each of the six sounds that were randomly presented.

2.2.1 Auditory Perception of Speech Features

OLIMSPAC [7] was developed for clinical use to assess young children's perception of phonological contrasts. It has been evaluated in both hearing children and those with CIs, and its validity as a measure of auditory perceptual capacity has been established [8]. OLIMSPAC uses 16 recorded nonsense vowel-consonant-vowel (VCV) stimuli to examine a child's ability to perceive six feature contrasts: two vowels and four consonant feature contrasts. The first eight trials simultaneously assessed vowel height, consonant voicing, and front consonant place, while the second eight trials assessed vowel place, consonant manner, and back consonant place. Because multiple contrasts are evaluated with each VCV stimulus, OLIMSPAC samples each contrast eight times. The. Present research focuses on the front place (e.g. /b/ differentiated from /d/), and back place (e.g. / (/ʃ/ differentiated from /s/) contrasts. The front feature contrast is considered more visible, while the back feature contrast is less visible.

In the present study, we used the OLIMSPAC program to present audio-only stimuli to children who were asked to imitate what they heard. Audio–video recordings were made of the child's productions for offline evaluation.

Three listeners scored the imitation by selecting the option from among the eight alternatives that best matched what they heard. This approach is similar to that used for the original Imitative SPAC test (e.g., [9]) and ensures greater validity of judgments than might be obtained from a single judge. The consensus among the three judges was considered the utterance that the child perceived for each of the eight front place and eight back place utterances.

2.2.2 Speech feature production

A picture-naming task was designed to elicit the production of contrasts of interest. We selected words that could be depicted by a picture and were familiar to 3-year-old children.

Two transcribers independently and blindly transcribed children's productions in IPA using broad transcription in PHON [10], a phonological analysis program. Consensus methods of transcription were used so that when the two transcriptions were in agreement, the transcription was taken as the actual production of the utterance. When the two transcribers disagreed, a third transcriber with considerable experience in transcribing children's speech listened to and transcribed the utterance. In all cases, the majority transcription was selected as the transcription of the actual production.

In the present study, a target segment in each word was used to determine if a particular contrast could be produced. For example, if a child correctly produced a front segment such as /b/ or /d/, this was considered evidence that the front place contrast was produced. Similarly, if a child produced a segment containing /ʃ/ or /s/, then the child was scored as a correct production of the back-place contrast.

2.3 Analysis

For each child, the proportion correct for perception and production at baseline and 1 year later for front and back consonants was calculated. Tables and graphs are used to display the results. Statistical significance for comparisons can be assessed using matched-pair t-tests with 9 degrees of freedom (n=10). Given the small number of subjects and the exploratory nature of the investigation, p-values are reported without adjustment. Results are descriptive and could suggest hypotheses for further investigation with a larger sample size.

3. Results

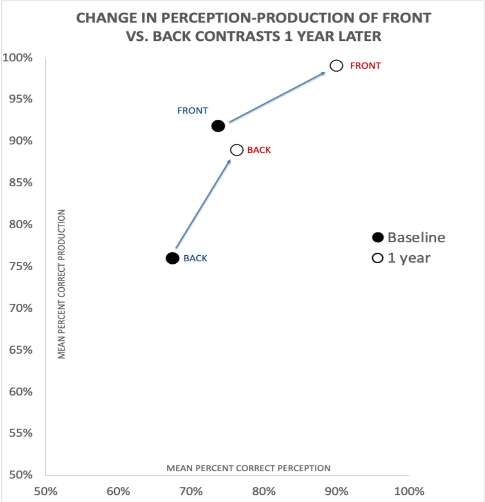

Figure 1: Baseline and 1 year auditory perception and production results.

Figure 1 shows a scatter plot of the accuracy scores for both the production and auditory perception of back and front contrasts. Scores shown at both baseline and 1 year later are the mean speech feature perception and production scores. Table 2 shows the results of the statistical testing conducted.

3.1 Comparison of feature production vs. auditory perception for each place contrast

At baseline, the children produced the front contrasts with significantly greater accuracy than they perceived them; front contrast production accuracy was 91.8% correct while auditory perception accuracy was 73.8% correct (t(9)=4.66, p<.01). A significant difference favoring production was also found after 1 year (t(9)=2.32, p<.05).

While a similar trend was observed for back consonant contrasts at baseline (production = 76%, auditory perception = 67.5%) the difference was not significant (t(9)= 0.87, p=.41). At one year, the difference between auditory perception and production accuracy of back features only approached significance (t(9) =1.92, p=.09).

Table 2: Statistical test results for various contrasts

Comparison	Mean difference	Test statistic	Unadjusted p-value
3.1 Production compared to auditory perception			
front, baseline	18.10%	4.66	0.001**
front, 1 year	9.10%	2.32	0.046*
back, baseline	8.50%	0.87	0.405
back, 1 year	12.80%	1.92	0.087^
3.2 Front compared to back production and auditory perception			
Production baseline	15.80%	2.33	0.044*
Production 1 year	10.10%	1.8	0.106^
Aud. Perception baseline	6.30%	0.79	0.453
Aud. Perception 1 year	13.80%	1.88	0.093^
3.3 Production and auditory perception baseline vs. year 1			
Production - front	7.30%	1.5	0.168
Production - back	13.00%	1.86	0.096^
Aud. Perception - front	16.20%	2.62	0.028*
Aud. Perception - back	8.80%	0.89	0.398

^ = less than 0.10; * = less than 0.05; ** = less than 0.01; *** = less than 0.001

3.2 Comparison of production and auditory perception accuracy for front vs. back place features

Feature production and auditory perception were compared for the two place categories at baseline, and 1 year later. At baseline, the children produced the front consonant contrast with a significantly higher level of accuracy (91.8%), than the back place contrast (76%) (t(9)= 2.33, p<.05). The front contrasts were auditorily perceived with somewhat greater accuracy than the back contrasts (73.8% vs. 67.5%, respectively) but was not statistically significant (t(9) =0.79, p=.45). At 1 year, feature production and auditory perception favored the front segments, but the differences were not significant.

3.3 Production and auditory perception change over 1 year

After 1 year, production of both the front and back place features improved. The mean production accuracy of the front contrasts improved 7.3% to 99.1%, while the mean production accuracy of the back feature contrasts increased by 13% to 89% over the year. While the difference in performance for front contrast production was not significant, auditory perception of the back contrast only approached significance (front: t(9)=1.50, p=0.17; back: t(9)=1.86, p=0.10).

Auditory feature perception also appeared to improve for both the front and back contrasts over 1 year. The largest performance increase was in the perception of front contrasts (16.2% increase in accuracy to 90% accurate). This improvement was statistically significant (t(9)=2.62, p<.05) The performance increase for the perception of the back contrast was considerably smaller after 1 year (8.8% to 76.2% accuracy) and did not reach statistical significance (t(9)=0.89, p=0.398).

Also noteworthy were the changes in variability among the children over the course of 1 year. Both production and auditory perception of both feature categories were more consistently produced by the group after 1 year (See Figure 2), as evidenced by the reduced variability among the group. This was particularly the case for the front contrasts.

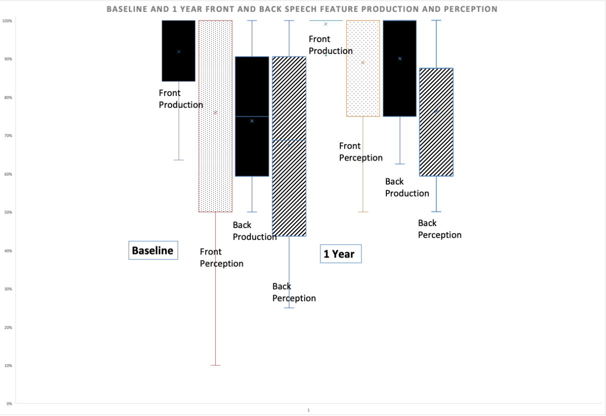

Figure 2: Distribution of mean production and auditory perception scores at baseline and 1 year later.

4. Discussion

This study aimed to examine the potential role of segment visibility on CWCI development of auditory perception and production skills over the course of a year. As noted in previous work [2,9,11], CWCI produce both categories of contrasts more accurately than they perceived them. The reason for the production of contrasts being greater than auditory perception could be the result of a number of factors. Among the factors suggested are the teaching of production and the visibility of segments.

This disparity was greatest for the more visible contrasts, and suggest that visual aspects of these contrasts may provide important information that assists production. In the present study, we found that while the children's production and feature perception improved over the course of 1 year, a disparity persisted between production and auditory perception that favored production; the more visible contrasts were produced and perceived more accurately by the group than the less visible contrasts. This again supports the potential role of vision in developing production and auditory perception.

The children demonstrated performance improvement in both contrast production and perception over the course of 1 year. However, the disparity between the front and back contrasts persisted for both production and auditory perception.

The role of vision in speech perception has been demonstrated by others. Few studies have examined the role that vision might play in the acquisition of speech contrasts, particularly for those with CIs for whom audition provides a more limited information. It is possible that the greater visibility of the front feature contrasts may have contributed to these features being produced and perceived more accurately at baseline. One year later, while feature perception through audition still lagged behind production, the children's ability to auditorily perceive these distinctions increased for both front and back consonants.

These findings lead us to posit a possible mechanism underlying the acquisition of these contrasts. For the front contrasts, children hear and see a distinction, leading to better accuracy at baseline than for the less visible back contrasts. During the year, children continue to develop auditory perception and production of front contrasts based on auditory *and* visual information. For the less visible segments, increased sensitivity to auditory cues is developing, but without the benefit of visual cues. The more limited available information influences both auditory perception and production accuracy, but to a lesser extent than the more visible contrasts. This may account for the 86% greater increase in mean auditory perception of front contrasts than back contrasts (16.2% vs. 8.7% increase for front vs. back contrasts, respectively). Thus the children continue to learn to auditorily perceive less visible contrasts of posterior segments based on less sensory information.

Other factors may also play a role. Overall maturation along with instruction, for example, could play a role in greater accuracy of production. It has also been suggested that production may, in fact, contribute to learning about auditory contrasts [2,8].

There are interventions, such as auditory verbal therapy or AVT, that deemphasize vision in order to promote a child's greater focus on auditory information. This position is not necessarily supported by research [12]. The present findings suggest that visual information may have a positive benefit on both production and perception. The complimentary aspect of visual input, and the importance of developing timely auditory visual fusion in CWCI may also be relevant considerations in intervention decisions.

The present study has limitations. While the sample studied is fairly homogeneous regarding age of implantation and audiological history, the sample size is small and the analysis intended to be suggestive of areas for future research. The findings, however, clearly suggest a difference in how the ability to produce and auditorily perceive these features differ based on the visibility of the contrast. Additional research is needed.

The approach used to obtain perceptual data was imitation. This is an accepted approach to obtaining perceptual information from young children (see [8]), assuming that production abilities do not prevent the child from saying what they hear. All the children in the present study demonstrated some ability to produce the contrasts, suggesting that production was not the limiting factor in the children's imitations.

5. Conclusions

Children produce these place feature contrasts more accurately than they perceive them. During the course of a year, the children as a group developed greater perceptual accuracy for the more visible front place contrasts than for the less visible back place contrasts. While they continued to develop auditory perception of less visible contrasts, the extent of development was less than that found for the front consonants. This difference is perhaps attributable to the increased information provided by vision for the front contrasts.

6. Acknowledgements

The contents of this study were developed under a grant from the National Institute on Disability, Rehabilitation Research (NIDRR grant H133E80006). We are grateful to the children, families, and centers who participated in this project. We also thank Dr. Cynthia Core for her role in developing the production stimulus set and assisting with the transcription of the children's utterances.

7. References

[1] Niparko, J. K., Tobey, E. A., Thal, D. J., et al.; CDaCI Investigative Team. (2010). Spoken language development in children following cochlear implantation. *JAMA*, *303*, 1498–1506.

[2] Mahshie J., Core, C., & Larsen, M. (2015). Auditory Perception and Production of Speech Feature Contrasts by Pediatric Implant Users. Ear & Hearing, 36, 653-663.

[3] Mahshie, J., Core, C., Rutkowski, R. (2011). Auditory and visual contributions to speech feature production in deaf children with cochlear implants. *Proc Int Congr Phon Sci*, *17*, 17–21.

[4] Bergeson, T. R., Pisoni, D. B., & Davis, R. A. O. O. (2005). Development of audiovisual comprehension skills in prelingually deaf children with cochlear implants. *Ear and Hearing*, *26*(2), 149–164.

[5] Schorr, E. A., Fox, N. A., Van Wassenhove, V., & Knudsen, E. I. (2005). Auditory-visual fusion in speech perception in children with cochlear implants. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(51), 18748–18750.

[6] Ling, D. (2002). *Speech and the Hearing-Impaired Child: Theory and Practice* (2nd ed.). Washington, DC: Alexander Graham Bell Association for the Deaf and Hard of Hearing.

[7] Eisenberg, L. S., Martinez, A. S., Boothroyd, A. (2003). Auditory-visual and auditory-only perception of phonetic contrasts in children. *Volta Rev*, *102*, 327–346.

[8] Kosky, C., & Boothroyd, A. (2003b). Validation of an on-line implementation of the Imitative test of Speech Pattern Contrast perception (IMSPAC). *J Am Acad Audiol*, *14*, 72–83.

[9] Boothroyd, A., Eran, O., Hanin, L. (1996). Speech perception and production in children with hearing impairment. In F. H. Bess, J. S. Gravel, A. M. Thorpe (Eds.), *Amplification for Children With Auditory Deficits* (pp. 55–74). Nashville, TN: Bill Wilkerson Center Press.

[10] Rose, Y., MacWhinney, B., Byrne, R., et al. (2006). Introducing Phon: A software solution for the study of phonological acquisition. In *Proceedings of the 30th Boston University Conference on Language Development* (pp. 489–500).

[11] Kishon-Rabin, L., Taitelbaum, R., Muchnik, C., et al. (2002). Development of speech perception and production in children with cochlear implants. *Ann Otol Rhinol Laryngol Suppl*, *189*, 85–90.

[12] McDaniel J, Camarata S, Yoder P. Comparing Auditory-Only and Audiovisual Word Learning for Children With Hearing Loss. J Deaf Stud Deaf Educ. 2018 Oct 1;23(4):382-398.