THE ACQUISITION OF ENGLISH SIBILANT FRICATIVES BY CHILDREN WITH BILATERAL COCHLEAR IMPLANTS

Patrick F. Reidy[a,b], *Mary E. Beckman*[b], *Ruth Y. Litovsky*[a], *& Jan Edwards*[a]

[a] Dept. of Communication Sciences & Disorders, University of Wisconsin–Madison, United States;
[b] Dept. of Linguistics, Ohio State University, United States

`reidy.16@osu.edu; beckman.2@osu.edu; litovsky@waisman.wisc.edu; jan.edwards@wisc.edu;`

ABSTRACT

The English sibilant fricatives /s/ and /ʃ/ are acquired late by normal hearing (NH) children, and pediatric cochlear implant (CI) users lag even further behind their NH peers. Previous work on the acquisition of sibilant fricatives by children with CIs has focused on their performance relative to NH controls, but the developmental trajectory of their sibilant fricative acquisition has not been investigated explicitly.

Productions of /s/ and /ʃ/ by children with bilateral CIs were analyzed with Dirichlet regression models in order to determine how their hearing experience and vocabulary development predict their accuracy and error patterns. Hearing age (i.e., total duration of CI use) best predicted the acquisitional trajectories. Neither age at implantation nor chronological age significantly improved the model fit, although receptive vocabulary score did.

Keywords: cochlear implants, sibilants, acquisition

1. INTRODUCTION

The production of /s/ or /ʃ/ requires a high degree of motor control for sibilant noise to be generated and sustained. A narrow constriction is formed between the tongue and the roof of the mouth, and as air moves through this constriction, its flow becomes turbulent, creating a turbulence noise source [11]. Simultaneously, the jaw rises, positioning the lower incisors downstream from the turbulent airflow [2], and another turbulence noise source is created when the flow of air collides with the incisors [6, 9]. To maintain the linguopalatal constriction, the tongue tip moves in relative opposition to the rising jaw [2]; thus, the turbulent airflow is sustained, rather than becoming occluded in the oral tract.

Given the gestural coordination required for a successful sibilant production, it is no surprise that the acquisition of English /s/ and /ʃ/ is protracted in children with normal hearing (NH). Smit and her colleagues [10] found that NH children reached 75% accuracy on /s/ between 3;6 and 6;0 and on /ʃ/ by 5;0 years;months. Furthermore, NH children have been found [8] to exhibit consonant-specific developmental patterns in their production of sibilants, suggesting that they learn different aspects of their articulation at different times.

Pediatric cochlear implant (CI) users have been found to lag behind their NH peers with respect to sibilant fricatives. Within a group of 181 eight- and nine-year-old pediatric CI users who had at least four years of experience with their prosthetic, Uchanski and Geers [13] found that only 49% of these children always produced /s/ and /ʃ/ as a fricative. Conversely, every child from a control group of chronological age-matched NH peers always produced these sibilants as a fricative. Separately, Todd and her colleagues [12] found that across a group of four- to nine-year-old CI children, productions were accurate, on average, 62% of the time for /s/ and 82.5% for /ʃ/. These accuracy rates were found to be similar to those of NH controls whose ages were matched to the CI children's duration of implant use, but were lower than those of age-matched NH peers.

While the findings in [12, 13] suggest that the acquisition of /s/ and /ʃ/ in pediatric CI users is more similar to NH children matched on hearing age, rather than chronological age, they do not address the developmental trajectory of this acquisition. The purpose of the current study is to determine how the accuracy and error patterns for target sibilant productions by children with bilateral CIs, vary with their hearing experience and language development.

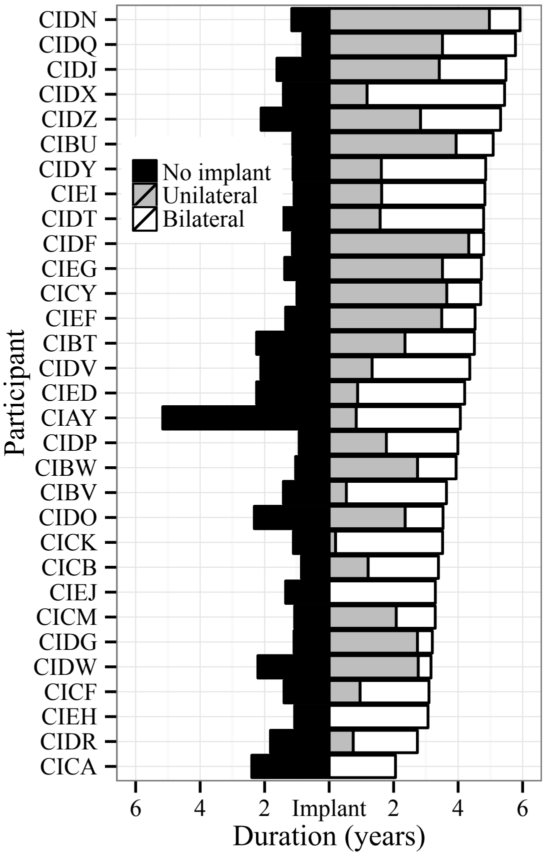
2. METHOD

2.1. Participants

Thirty-one pediatric bilateral CI users, who were born congenitally deaf, but who otherwise had no diagnoses for developmental disorders, participated in a picture-prompted word-repetition task. The participants were recruited from throughout the United States and tested at the University of Wisconsin–Madison. All spoke English as their first language.

Each child's hearing experience is summarized in Fig. 1. Age at implantation ranged from 0;9 to 5;1 (M = 1;6), while chronological age at test ranged from 4;1 to 9;2 (M = 5;8). Duration of unilateral CI use ranged from 0;0 to 4;11 (M = 2;0); and of bilateral CI use, from 0;4 to 4;3 (M = 2;1). Hearing age—i.e., the sum of the durations of unilateral and bilateral CI use—ranged from 2;0 to 5;11 (M = 4;2).

Figure 1: The implant history of each participant.



Age-standardized scores on the Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4) [1] ranged from 76 to 126 (mean = 97.97). Twenty-five of the participants had scores that were at least within one standard deviation (i.e., ≥ 85) of the normative mean for NH children of the same age.

2.2. Materials

The target words for the word-repetition task were 18 /s/- or /ʃ/-initial real English words that would be familiar to young children. The target sibilants occurred pre-vocalically, the vowels grouped into three classes, {i, ɑ, u} that represented the corner vowels. Specifically, {i} comprised /i, ɪ/; {ɑ}, /ɑ, ʌ, ɔ/; and {u}, /u, ʊ/. The list of sibilant-initial target words is given in Table 1. The word list also included stop-initial words, intermixed with the target words.

Table 1: The /s/- and /ʃ/-initial target words used in the word-repetition task.

	{i}	{ɑ}	{u}
/s/	seal seashore sister	sauce soccer sun	soup suitcase super
/ʃ/	sheep shield ship	shark shop shovel	chute shoe sugar

Three repetitions of each test word were spoken in a child-directed register by an adult female phonetician and recorded digitally at 22.5 kHz. From these recordings, three lists of auditory stimuli were assembled, and the order within each list was randomized. Finally, the auditory stimuli were paired with digital images of the target word, and these pairs were used as prompts in the repetition task.

2.3. Procedure

Testing took place in a sound-attenuated room. Prior to the task, the children were instructed that they would be shown images on a computer screen and played recordings of spoken words, and that they should repeat those words. Practice trials were completed before testing, so that the children would be familiar with the task. In some instances, a child's repetition overlapped the audio stimulus or was inaudible, in which case the prompt was replayed to elicit an audible, isolated production. The children's repetitions were recorded digitally at 44.1 kHz.

2.4. Transcription

A trained phonetician, who had no prior exposure to the speech of cochlear implant users, transcribed the initial sibilant of each target word using a custom Praat script, which allowed her to listen to each production and to visualize its waveform and spectrogram before transcribing it. For trials where more than one repetition of the target was elicited, the transcriber first chose which to transcribe. Here, the earliest audible attempt at the target was transcribed; a speech production error was not reason to skip an earlier repetition for a later, correct one. Next, the transcriber judged the consonant's type as either a 'sibilant fricative', a 'sibilant affricate', a 'non-sibilant fricative', a 'non-sibilant plosive', or 'other'. This last category was a catch-all for any production that was not easily classified into any of the other categories, and was used on only one trial. Lastly, if the consonant was judged to be a 'sibilant fricative', it was transcribed phonetically.

2.5. Error profiles

The consonant-type judgments and transcriptions were used to classify each target repetition as either a *phonetically correct* production or one of three errors that reflect the ways that sibilant noise may fail to be generated and sustained. First, *fortition errors* denote productions where a full occlusion is made in the oral tract, and as a consequence, the consonant has a plosive onset; 'sibilant affricates' and 'non-sibilant plosives' were both classified as fortition errors.

Next, *lenition errors* refer to productions where the constriction is made at a location in the oral tract such that the airflow exiting the constriction does not strike the incisors with enough force to generate a secondary noise source. This comprises both 'fronting' and 'backing' errors since a [θ] or a [x] substitution would count as a lenition error. 'Non-sibilant fricatives' were classified as lenition errors.

Lastly, *sibilant errors* denote productions where sibilant noise was generated, but where the constriction was not made at the target place of articulation. Again, this category cuts across the traditional fronting/backing dichotomy since [s] for /ʃ/ and [ʃ] for /s/ substitutions were both counted as sibilant errors. 'Sibilant fricatives' whose transcription did not match the target were counted as sibilant errors.

Each child's productions of /s/ and /ʃ/, respectively, were represented by a four-dimensional vector—referred to as an *error profile*—whose components denote the proportions of correct productions and fortition, lenition, and sibilant errors. Importantly, each of these vectors is compositional, in the sense that its components sum to one.

2.6. Dirichlet regression

Dirichlet regression (see [4]) assumes that the observed N-dimensional data are sampled from a Dirichlet distribution, $y \sim \text{Dirichlet}(\alpha)$, which assigns probability to compositional vectors y, such that $\mathbb{E}(y_i) = \alpha_i / \sum_n^N \alpha_n$. Each component of the parameter vector α is modeled as a function of the predictor variables $X^{(i)}$: $\log(\alpha_i) = X^{(i)}\beta^{(i)}$. That is, different estimates $\beta^{(i)}$ are computed for each dimension of the data, y_i.

When fitting a Dirichlet regression model to the error profiles for /s/ and /ʃ/, each parameter of a four-dimensional Dirichlet distribution was associated to a production category (i.e., correct or fortition, lenition, or sibilant error). Maximum likelihood estimates for the $\beta^{(n)}$ coefficients were computed using the DirichReg package [5] in R.

Models were built using a stepwise, forward selection protocol. The independent variables considered were chronological age, age at first implant, duration of unilateral and of bilateral CI use, hearing age, and (standardized PPVT-4) vocabulary score. At each step, nested models were compared using a likelihood-ratio test, and unnested models were compared in terms of their Cox & Snell pseudo-R^2.

3. RESULTS

The base model included only a main effect of consonant. The model fit was not improved by adding bilateral CI use ($\chi^2(4) = 0.8093$, $p > 0.93$), but was significantly improved by a main effect of any other variable, among which hearing age had the greatest pseudo-R^2 (hearing age: 0.7196, chronological age: 0.6124, unilateral CI use: 0.4981, age at implantation: 0.4133, vocabulary score: 0.3851). The interaction between consonant place and hearing age did not improve the model ($\chi^2(4) = 3.4116$, $p > 0.49$).

After adding a main effect of hearing age, the model was not improved by the addition of chronological age ($\chi^2(4) = 4.2163$, $p > 0.37$), of unilateral CI use ($\chi^2(4) = 0.6657$, $p > 0.95$), or of age at implantation ($\chi^2(4) = 4.2084$, $p > 0.37$). A main effect of vocabulary score significantly improved the model fit ($\chi^2(4) = 13.829$, $p < 0.01$), but no binary nor ternary interaction involving vocabulary score offered any further improvement ($p > 0.10$).

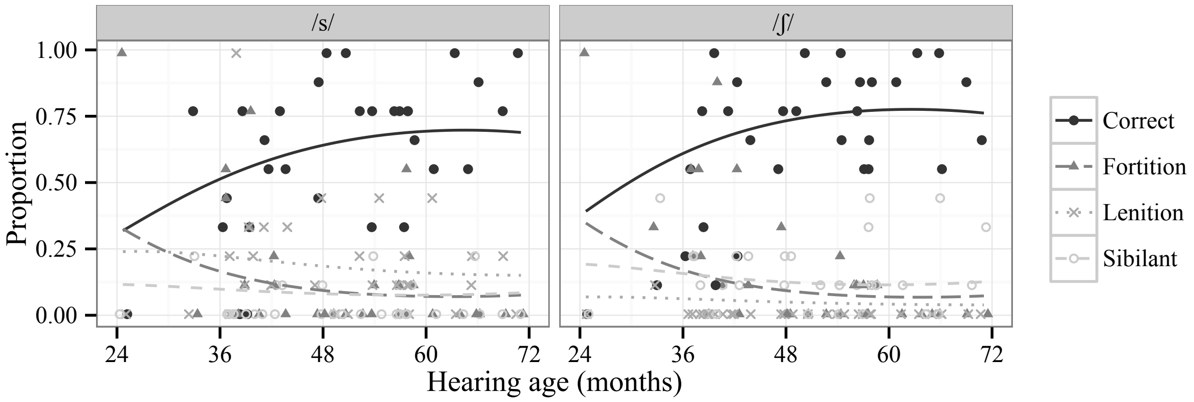
Lastly, quadratic hearing age significantly improved the model fit ($\chi^2(4) = 10.551$, $p < 0.05$), but quadratic vocabulary score did not ($\chi^2(4) = 1.2643$, $p > 0.86$). The final Dirichlet regression model had the following structure:

(1) $\text{Consonant} + \text{HearAge} + \text{HearAge}^2 + \text{Vocab}.$

Correct productions were significantly more common for /ʃ/ ($\hat{\beta} = 0.6337$, $SE = 0.2723$, $p < 0.05$). The coefficients for linear ($\hat{\beta} = 6.1123$, $SE = 1.0811$, $p < 0.001$) and quadratic ($\hat{\beta} = -2.6559$, $SE = 1.1059$, $p < 0.05$) hearing age were both significantly different from zero; the combination of a positive linear term and a negative quadratic term suggests that the rate of increase in accuracy slows with prolonged CI use, as is shown in Fig. 2. Sibilant accuracy was also found to increase significantly with expanding vocabulary size ($\hat{\beta} = 3.3749$, $SE = 1.0871$, $p < 0.01$).

Fortition errors decreased significantly with increasing hearing age ($\hat{\beta} = -3.1547$, $SE = 1.1415$, $p < 0.01$). Furthermore, Fig. 2 suggests that this decrease was greatest between the second and fourth

Figure 2: The marginal predicted values of the fitted model plotted against hearing age.



year post-implantation, after which only small reductions are seen.

Lenition error proportions were significantly lower for /ʃ/ ($\hat{\beta} = -0.8222$, $SE = 0.2572$, $p < 0.01$). Conversely, sibilant errors were more common for /ʃ/ ($\hat{\beta} = 0.9287$, $SE = 0.2562$, $p < 0.001$). While the coefficients for linear and quadratic hearing age were neither significantly different from zero for either error type, Fig. 2 suggests that these errors decrease as hearing age increases.

4. DISCUSSION

Hearing age best predicted the error profiles for sibilant fricatives. Furthermore, none of the other hearing experience variables improved model fit after a main effect of hearing age was added. In particular, it was surprising not to find a significant effect of age at implantation, given that previous research [7] has found that CI children's language outcomes improve with younger implantation ages, when controlling for hearing age. It is possible that age at implantation had no effect because these participants exhibited a narrow range of ages: excluding child CIAY, implantation age only ranges from 0;9 to 2;4.

The significant, positive effect of vocabulary score on the proportion of correct productions suggests that larger vocabularies are associated with greater production accuracy for sibilants; however, the causal mechanisms underpinning this association may be bidirectional. On the one hand, knowledge of a phonological category comprises generalizations over the motor patterns that must be executed during speech production, as well as over the representations of lexical items that are themselves abstracted from somatosensory patterns. Thus, children with larger vocabularies will have a broader base of examples from which to learn the articulatory schemes for a phonological category. On the other hand, the phonological encoding of a newly encountered word is facilitated by the presence of robust perceptual categories [3]. Thus, to the extent that production accuracy indicates the development of perceptual categories for the same phone, children with higher accuracy rates would be expected to have larger vocabularies.

The fitted model included a main effect of consonant, but no interaction between consonant and hearing age; thus, accuracy rates were predicted to be greater for /ʃ/ than for /s/ across the entire hearing age range. This differs from the trend that has been observed [8, 10] in NH children, who are more accurate on /s/ between two and four years of age, but who then become more accurate on /ʃ/.

The developmental perspective of the Dirichlet regression also illuminated CI children's error patterns in new ways. Todd and her colleagues [12] reported that the most common errors for /s/ were [f], [θ], and stop substitutions, and for /ʃ/ were [s] and [tʃ] substitutions, which in the current typology, correspond to lenition and fortition errors for /s/ and to sibilant and fortition errors for /ʃ/. The fitted Dirichlet regression model reveals that these errors do not occur uniformly across hearing ages, but rather that for /s/ and /ʃ/ the fortition errors occur primarily in CI children of younger hearing ages, while lenition or sibilant errors, respectively, occur more frequently than fortition errors after the fourth year post-implantation.

5. ACKNOWLEDGMENTS

Data collection for this project was supported by NIH Grant No. R01DC02932 to Jan Edwards and NIH Grant No. R01DC008365 to Ruth Y. Litovsky.

6. REFERENCES

[1] Dunn, L. M., Dunn, D. M. 2007. *Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4)*. Minneapolis, MN: Pearson Assessments.

[2] Iskarous, K., Shadle, C. H., Proctor, M. I. 2011. Articulatory–acoustic kinematics: The production of American English /s/. *J. Acoust. Soc. Am.* 129(2), 944–954.

[3] Keren-Portnoy, T., Vihman, M. M., DePaolis, R. A., Whitaker, C. J., Williams, N. M. 2010. The role of vocal practice in constructing phonological working memory. *J. Speech Lang. Hear. R.* 53, 1280–1293.

[4] Maier, M. J. 2014. *DirichletReg: Dirichlet regression for compositional data in R*. Institute for Statistics and Mathematics, Vienna University of Economics and Business. Research Report Series, Report 125.

[5] Maier, M. J. 2014. *DirichletReg: Dirichlet regression in R*. R package version 0.5–2.

[6] Naryanan, S. S., Alwan, A. A. 2000. Noise source models for fricative consonants. *IEEE Trans. Speech Audio Proc.* 8(2), 328–344.

[7] Nicholas, J. G., Geers, A. E. 2007. Will they catch up? The role of age at cochlear implantation in the spoken language development of children with severe to profound hearing loss. *J. Speech Lang. Hear. R.* 50, 1048–1062.

[8] Reidy, P. F. 2014. Moving targets and unsteady states: "Shifting" productions of sibilant fricatives by young children. *J. Acoust. Soc. Am.* 136(4), 2262.

[9] Shadle, C. H. 1991. The effect of geometry on source mechanisms of fricative consonants. *J. Phonetics* 19, 409–424.

[10] Smit, A. B., Hand, L., J., F. J., Bernthal, J. E., Bird, A. 1990. The Iowa Articulation Noems Preject and its Nebraska replication. *J. Speech Hear. Disord.* 55, 779–798.

[11] Stevens, K. N. 1998. *Acoustic Phonetics*. Cambridge, MA: MIT Press.

[12] Todd, A. E., Edwards, J. R., Litovsky, R. Y. 2011. Production of contrast between sibilant fricatives by children with cochlear implants. *J. Acoust. Soc. Am.* 130(6), 3969–3979.

[13] Uchanski, R. M., Geers, A. M. 2003. Acoustic characteristics of the speech of young cochlear implant users: A comparison with normal-hearing age-mates. *Ear Hearing* 24(1S), 90S–105S.