

Some Effects of the Prosodic Hierarchy on the Perception and Production of Spoken Language in Children Who Use Cochlear Implants

Jill Titterton[1,2], Alison Henry[2] and Joe Toner[1]

[1]Belfast City Hospital and [2]University of Ulster

1. Introduction

Many individuals with severe to profound sensorineural hearing loss can derive little benefit, in terms of speech perception, from conventional hearing aids. For this group, the use of a cochlear implant can facilitate perception of speech by electrically stimulating the auditory nerve. This technological development perhaps has most implications for children who are prelingually deafened, allowing access to speech perception and thus acquisition of spoken language. However, much research into this area shows that there is notable variation of spoken language ability between children who use cochlear implants (Svirsky et al, 2000). Whilst we know that certain factors such as age at onset of deafness, age of implantation, and duration of deafness, affect the development of spoken language skills in these children, it is not as yet clear what underlying processing or cognitive factors may be contributing to this variance (Pisoni, 2000). It is possible that the structure or processing within short term auditory sensory memory (Cowan, 1984) of children with cochlear implants (CI) is in some way different to that of children with normal hearing (NH), (Cleary et al, 1999). We wanted to investigate this individual variance (and what may underlie it) in more depth, and as a starting point noted that many CI children tend to omit weak syllables from their speech productions.

In early language development, NH children omit more weak than strong syllables (Fee, 1997). Furthermore, Gerken (1994 a and b, 1996), has shown that this tends to affect certain weak syllables over others within single and multiword utterances around 23-31 months of age in NHs, for example:

* The Freemasons of Ireland kindly provided grant support for this research. Thanks to staff at the Ulster and Belfast City Hospitals, Stranmillis Primary School and Nursery, and Little Rascals Playgroup, for facilitating data collection. Thanks also to all children and parents who participated in this project, and to the Belfast Cochlear Implant Team, Martin Kraemer (phonologist), Mike Stevenson (statistician), and David Watson (scientist) for all their time, support and help.

(1) a. /bə'na:ne/ → /'næ:ne/ →
'banana' → 'nana'

b. ((Tom)_{PW})_{PHP} (is (washing)_{PW})_{PHP} →
((Tom)_{PW})_{PHP}((washing)_{PW})_{PHP}

Gerken's (1994 a and b) Strong-(weak) Production Template Hypothesis posits an explanation for this pattern of weak syllable omissions by suggesting that the underlying prosodic hierarchy is influencing children's productions. It appears that at this early stage of language development, the optimum metrical structure children can produce is a strong-(weak) metrical foot. Furthermore, from the perspective of Optimality Theory, Gerken (1996) suggests that this strong-(weak) trochaic binary foot structure is optimised because the use of footed versus unfooted weak syllables within the prosodic hierarchy eases processing in STASM thus taking precedence over accurate target realisation for early language learners.

Our interest in this study then, was to investigate whether Gerken's account for weak syllable omissions in young NHs could also account for the weak syllable omissions of CI children, and to see whether (as in hearing children from Gerken et al, 1993) this was purely a production phenomenon for CI children. We were also interested in the interaction of memory load on weak syllable omissions. When looking at these aspects, we wanted to see how CI children performed in comparison to NH children matched by chronological and language age.

2. Experiment Methodology

2.1 Subjects

Three groups of children were used: a cochlear implant (CI) group, a chronological age match (AM) group, and a language match (LM) group. To date the CI and LM group contain 15 children, whilst the AM group contains 14. The CI and AM children were aged between 7 - 14 years, with a mean age of 9 years. The LM group (aged 3 - 6 years) was matched to the CI group by receptive vocabulary age as measured by the British Picture Vocabulary Scales II, Dunn et al (1997). For both groups of NHs hearing was screened at 20 dBHL (with 25 dBHL accepted for 250 Hz in test rooms without sound proofing) for 250, 500, 1000, 2000, 4000 and 6000 Hz. A screening questionnaire and school placements were used to ensure that NH children exhibited no evidence of developmental or speech and language delay.

The CI group were all prelingually deafened, with the aetiology for all but one of the children (who had suffered from meningitis at 7 months) being congenital. All the CI children had used their device consistently for at least three years, and the mean pure tone average for the group with device in use was 37 dBA. Educational Psychology, or teacher report ensured that the non-verbal performance of all the CI children was within normal limits for their

chronological age. All but one of the CI children (this boy being educated in a total communication environment) were educated in an oral-aural environment.

2.2. Procedures common to all three experiments

2.2.1. Recording of materials

All nonword stimuli (experiments 1 and 2), were carefully digitally recorded using a lapel mounted omnidirectional microphone ('realistic' electret tie pin, 33-1063) onto a Sony digital audio recorder (TCD-D7), by a standard Northern Irish speaker (female) at conversational speech levels (65-70 dBA) in a sound-attenuated, double-walled anechoic recording chamber using a sound level meter (Quest model 1700). As far as possible all syllable types in terms of stress (strong or weak), were produced equivalently (volume and duration). Pitch changes on stressed syllables were also carefully monitored so that they were appropriate for position within the nonword.

Sentences were digitally video recorded (Canon XL1) under similar conditions.

2.2.2. Material presentation and recording of responses

All stimuli were presented from a lap top Gateway Solo computer. Audio output from the computer was via a single loudspeaker (Yamaha Monitor speaker MS 101II) placed directly in front of the subject at a distance of 5 feet (1.50 cm). The presentation of all stimuli was set to approximately 70 dBA as measured by a sound level meter (Quest Model 1700) at the level of the child's ear.

All responses were recorded digitally through an omnidirectional microphone (ECM-MS907) using a Sony minidisc recorder (MZ-R700).

2.3. Experiment 1: Speech Memory Probe Test

This test of speech perception was designed to investigate how subjects performed on a simple same/different auditory only discrimination task using disyllabic nonwords which were manipulated to look at the effects of: footedness of weak syllables using strong-weak (trochaic feet) and weak-strong (unfooted weak syllables) stress patterns, length of memory trace for a given representation (using longer interstimulus intervals known to stretch short term auditory sensory memory capacity in young children, Ceponi  ne et al, 1999), and the effect of the position of onset consonants to be discriminated within the nonwords. Nonwords were used to avoid the possible interaction of morphosyntactic and semantic cues on phonological processing.

2.3.1 Stimulus materials

Children were asked to discriminate between a /p/ and /d/ consonant contrast placed in the onset position of a word initial or word medial syllable. All the phonemes used in the disyllabic nonwords were selected because analysis of the CI group's productions showed them to be produced correctly at least 60 % of the time. Accurate production of phonemes in hearing impaired children has been noted to always correlate positively with accurate perception of the same targets (Osberger, 1986). For the consonants to be discriminated between, we chose a contrast which would stretch discrimination ability but be easy for the CI children to perceive in simple tasks.

Two different stress patterns were used for the disyllabic nonwords. These were either trochaic (strong-weak) or iambic (weak-strong). After recording the stimuli they were then copied onto the Syntrillium Cool Edit programme using a sampling rate of 32 KHz with 16-bit amplitude quantization. The materials were subsequently analysed and (if necessary) adapted to ensure that indeed all the syllable types (strong or weak stress) were of equivalent volume and duration. This made sure that although the nonword pairs to be discriminated between exhibited the appropriate consonant contrast, the stress pattern cues of volume and duration remained the same.

Due to ethical limitations (relative to the amount of testing permissible for hearing controls), not all variations on this speech perception task could be investigated. The priorities for this study therefore, were to look at whether children received an advantage in discrimination if the target consonants were at the onset of the head of a trochaic foot versus a general advantage relative to headedness because the consonants to be discriminated were at the onset of the head of a monosyllabic strong foot, for example:

(2) Discrimination at the onset of the head of a trochaic foot: /pe'bm/

(3) Discrimination at the onset of the head of a monosyllabic strong foot within the nonword: /be'pim/

If this was shown to be the case, we then needed to look at whether the trochaic condition only provided an advantage because the consonant to be discriminated was also in word onset position, therefore we looked at discrimination of consonants at the onset of nonwords beginning with a weak syllable forming an iambic pattern, for example:

(4) Discrimination at the onset of an iambic nonword: /pe'lbim/

2.3.2. Procedure and scoring

Before completing the Speech Memory Probe tests children had to score at least 5/6 correct in a minimal pairs auditory only discrimination task where /p/ and /d/ were the contrasting consonants in single syllable real word pairs.

The Speech Memory Probe test consisted of an auditory only same/different discrimination task using the three conditions as noted in examples (2) to (4). Within each condition children were asked to discriminate between 12 nonword pairs delivered in a previously programmed random order under the control of the tester. Pictures of same or different cartoon characters were placed on either side of a roller ball mouse, and each child was asked to click the button on the mouse which corresponded to the same or different picture sequence depending on whether the two 'silly' words heard were the same or different. After each response the child received positive reinforcement. Before each testing set, subjects could have up to 9 practice items in order to learn how to do the same/different task.

The Memory Probe tests were scored in terms of the number of correct discriminations made in each test condition.

2.3.3. Results for Experiment 1: Speech Memory Probe Test

This experiment was conceptually quite difficult for the younger children, consequently, only 7 of the language match group were able to complete it. The chronological age matched group performed at ceiling level for all conditions in the Speech Memory Probe test.

2.3.3.1 Cochlear Implant Group and Language Match Group

Figures 1 and 2 show the three conditions of the Speech Memory Probe test for the CI and LM groups. The first error bar shows discrimination at the onset of the head of a trochaic foot, the second error bar shows discrimination at the word onset of an iambic nonword, and the third error bar shows discrimination at the onset of the head of a monosyllabic strong syllable within a nonword.

Figure 1 shows the results of the Speech Memory Probe tests for the CI group. Firstly, let us look at the first and third error bars, comparing discrimination ability at the head of a trochaic foot versus the head of a monosyllabic foot within a nonword. The CI group performed significantly better when discriminating the onset consonant at the head of a trochaic foot (Univariate ANOVA, $p < .001$).

How do we know that this better performance is not purely due to the target consonants being at the onset of the entire word? Looking now at the first and

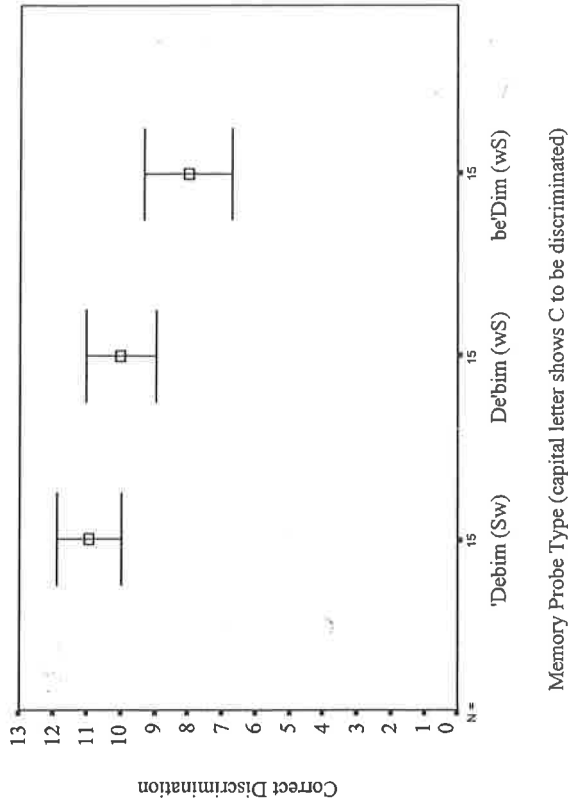


Figure 1. Discrimination scores for Memory Probe Tests (CI group)

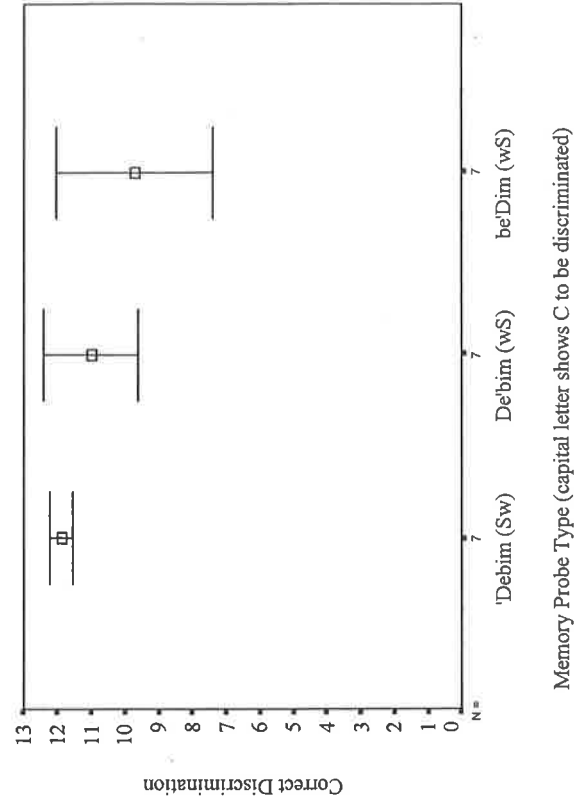


Figure 2. Discrimination scores for Memory Probe Tests (LM group)

second error bars in figure 1, we can compare discrimination ability where the target consonants are at the head of a trochaic foot in word onset position, and at the word onset of an iambic nonword (the onset of a weak unfooted syllable). Two-tailed Univariate ANOVA showed no significant difference between the two conditions ($p = .093$). However, if we look at our directional hypothesis which suggests that these CI children may perform better with a trochaic foot, and use a one-tailed Univariate ANOVA, a significant difference is found between the two conditions ($p = .047$). It seems that both a trochaic stress pattern and word onset position gives an advantage to the CI group for these Speech Memory Probe tests.

Figure 2 shows results for the three conditions in the LM group. As can be seen there is large variability (although this is least for the trochaic condition), probably because the task was difficult for these younger children, and because the group is so small. Whilst it is interesting to note that the trend seems to follow that of the CI group, a Univariate ANOVA showed no significant differences between any of the conditions ($p = .087$).

2.4. Experiment 2: Nonword Repetition Test

An auditory only Nonword Repetition test was devised to investigate the effect of increasing word length (and hence increasing memory load) on the processing of different stress patterns (particularly unfooted weak syllables) in short term auditory sensory memory. Again, nonwords were used to avoid the possible interaction of morphosyntactic and semantic cues on phonological processing.

2.4.1. Stimulus Materials

Ten 2, and twenty 4 syllable nonwords were created. In order to reduce the impact of speech production and discrimination limitations on the CI group, consonants were used which previous analysis of the CI groups' productions showed to be correct at least 60 % of the time (as has been said earlier, accurate production is a good estimate of perceptual skills). Furthermore, distinctive feature differences between consonants within nonwords were chosen so as to optimise discrimination. Short vowels were used so that they could equally take strong or weak stress, therefore not pre-empting results by using schwa vowels in weak syllable position.

At the disyllabic level, the nonwords were split into trochaic (strong-weak) or iambic (weak-strong) stress patterns, thus targeting both footed and unfooted weak syllables, for example:

(5) /'bɛwæp/ versus /bɛ'wæp/
S-(w) w-S

Whilst at the four syllable level, the nonwords were again split into pairs, giving each nonword the opportunity for the use of one footed and one unfooted weak syllable (see Gerken, 1994b re rationale for stress patterns), for example:

(6) /dɛmʌbæ'wɪp/ versus /dɛ'mʌbæ'wɪp/
S-(w) * S-(w) * S-(w) S-(w)

These recorded stimuli were subsequently analysed through Syntrillium's cool edit programme to ensure that difference in amplitude between unfooted and footed weak syllables would not be a factor influencing results. As expected, two-tailed paired samples t-tests showed no significant amplitude differences between the footed and unfooted weak syllables for the two syllable ($p = .76$), and four syllable non words ($p = .94$).

2.4.2. Procedure and scoring

Children were asked to repeat back the 'silly' words they heard through the loudspeaker. Before beginning the test, children had several opportunities to practice repeating nonwords of different lengths - the importance of repeating the target word as accurately as possible was stressed. Once the test commenced, the child was praised after repeating each nonword.

The Nonword Repetition tests were transcribed by a Speech and Language Therapist (both live and recorded responses) with specialist experience of the speech of hearing impaired individuals. These transcriptions were then validated by a trained Phonologist. Interestingly, many of the CI children in this study tended to produce nonwords which were often quite different both segmentally and prosodically to the targets. As a result it was decided to look at the results from the perspective of the mean number of syllable type produced considering: trochaic strong, monosyllabic strong, footed weak and unfooted weak, for the disyllabic and four syllable nonwords.

2.4.3. Results for Experiment 2: Nonword Repetition Test

At this stage the results for the CI group alone will be reported, as this data has not yet been analysed for the hearing control groups. The CI children realised the target stress patterns of the nonwords correctly significantly more in disyllabic than in four syllable nonwords (Wilcoxon Signed Ranks Test, $p < .001$).

In the disyllabic condition the CI group exhibited little difference between the use of footed and unfooted weak syllables. Figure 3 shows the syllable types used in four syllable nonwords. The trochaic strong, monosyllabic strong and footed weak syllables were all used similarly. However, with increasing syllable number, unfooted weak syllables were used significantly less than footed weak syllables (Wilcoxon Signed Ranks Test, $p = .001$).

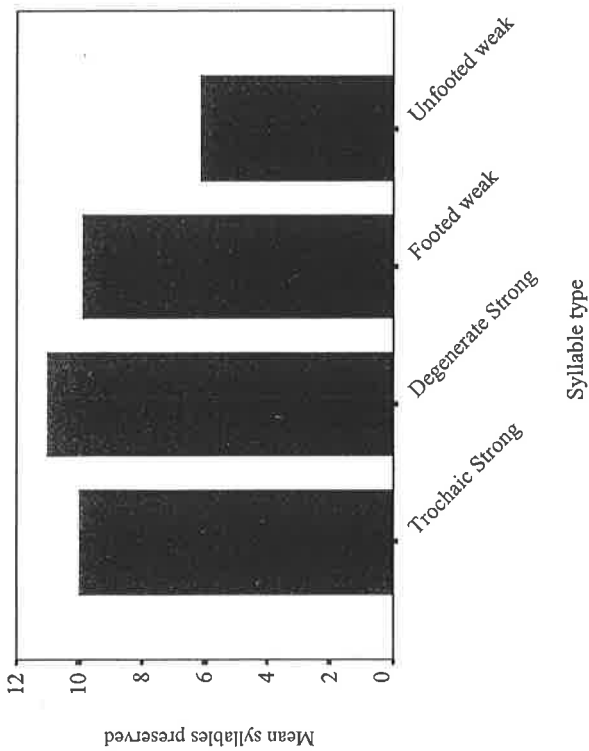


Figure 3. Syllable type used in four syllable nonwords (CI group)

2.5. Experiment 3: Sentence Imitation Test

This auditory plus visual test was developed to look at preservation of weak syllables (looking at footed versus unfooted weak syllables) in sentences where memory load was manipulated by changing sentence length and syntactic complexity.

2.5.1. Stimulus Materials

Sixteen sentences were developed for this sentence imitation task. The sentences used familiar vocabulary and were split into eight simple short (4 syllables in length) and eight more complex long (7 syllables in length) sentences. For both long and short sentence types, one of the weak syllables was footed.

All lexical words were given primary stress, whilst function morphemes were weakly stressed. Pitch changes were appropriate for Northern Irish speech. The prosodic structure for the target sentences is noted in the examples below, where the four standard sentence types (two short and two long) are shown (the final test contained four examples of each type):

(7) ((Tom)_{PW})_{PHP} (is (washing)_{PW})_{PHP}

(He (climbed)_{PW})_{PHP} (the (tree)_{PW})_{PHP}

((Tom)_{PW})_{PHP} (is (washing)_{PW})_{PHP} (the (fat)_{PW} (sheep)_{PW})_{PHP}

(He (climbed)_{PW})_{PHP} (the (stairs)_{PW})_{PHP} (in the (house)_{PW})_{PHP}

This prosodic structure was based on Selkirk's (1996) criteria, where nearly all the weak syllables in these sentences function as free clitics.

Once digitally video recorded, the materials were edited, making each sentence into a clip which was then transferred to a computer for MPEG1 encoding. These MPEG1 files were then inserted as individual slides in a power point presentation format.

2.5.2. Procedure and scoring

Children were asked to watch and listen carefully to the target sentences from the video, and repeat back exactly what they heard. Each child had three demonstration items before commencing. Once the test had begun, the child was positively reinforced after each imitation.

Preservation of weak syllables was looked at in terms of whether they were footed or unfooted in simple short, or long more complex sentences.

2.5.3. Results for Experiment 3: Sentence Imitation Test

As before, the AM group performed at ceiling levels.

2.5.3.1. Cochlear Implant Group and Language Match Group

Figures 4 and 5 show preservation of footed versus unfooted weak syllables in firstly the short, syntactically more simple sentences and then in the long, more complex sentences for the CI group and the LM group. A Univariate ANOVA showed that the CI group in general, preserved significantly fewer weak syllables, particularly in longer sentences when compared to the LM group ($p < .001$). When compared to the LM group, the CI group produced significantly more footed than unfooted syllables ($p = .011$), and their shorter sentences contained significantly more unfooted weak syllables than their longer sentences ($p = .005$).

4. Discussion and Conclusion

This project addressed three issues: do CI children process speech optimally through a trochaic binary foot, is this purely a production phenomenon, and how is this influenced by load in short term auditory sensory memory?

The results of the Speech Memory Probe test suggest that the CI group perform best perceptually, when processing trochaic feet. To a lesser extent, it was also clear that when target consonants were in the onset position of the nonwords, this facilitated discrimination. Interestingly, the LM group showed a similar trend to that of the CI group. Whilst these results suggest that binary trochaic feet may ease perceptual processing load in short term auditory sensory memory for CI children, further and more detailed investigation is required in this area.

The Nonword Repetition results showed that as memory load increased, the CI children began to use more footed than unfooted weak syllables. This suggests that at times the CI group can adequately process unfooted weak syllables in their productions, but only as long as memory isn't overloaded. This inconsistency of processing must in itself make language acquisition challenging. It also suggests that factors influencing processing in short term auditory sensory memory may have a major impact on the language abilities of these children. Titterton et al (2002), used an evoked response potential measure of short term auditory sensory memory (STASM) which suggested that whilst STASM capacity in CI children is similar to that of NHs, it is the encoding and processing of the signal that appears to be problematic (despite allowing for accurate discrimination). This seems to result in less efficient processing of language in STASM, hence forcing children to use facilitative strategies such as optimising weak syllables in trochaic feet.

The Sentence Imitation test results showed that the CI children tended to preserve more footed than unfooted weak syllables in sentences, particularly as sentence length increased. However, in the short sentences CI children still used more footed than unfooted weak syllables suggesting (unsurprisingly) that at the level of sentence processing, memory resources are beginning to optimise processing in trochaic feet even in short sentences.

In conclusion, when taken together these results support the possibility that CI children are influenced by the principles of the prosodic hierarchy, therefore optimising successful processing of weak syllables fitting into trochaic feet over those weak syllables which are unfooted. With these children, this phenomenon appears to be reflected in the processing of both perception and production of speech, and importantly, is strongly influenced by load in short term auditory sensory memory. This finding could have important diagnostic and treatment implications for children with CIs.

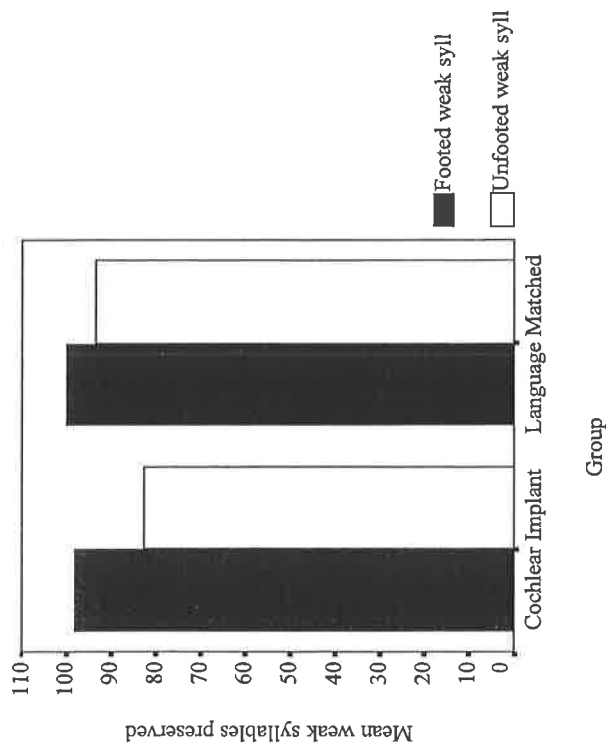


Figure 4. Preservation of footed and unfooted weak syllables in short sentences for the CI and LM groups

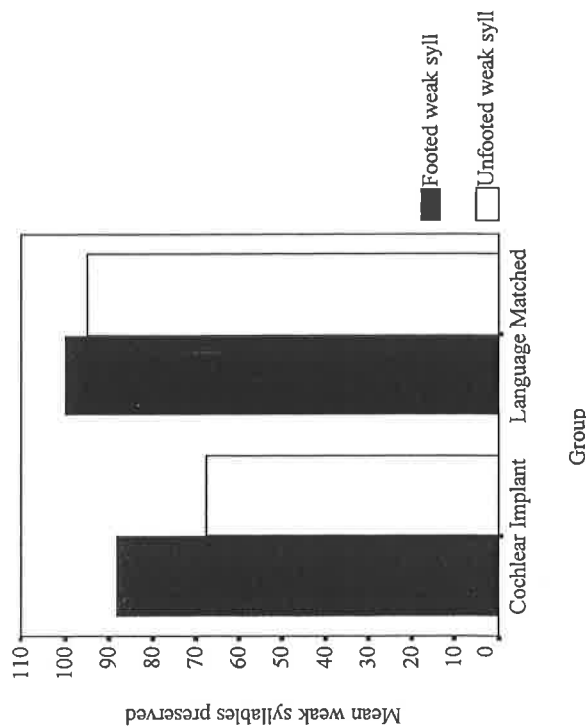


Figure 5. Preservation of footed and unfooted weak syllables in long sentences for the CI and LM groups

References

- Bishop, Dorothy, V.M. (1989). *Test for the reception of grammar 2nd edition*. Chapel Press: England.
- Ceponiene, Rita., Service, E., Kurjenluoma, S., Cheour, M., and Naatanen, R. (1999). Children's performance on pseudoword repetition depends on auditory trace quality: evidence from event-related potentials. *Developmental Psychology*, 35, 3, 709-720.
- Fee, E. Jane. (1997). The Prosodic Framework for Language Learning. *Topics in Language Disorders*, 17, 4, 53-63.
- Cleary, Miranda., Pisoni, D., and Geers, A.E. (1999). Some measures of verbal and spatial working memory in eight- and nine-year-old hearing-impaired children with cochlear implants. *Research on Spoken Language Processing Progress Report No. 23*. Bloomington, IN: Speech Research Laboratory, (pp. 52-79).
- Cowan, Noel., (1984). On short and long auditory stores. *Psychological Bulletin*, 96, 341-370.
- Dunn, Loyd, M., Dunn, L.M., Whetton, C., and Burley, J. (1997). *The British Picture Vocabulary Scale 2nd ed.* NFER-NELSON: Great Britain.
- Gerken, Lou Ann., and McIntosh, B.J. (1993). Interplay of function morphemes and prosody in early language. *Developmental Psychology*, 29, 3, 448-457.
- Gerken, Lou Ann. (1994a). Young children's representations of prosodic phonology: evidence from English-speakers' weak syllable productions. *Journal of Memory and Language*, 33, 19-38.
- Gerken, Lou Ann. (1994b). A metrical template account of children's weak syllable omissions from multisyllabic words. *Journal of Child Language*, 21, 565-584.
- Gerken, Lou Ann. (1996). Prosodic structure in young children's language production. *Language*, 72, 4, 683-712.
- Osberger, Mary Joe. (1986). Understanding speech intelligibility. *Topics in Language Disorders*, 6, 3, 47-59.
- Pisoni, David. (2000). Cognitive factors and cochlear implants: some thoughts on perception, learning, and memory in speech perception. *Ear and Hearing*, 21, 1, 70-78.
- Selkirk, Elizabeth. (1996). The prosodic structure of function words. *From signal to syntax: Bootstrapping from speech to grammar in early acquisition*, ed. by Jim Morgan and Katherine Demuth, 187-214. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Svirsky, Mario.A., Robbins, A.M., Kirk, K., Pisoni, D.B., and Miyamoto, R.T. (2000). Language development in profoundly deaf children with cochlear implants. *Psychological Science*, 11,2, 153-158.
- Titterton, Jill., Watson, D., Toner, J.G.T., and Henry, A. (2002). The use of the mismatch negativity (MMN) as an objective assessment of short-term auditory sensory memory function in cochlear implant children. Poster presented at the VIIIth International Cochlear Implant Conference, Manchester, England.
- Second-Language Sound Learning in Children and Adults: Learning Sounds, Words, or Both?**
- Pavel Trofimovich¹, Wendy Baker², James E. Flege³, Molly Mack¹
¹University of Illinois, ²Brigham Young University, ³University of Alabama
- 1. Introduction**
- What factors influence children's and adults' production of second-language sound segments (or "sounds" for short)? Previous research has identified at least two, among others. The first factor, "cross-language similarity," refers to how perceptually similar sounds are in a native (L1) and second (L2) language. That is, the degree of perceived dissimilarity (or similarity) between L1 and L2 sounds determines how L2 sounds are perceived and produced (Flege, 1995). For example, Japanese learners may produce the English /ɹ/ more accurately than the English /l/ (Flege, Takagi, & Mann, 1995) because they are more likely to perceptually differentiate the English /ɹ/, but not /l/, from the Japanese /r/ (Aoyama, Flege, Guion, Yamada, & Akahane-Yamada, 2002). Japanese learners may also produce the English /t/ more accurately than the English /θ/ because they are more likely to perceptually equate the English /t/, but not /θ/, with the similar Japanese /t/ (Guion, Flege, Akahane-Yamada, & Pruitt, 2000). L2 production thus depends on the perceived distance between L1 and L2 sounds. That is, depending on the particular relationship between individual L1 and L2 sounds, cross-language similarity can either help or hinder L2 production.
- The second factor that influences L2 production is related to the phonetic, syllabic, phonotactic, or prosodic "context" in which L2 sounds occur. That is, learners may have difficulty producing an L2 sound when it occurs in the context of certain sounds or in certain word- or phrase-stress conditions (Strange, Akahane-Yamada, Kubo, Trent, et al., 1998). It is known, for example, that English liquids differ in phonetic realization in word-initial and word-final position and, perhaps because of this, Japanese adults differ in the ability to produce /ɹ-/l/ distinctions as a function of word position (Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997). Both factors—perceived cross-language similarity and phonetic context—may be thought of as segmental or "sound-related" factors.

* This research was supported by University of Illinois at Urbana-Champaign Graduate College grants and NIH Grant 501 DC02892-05; CFDA#93.173.

