The Acoustics of Word-Initial Fricatives and Their Effect on Word-Level Intelligibility in Children With Bilateral Cochlear Implants

Patrick F. Reidy,[1,2] Kayla Kristensen,[1,2] Matthew B. Winn,[1,3,4] Ruth Y. Litovsky,[1,2,3] Jan R. Edwards[1,2]

Objectives: Previous research has found that relative to their peers with normal hearing (NH), children with cochlear implants (CIs) produce the sibilant fricatives /s/ and /ʃ/ less accurately and with less subphonemic acoustic contrast. The present study sought to further investigate these differences across groups in two ways. First, subphonemic acoustic properties were investigated in terms of dynamic acoustic features that indexed more than just the contrast between /s/ and /ʃ/. Second, the authors investigated whether such differences in subphonemic acoustic contrast between sibilant fricatives affected the intelligibility of sibilant-initial single word productions by children with CIs and their peers with NH.

Design: In experiment 1, productions of /s/ and /ʃ/ in word-initial prevocalic contexts were elicited from 22 children with bilateral CIs (aged 4 to 7 years) who had at least 2 years of CI experience and from 22 chronological age-matched peers with NH. Acoustic features were measured from 17 points across the fricatives: peak frequency was measured to index the place of articulation contrast; spectral variance and amplitude drop were measured to index the degree of sibilance. These acoustic trajectories were fitted with growth-curve models to analyze time-varying spectral change. In experiment 2, phonemically accurate word productions that were elicited in experiment 1 were embedded within four-talker babble and played to 80 adult listeners with NH. Listeners were asked to repeat the words, and their accuracy rate was used as a measure of the intelligibility of the word productions. Regression analyses were run to test which acoustic properties measured in experiment 1 predicted the intelligibility scores from experiment 2.

Results: The peak frequency trajectories indicated that the children with CIs produced less acoustic contrast between /s/ and /ʃ/. Group differences were observed in terms of the dynamic aspects (i.e., the trajectory shapes) of the acoustic properties. In the productions by children with CIs, the peak frequency and the amplitude drop trajectories were shallower, and the spectral variance trajectories were more asymmetric, exhibiting greater increases in variance (i.e., reduced sibilance) near the fricative–vowel boundary. The listeners' responses to the word productions indicated that when produced by children with CIs, /ʃ/-initial words were significantly more intelligible than /s/-initial words. However, when produced by children with NH, /s/-initial words and /ʃ/-initial words were equally intelligible. Intelligibility was partially predicted from the acoustic properties (Cox & Snell pseudo-R^2 > 0.190), and the significant predictors were predominantly dynamic, rather than static, ones.

Conclusions: Productions from children with CIs differed from those produced by age-matched NH controls in terms of their subphonemic acoustic properties. The intelligibility of sibilant-initial single-word productions by children with CIs is sensitive to the place of articulation of the initial consonant (/ʃ/-initial words were more intelligible than /s/-initial words), but productions by children with NH were equally intelligible across both places of articulation. Therefore, children with CIs still exhibit differential production abilities for sibilant fricatives at an age when their NH peers do not.

Key words: Cochlear implants, Intelligibility, Sibilant fricatives, Speech production

(Ear & Hearing 2016;XX;00–00)

INTRODUCTION

Since cochlear implants (CIs) were approved for use in children nearly 30 years ago, the speech and language outcomes for prelingually deaf children have improved dramatically. In virtually every aspect of speech and language, prelingually deaf children who receive a CI perform better than children with similar levels of hearing loss who use a hearing aid (Osberger et al. 1993; Spencer et al. 1998, 1999; Tomblin et al. 1999; Geers et al. 2003). Despite this marked improvement in speech and language outcomes, children with CIs often continue to perform at a lower level than their peers with normal hearing (NH) on a number of speech and language measures, including vocabulary size, speech perception and production, and phonological awareness (Spencer et al. 2004; James et al. 2005; Connor et al. 2006; Nittrouer et al. 2012).

This article focuses on the speech production skills of children with CIs and those with NH, which are important for several reasons. First, school-age children need to be able to communicate effectively to their caregivers, teachers, and peers, and even small deficits in the intelligibility of a child's speech in quiet will be compounded in a noisy setting, such as a classroom. Second, speech production ability has been linked to phonological awareness (Bird et al. 1995), which is a strong predictor of literacy outcomes in children with NH (e.g., Lonigan et al. 2000; Melby-Lervåg et al. 2012) and in children with hearing loss (Cupples et al. 2014; Webb & Lederberg 2014).

Children with CIs produce speech sounds, especially consonants, less accurately than their peers with NH. Connor et al. (2006) assessed consonant production accuracy with two standardized articulation tests, the Arizona Articulation Proficiency Scale (Fuddle 1974) and the Goldman-Fristoe Test of Articulation (Goldman & Fristoe 2000). They found that children with CIs who were implanted between age 1 and 2.5 years were less accurate than their NH peers at age 7. Similar findings have been observed in other studies with smaller sample sizes (e.g., Warner-Czyz & Davis 2008; Ertmer & Goffman 2011).

The degraded nature of the CI signal is an important factor leading to poorer speech production skills in children with CIs. First, compared to the frequency tuning of a healthy auditory system, the frequency tuning of electrically stimulated nerve

cells is much broader (Raggio & Schreiner 2003; Middlebrooks et al. 2005); hence, frequency resolution is poorer with electrical hearing than with acoustic hearing. Spectral contrasts, such as place of articulation (e.g., /s/ vs. /ʃ/), relative to voicing (e.g., /t/ vs. /d/) and manner contrasts (e.g., /t/ vs. /s/) are particularly impacted by this poor frequency resolution (Friesen et al. 2001; Iverson 2003; Munson et al. 2003). Another deficit of the CI signal is that the range of audible frequencies is generally more restricted for CI users because the processor of commercially available CIs extends up to only 8 kHz (Loizou 2006). This cutoff frequency may complicate CI users' perception of anterior consonants, such as the sibilant fricative /s/, which typically have energy concentrated in the 5–10 kHz range.

This article focuses on one particular spectrally cued consonant contrast between the voiceless sibilant fricatives in English: /s/ and /ʃ/. Based on the poorer frequency resolution and reduced frequency bandwidth of the CI signal alone, this should be a challenging contrast for children with CIs to acquire. The objective of the present study was to compare the /s/ versus /ʃ/ contrast in children with CIs to a group of age-matched controls with NH. This comparison was made at two levels: first, in terms of the subphonemic acoustic properties of the children's productions of the fricatives, and second, in terms of the intelligibility of the children's productions of sibilant-initial words, when these words were presented in background noise to naïve adult listeners.

EXPERIMENT 1

Motivation

The sibilants /s/ and /ʃ/ are acquired over an extended period of time in children with NH (Nittrouer et al. 1989; Nittrouer 1995). In a large-scale cross-sectional study of children with NH, Smit et al. (1990) found that phonetically trained clinicians judged as accurate 70% of target /s/ and 77% of target /ʃ/ tokens produced by 4-year-old children. For the productions by 7-year-old children, 86% of target /s/ and 94% of target /ʃ/ tokens were judged accurate. Complete mastery of these consonants may, thus, not occur until preadolescence for some speakers. Furthermore, during acquisition, covert contrast (a systematically produced subphonemic contrast that is not reliably perceived by a trained phonetician) is observed in the [s] and [ʃ] substitution patterns for at least some children with NH between the ages of 2 and 5 years (Li et al. 2009), indicating that subphonemic acoustic contrast varies across young children in gradient ways that may not be detected by narrow transcription. Finally, deficits in a child's perception of /s/ and /ʃ/ have been linked to delays in the accurate production of these consonants (Rvachew & Jamieson 1989; Rvachew et al. 2004). The extended period of refinement, the existence of covert contrast, and the perception–production link suggest the need for auditory feedback and self-monitoring to acquire these fricatives. A reliance on the auditory–verbal feedback loop frames /s/ and /ʃ/ as interesting objects of investigation in children with CIs because this population has delayed access to auditory input and a degraded and band-limited auditory signal.

Acquisition of /s/ and /ʃ/ is more protracted in children with CIs, relative to children with NH. In a longitudinal study that tracked phonological development in children with CIs from 6 months preimplant to 6 years postimplant, Serry and Blamey (1999) found that /ʃ/ was produced accurately at least 50% of the time by 48 months post implant. However, /s/ was not produced with comparable accuracy even after 72 months of implant use (Blamey et al. 2001). Chin (2003) found, in a cohort of children who had at least 5 years of implant experience, that /ʃ/ was produced accurately at least 75% of the time, but /s/ was not. Thus, when comparing the duration of the CI users' experience to the age of the children with NH (cf. Smit et al. 1990), the acquisition of /ʃ/ by the CI users appears to lag that by NH children by approximately 1 year, but /s/ is even more delayed.

Acoustic analyses have indicated that, relative to their NH peers, children with CIs produce /s/ and /ʃ/ with atypical acoustic properties and with diminished acoustic contrast. Uchanski and Geers (2003) recorded productions of word-initial target /s/ and /ʃ/ from 8- and 9-year-old children with CIs with at least 4 years of experience with their device and from a cohort of NH controls of the same age. One hundred percent of the NH children produced all target /s/ and /ʃ/ tokens as fricatives, but only 49% of the children with CIs did so. When aggregated across children with CIs, 82% of target /s/ and /ʃ/ tokens were produced as fricatives; hence, while individual talkers with CIs produced manner errors that were atypical relative to their NH peers, as a group, children with CIs exhibited high accuracy in manner of articulation. Centroid frequency was estimated from each target production whose manner of articulation was a fricative. From each NH child's productions, the mean centroid was computed separately for /s/ and /ʃ/, as was the difference between these two means as an acoustic measure of the /s/ versus /ʃ/ contrast. These values were used to demarcate the normal limits of the spectral properties of the sibilant categories and the sibilant contrast. Of the children with CIs, 63% produced /s/ within normal limits, 86% produced /ʃ/ within normal limits, and 71% produced the /s/ versus /ʃ/ contrast within normal limits.

One limitation of Uchanski and Geers's (2003) analysis was that to be included in their acoustic analysis, a production needed only to have a fricative manner of articulation; hence, place errors, such as an [s] for target /ʃ/ substitution, would have been included. Since the children with CIs were less accurate than their NH peers at producing the target sibilants, it is possible that the reduced acoustic contrast was because of more productions with place errors being included for the children with CIs. To control for this potential confound, Todd et al. (2011) recorded productions of target /s/ and /ʃ/ in word-initial position from children with bilateral CIs between the ages of 4 and 7 years, who had at least 2.5 years of experience with their first device, and from their chronological age–matched peers. Productions of /s/ and /ʃ/ were included in the acoustic analysis only if they were judged by a trained phonetician to be phonemically accurate productions of the target sound. Peak frequency was computed from the middle 40 ms of each production, and a linear mixed-effects model revealed a significant interaction between group and consonant: the acoustic contrast between /s/ and /ʃ/ was smaller in the productions by the children with CIs, even when only correct productions were analyzed.

One limitation of Uchanski and Geers (2003) and Todd et al. (2011) is that both studies investigated the acoustics of /s/ and /ʃ/ in terms of a single acoustic feature that indexed the difference in place of articulation. Because /s/ is produced with a more anterior place of articulation than /ʃ/, the front cavity is smaller and has higher resonant frequencies; therefore, energy is concentrated at higher frequencies in the spectrum (Hughes & Halle 1956; Forrest et al. 1988). The frequency

location of this energy concentration may be reliably indexed with either centroid frequency (e.g., Uchanski & Geers 2003) or peak frequency (e.g., Todd et al. 2011)—two measures that may be thought of, respectively, as the mean or mode frequency at which energy is concentrated. When /s/ and /ʃ/ are the only fricatives analyzed, the contrast between them can be captured with a single acoustic measure of energy concentration, such as centroid or peak frequency (Jongman et al. 2000); however, these measures fail to capture other aspects of /s/ and /ʃ/, such as their degree of sibilance. Sibilant frication is generated when the flow of air collides with a downstream obstacle, such as the incisors, producing noise whose spectrum is characterized by a high-frequency peak (Narayanan & Alwan 2000). By contrast, nonsibilant frication (e.g., /f, θ/) is generated without such an obstacle noise source, and the resulting spectrum is more diffuse in shape, often lacking a prominent high-frequency peak (Shadle 1985, 1990). Two measures have been reported as indices of degree of sibilance. The first is the variance of the spectrum, which is computed by normalizing its amplitude values, so that they sum to one, and then computing the variance of this normalized spectrum as if it were a discrete probability mass function (Forrest et al. 1988; Jongman et al. 2000). The second measure is the difference in decibels between the higher-frequency maximum amplitude and lower-frequency minimum amplitude (Koenig et al. 2013).

A second limitation of the methods in Uchanski and Geers (2003) and Todd et al. (2011) is that centroid and peak frequency were measured at only one point in the fricative; however, the spectral properties of /s/ and /ʃ/ are known to vary temporally in both adults' and children's productions (Soli 1981; Munson 2004; Iskarous et al. 2011; Koenig et al. 2013). Furthermore, /s/ and /ʃ/ contrast in terms of how peak frequency varies temporally during the fricative. Reidy (2015) analyzed peak frequency trajectories from adults' and children's productions of /s/ and /ʃ/. In adults, /s/ and /ʃ/ differed in terms of trajectory level and shape—that is, in terms of both static (level) and dynamic (shape) aspects of peak frequency. The children developed toward this adult-like contrast. The youngest children differentiated /s/ and /ʃ/ only in terms of trajectory level; older children differentiated the sibilants in both level and shape, but to a lesser extent than the adults. While there is no published work, to our knowledge, on the spectral dynamics of productions of target sibilants by children with CIs, it has been observed that children with CIs speak more slowly than their NH age peers (Burkholder & Pisoni 2003), which suggests that there may be differences in spectral dynamics across the two groups.

Experiment 1 investigates the acoustics of phonemically correct productions of /s/ and /ʃ/ by children with CIs and children with NH, who were aged between 4 and 7 years. Children within this age group were targeted because according to the normative data from Smit et al. (1990), they would be expected to produce a majority (at least 75%) of the target sibilant tokens accurately. At the same time, because the children are still developing toward adult-like proficiency, it is likely that acoustic analysis will reveal greater detail of these children's incipient contrastive categories. In particular, experiment 1 extends previous work on the acoustics of productions of /s/ and /ʃ/ by children with CIs and children with NH in three respects. First, in addition to computing peak frequency to index the contrast between /s/ and /ʃ/, we also computed spectral variance and amplitude drop to index the degree of sibilance of the children's productions.

Second, these three acoustic measures were computed from psychoacoustic spectra, which were computed by applying a filter bank model of the auditory system to the Hertz spectrum. Third, we examined the dynamics of the three measures over the duration of the fricative noise instead of focusing on a static single-point representation of them. The subphonemic psychoacoustic properties were measured from phonemically correct productions of /s/ and /ʃ/ by children with CIs and children with NH, and the temporal variation in these properties across the fricatives was analyzed with polynomial growth curve models. The experiment was conducted with approval from the Institutional Review Board at the University of Wisconsin–Madison.

MATERIALS AND METHODS

Participants

Twenty-two congenitally deaf children with bilateral CIs participated in Experiment 1. All children were between the ages of 4;1 and 7;8 (years; months). These participants were recruited throughout the United States and tested at the University of Wisconsin–Madison. The speech and language skills of some of these children have been previously reported in Hess et al. (2014) and Todd et al. (2011), and the speech perception skills of some of these children have been reported in Misurelli and Litovsky (2012, 2015). Before testing, parents were asked about their child's speech, language, and developmental history. Children with CIs were not included in the study if they had any developmental problems other than hearing loss. All children had received their first CI by 30 months and had at least 2 years of experience with at least one implant, but otherwise exhibited varying amounts of hearing experience: age at implantation ranged from 10 to 28 months (M = 15.95, SD = 4.79); duration of unilateral hearing experience ranged from 0 to 51 months (M = 23.28, SD = 16.07); duration of bilateral hearing experience ranged from 5 to 39 months (M = 23.49, SD = 11.18); and hearing age (time with at least one device) ranged from 24 to 77 months (M = 46.77, SD = 12.14). Device manufacturer was not controlled across the children with CIs.

Twenty-two children with NH also participated in the study. They were matched to the participants with CIs by age and sex. The children with NH were recruited from schools and day care centers in Columbus, Ohio. All children with NH were typically developing native English speakers based on parent report and passed a hearing screening which consisted of either otoacoustic emissions within normal range at 2000, 3000, 4000, and 5000 Hz or pure-tone audiometry thresholds within normal limits at 500, 1000, 2000, and 4000 Hz. Table 1 provides demographic information for both groups of participants. A Wilcoxon signed-rank test revealed no significant difference (T = 163, p = 0.25) between the mean ages of the two groups (CI: M = 62.72, SD = 11.31 months; NH: M = 61.89, SD = 10.22 months). Furthermore, a Kolmogorov–Smirnov test revealed no significant difference in the age distribution of the two groups (D_{22} = 0.27, p = 0.39). To assess their receptive vocabulary skills, the children with CIs were administered the Peabody Picture Vocabulary Test—Fourth Edition (Dunn & Dunn 2007) and the children with NH were administered the Receptive One Word Picture Vocabulary Test—Second Edition (Brownell 2000). The two groups of children were administered different instruments because they were tested at different

TABLE 1. Demographic characteristics of the two groups of children who participated in Experiment 1

Group	Number of Speakers	Mean Chronological Age in Years;Months (SD, Range)	Mean Hearing Age* (SD, Range)	Number of Male:Female Speakers	Average Receptive Vocabulary Standard Score (SD, Range)†
CI	22	5;2 (0;11, 4;1—7;8)	3;10 (1;0, 2;0–6;5)	11:11	100 (16) 76—126
NH	22	5;1 (0;10, 4;0–7;9)	—	11:11	112 (13) 85—145

CI indicates cochlear implant; NH, normal hearing.
*Hearing age was calculated by subtracting age at activation of the first CI from chronological age at time of test.
†Receptive vocabulary was assessed by the Peabody Picture Vocabulary Test (PPVT-4, Dunn & Dunn 2007) for the children with CIs and by the Receptive One Word Picture Vocabulary Test (ROWPVT-2, Brownell 2000) for the children with NH.

sites. The groups' age-standardized receptive vocabulary scores are reported in Table 1. All children scored no less than two standard deviations below the normative mean standard score (M = 100, SD = 15 for both instruments). Because different instruments were administered, the standard scores were not compared across the two groups.

Stimuli

The stimuli were monosyllabic and bisyllabic words that could be represented by pictures. There were 15 /s/-initial words and 15 /ʃ/-initial words. The children with NH were participating in a larger study on the acquisition of consonants, and the word list for these children included 78 other words that began with other lingual obstruents (/t/, /k/, /d/, /g/, /tʃ/, or /θ/). The word list for the children with CIs was a subset of the words presented to the children with NH. There were 9 words for each fricative—those words in a high front, low central, or high back vowel context, as listed in Table 2—and 45 other words that began with /t/, /k/, /d/, or /g/.

A young adult female speaker of Mainstream American English produced multiple repetitions of each word in a child-directed register. These productions were recorded digitally at 22.5 kHz. Three repetitions of each word were chosen to combine with other words to create six lists of auditory stimuli (i.e., two ordered lists for each of the three sets of audio recordings of the words). The order within each list was pseudorandomized, so that the words for each target sibilant–vowel pair were distributed evenly across the list. The words were normalized for amplitude within each list. Each word was paired with a color digital image that depicted its referent, and these audiovisual pairs were used as stimuli in the repetition task.

Procedure

The children completed a picture-prompted auditory word repetition task in a quiet room. A prompted repetition task was chosen, as opposed to an unprompted naming task, because in the larger cross-sectional study of children with NH, the same task was used to elicit productions from very young children

(2 and 3 years old) and because we wanted to minimize task demands as much as possible. Children completed the task seated at a table in front of a computer screen, loudspeakers (such as Audix PH5-VS), and an AKG C59000M microphone (cardioid response). Before beginning the task, the children were instructed that they would see pictures on the computer screen and would hear words through the loudspeakers and that it was their job to repeat those words into the microphone. The task was implemented with a custom program, written in the Tcl/Tk programming language (www.tcl.tk), that presented the digital images on the computer screen and, after a 300 ms delay, played the audio recordings of the words through the loudspeakers. The children completed the task in the presence of an adult experimenter who controlled the computer program, which allowed the experimenter to replay the audio prompt if the first presentation did not elicit a clear repetition of the target word (e.g., if the child did not respond at all, produced a word other than the target, or talked over the audio prompt). The entire session of the repetition task was digitally recorded at a 44.1 kHz sampling rate onto a Marantz PMD660 flash card recorder for subsequent transcription and acoustic analysis.

Analysis

A group of trained phoneticians, who had no prior exposure to the speech of children with CIs, phonemically transcribed the initial consonants of the productions of the target words using a custom Praat script (Boersma & Weenink 2015), which allowed the user to listen to the auditory signal of each produced word and to visualize its waveform and spectrogram before transcribing its initial consonant. Only phonemically correct productions of the 9 /s/-initial words and the 9 /ʃ/-initial words that were elicited from both groups were included in the acoustic analysis. Further, for each child, at most one production of each target word was analyzed. Of the 396 target sibilant productions elicited from each group of children (22 × 18 = 396), the children with CIs produced 273 phonemically correct tokens (127 /s/ and 146 /ʃ/), and the children with NH produced 324 phonemically correct tokens (146 /s/ and 178 /ʃ/).

TABLE 2. List of /s/ and /ʃ/-initial words elicited during the word-repetition task

Vowel Context		/s/-Initial Words	/ʃ/-Initial Words
High front	/i/, /ɪ/	sister, seal, seashore	sheep, shield, ship
Mid front	/e/, /ɛ/	safe, same, seven	shape, shell, shepherd
Low central	/ɑ/, /ʌ/, /ɔ/	sauce, soccer, sun	shark, shop, shovel
Mid back	/o/	soak, sodas, soldier	shore, shoulder, show
High back	/u/, /ɝ/	soup, suitcase, super	chute, shoe, sugar

The onset of frication and the fricative–vowel boundary were marked after visually inspecting the waveform and spectrogram simultaneously in a Praat editor window. Frication onset was marked at the earliest point where high-frequency energy (>2.5 kHz) became visible in the spectrogram. The fricative–vowel boundary was marked at the waveform's zero-crossing nearest to the onset of a visible voicing bar in the spectrogram.

The times marked for frication onset and fricative–vowel boundary were used to define a sequence of seventeen 20-ms analysis windows, spaced evenly across the fricative, from which the psychoacoustic properties of the fricative productions were measured. This number of analysis windows was chosen in order to maintain consistency with previous work on the spectral and articulatory dynamics of sibilant fricatives. Both Iskarous et al. (2011) and Zharkova et al. (2014) used 9 windows; the current method generalizes this by adding a window between every sampling point used by prior studies. The waveform within each window was pre-emphasized (a = 0.95), and then the spectrum of the preprocessed waveform was estimated with an eighth-order multitaper spectrum (Thomson 1982).

Each spectral estimate was then passed through a filter bank that modeled the frequency selectivity of the auditory periphery. This filter bank comprised 361 fourth-order gammatone filters, whose center frequencies were spaced evenly from 3 to 39 (i.e., 0.1 spacing between adjacent channels) on the ERB number scale (ERB number = $21.4 \times \log10(1 + 0.00437f)$, with f in Hertz), a psychoacoustic frequency scale that models how the Hertz scale is compressed logarithmically and represented tonotopically on the basilar membrane (Moore & Glasberg 1983; Greenwood 1990). The bandwidth bw of each channel was proportional to its center frequency f in Hertz (bw = $1.019 \times ERB(f)$, where $ERB(f) = 24.7 \times (0.00437 \times f + 1)$). Hence, the gammatone filters in the filter bank were wider at higher frequencies than at lower frequencies, consistent with the frequency tuning of the auditory periphery. Once an input sound is passed through the gammatone filter bank, the output of each channel is summed to derive the total energy (or "excitation") within the channel in response to the input spectrum. Plotting the output excitations of the channels as a function of their center frequencies yields a psychoacoustic spectrum.

Three features were computed from each psychoacoustic spectrum to index sibilant contrast and degree of sibilance. First, the peak frequency of the psychoacoustic spectrum was computed by finding the ERB number of the filter channel with the maximum excitation level. Second, the variance of the psychoacoustic spectrum was computed by normalizing its excitation levels so that they summed to one and then treating this normalized psychoacoustic spectrum as a discrete probability mass function over ERB numbers and computing variance in the traditional way. Third, the difference in decibels between the maximum high-frequency (24.5–39 ERB numbers; 3–15 kHz) excitation level and the minimum low-frequency (3–24.5 ERB numbers; 0.85–3 kHz) excitation level was calculated. These three measures were computed from 17 psychoacoustic spectra estimated across the duration of each production; thus, each production was represented by three 17-point trajectories of psychoacoustic features.

The shapes of these trajectories were fitted with orthogonal polynomial growth-curve models (one model for each feature) to analyze temporal variation in these features. The growth-curve models included fixed effects of intercept, linear, quadratic, and cubic powers of time. These fixed effects denote, respectively, the level, slope, concavity, and jerk (or asymmetry across tails) of the trajectories. Additionally, fixed effects of consonant (/s/ or /ʃ/) and group (NH or CI), as well as binary and ternary interactions between them and any one of the temporal terms, were included to test differences in trajectory level or shape across consonants or groups. Finally, the growth-curve models included uncorrelated random effects of each temporal term by participants and by consonant-within-participants (to account for the repeated measures design of the repetition task). The significance of the fitted coefficients was assessed by bootstrapping 95% confidence intervals with 1000 replicates; a coefficient was considered significant if its confidence interval did not include 0, and in general, only significant effects are discussed below due to space limitations. Prediction intervals for each model were also bootstrapped with 1000 replicates, each time re-estimating the random effects of the model.

RESULTS

Figure 1 plots the mean duration of each target fricative separated by participant group. Productions of target /s/ by children with CIs (M = 231.72 ms, SD = 61.52 ms) were on average longer than those by children with NH (M = 203.99 ms, SD = 52.82 ms). Likewise, productions of target /ʃ/ by children with CIs (M = 256.15 ms, SD = 70.49 ms) were on average longer than those by children with NH (M = 211.32 ms, SD = 47.63 ms). The durations of the productions were modeled by a linear mixed-effects model with fixed effects of consonant (/s/ vs. /ʃ/), group (CI vs. NH), and a consonant-by-group interaction and with random effects of intercept by participant. The levels of the consonant and group effects were mapped to indicator variables in two different ways to explore all binary contrasts. Confidence intervals for the fitted models' coefficients were bootstrapped with an adjusted alpha level of 0.025 to account for multiple models. The first model, with
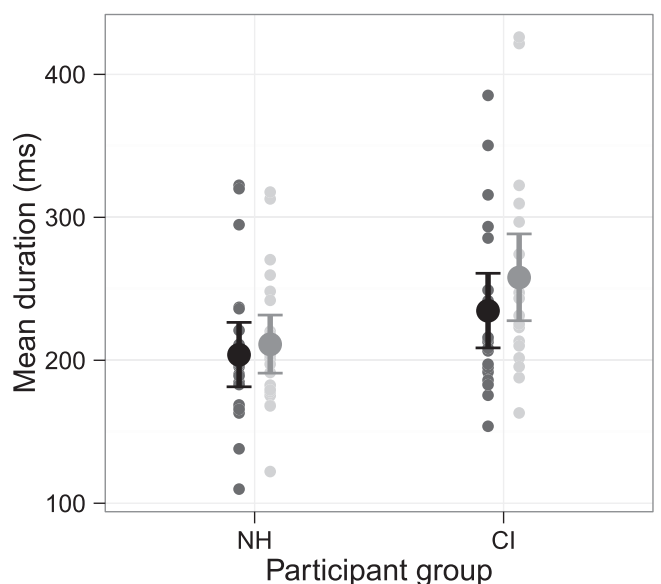

Fig. 1. Mean durations of productions of /s/ (black) and /ʃ/ (gray) for each group of children (large circles), shown with ±2 standard error bars. Mean durations for individual children are shown as backgrounded smaller circles. CI indicates cochlear implant; NH, normal hearing.

TABLE 3. Summary of phonemic accuracy judgments for each group and target consonant

Group	Target Consonant	Number of Phonemically Correct Tokens (% of 198 Possible)	Mean Accuracy Within Participant (SD; Range)
CI	/s/	127 (64%)	67% (20%; 33–100%)
	/ʃ/	146 (73%)	77% (22%; 22–100%)
NH	/s/	146 (73%)	74% (23%; 22–100%)
	/ʃ/	178 (90%)	90% (16%; 44–100%)

CI indicates cochlear implant; NH, normal hearing.

CI = 1 and /ʃ/ = 1, indicated that the two groups did not differ in the durations of their /s/ productions (β = 28.42, SE = 18.02, CI = [−8.76, 70.43]) and that the two target fricatives did not differ in their durations when produced by children with NH (β = 9.06, SE = 6.85, CI = [−5.05, 23.34]). The second model, with NH = 1 and /s/ = 1, indicated that the duration of /ʃ/ was significantly shorter when it was produced by children with NH (β = −45.69, SE = 17.77, CI = [−81.43, −5.36]) and that /s/ was significantly shorter than /ʃ/ when produced by children with CIs (β = −26.33, SE = 7.49, CI = [−43.51, −10.17]).

Table 3 summarizes the accuracy of participants within each group on each target consonant. The accuracy judgments were modeled with logistic mixed-effects regression to test the significance of these differences between groups or target consonants and to test the additional effect of age on phonemic accuracy. The model included simple fixed effects of group, consonant, and age; random effects of intercept by participant were included as well. As with the durational analysis above, the levels of the consonant and group effects were mapped to indicator variables in two different ways to explore all binary contrasts. An adjusted alpha level of 0.025 was used to account for multiple models. The first model, with CI = 1 and /ʃ/ = 1, indicated that on productions of target /s/, the children with CIs were significantly less accurate than the children with NH (β = −0.61, SE = 0.22, p < 0.01) and that the children with NH were significantly more accurate on target /ʃ/ than on target /s/ (β = 0.81, SE = 0.18, p < 0.001). The second model, with NH = 1 and /s/ = 1, indicated that on productions of target

/s/, the children with NH were significantly more accurate than the children with CIs (β = 0.61, SE = 0.22, p < 0.01) and that the children with CIs were significantly less accurate on target /s/ than on target /ʃ/ (β = −0.81, SE = 0.18, p < 0.001). Both models indicated that accuracy increased with age, but that this effect was not significant relative to the adjusted alpha level (β = 0.02, SE = 0.01, p > 0.03).

Figure 2 shows the peak ERB trajectories for word-initial /s/ and /ʃ/ productions for both the children with CIs (right panel) and their NH peers (left panel). These trajectories indicate that in the /s/ and /ʃ/ productions by both groups of children, peak ERB number rose across the first half of the fricative and fell across the second. In the fitted model, these temporal trends were indicated by a positive simple effect of linear time (β = 2.59, SE = 0.48, CI = [1.68, 3.52]) and a negative simple effect of quadratic time (β = −2.82, SE = 0.44, CI = [−3.65, −1.96]). The fitted model also included negative simple effects of consonant (β = −3.79, SE = 0.36, CI = [−4.56, −3.11]) and group (β = −1.88, SE = 0.46, CI = [−2.79, −1.02]), as well as a positive interaction between these simple effects (β = 2.28, SE = 0.52, CI = [1.28, 3.41]). These effects involving consonant and group indicated that in the productions by the children with NH, the peak frequency trajectory of /ʃ/ was approximately 3.79 ERB numbers lower than that for /s/. Contrastingly, in the productions by the children with CIs, the peak frequency trajectories of /s/ and /ʃ/ were separated only by approximately 1.51 ERB numbers. To further explore this interaction between consonant and group, post hoc models were built within each
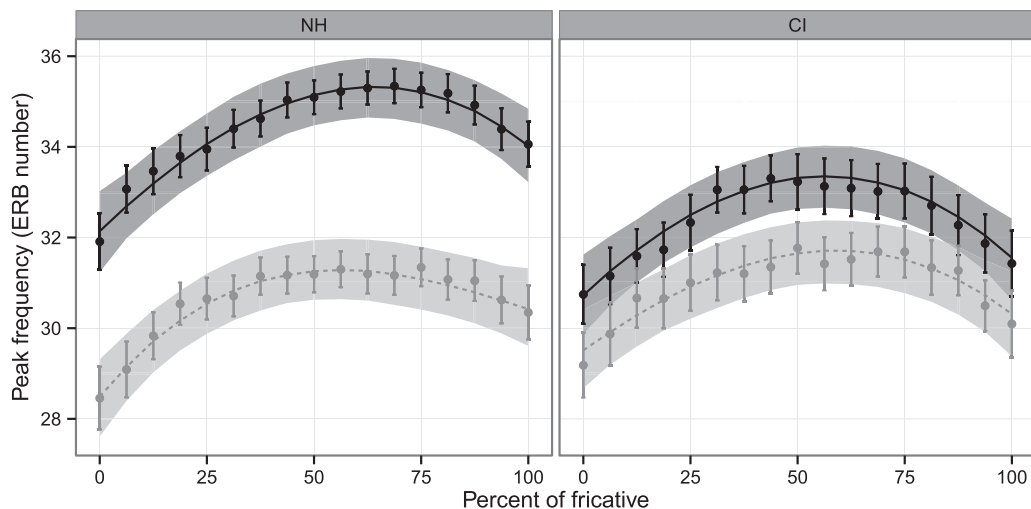

Fig. 2. Peak ERB number trajectories of /s/ (solid, black) and /ʃ/ (dashed, gray). Points and error bars denote the mean and standard errors of the data, respectively. The lines and ribbons denote the median predicted trajectory and the 95% prediction interval, bootstrapped from 1000 replicates. CI indicates cochlear implant; NH, normal hearing.

consonant. In the model fitted to just the productions of /s/, the simple effect of group was significant and negative ($\beta = -1.87$, SE = 0.45), whereas in the /ʃ/ model, the simple effect of group was positive but not significant ($\beta = 0.41$, SE = 0.47). These post hoc models thus suggest that the diminished psychoacoustic contrast in the productions by children with CIs was due primarily to their /s/ productions having a relatively lower peak frequency. Finally, a small negative interaction between linear time and group ($\beta = -1.43$, SE = 0.70, CI = [−2.74, −0.02]) indicated that in the productions by children with CIs, the overall linear trend of the peak ERB number trajectories was closer to zero. Hence, peak frequency followed a more symmetric rising-then-falling trajectory in these children's productions than in those by the children with NH. All other effects in the model were very small in magnitude ($|\beta| < 0.68$), and their confidence intervals all contained 0.

Figure 3 shows the variance trajectories for the same productions for both groups of children. The shapes of these trajectories differed more across groups than across consonants. For the children with NH, the variance trajectories of both /s/ and /ʃ/ followed convex curves that were, for the most part, symmetric about the temporal midpoint of the fricative. This shape was indicated in the fitted model by a positive simple effect of quadratic time ($\beta = 14.61$, SE = 1.89, CI = [10.75, 18.49]), but no significant effects of either linear ($\beta = -3.45$, SE = 1.83, CI = [−7.03, 0.26]) or cubic time ($\beta = 0.02$, SE = 0.98, CI = [−1.90, 1.98]). For the children with CIs, the variance trajectories of both /s/ and /ʃ/ showed greater asymmetry across the temporal midpoint, increasing across the second half of the fricative, as indicated by positive interactions of group with linear ($\beta = 8.98$, SE = 2.63, CI = [3.61, 14.29]) and cubic time ($\beta = 6.87$, SE = 1.43, CI = [3.98, 9.57]). A positive simple effect of consonant ($\beta = 4.92$, SE = 0.92, CI = [3.06, 6.76]) indicated that within each group, the /s/ and /ʃ/ trajectories were similar in shape, but that the psychoacoustic spectrum of /ʃ/ exhibited greater variance than that of /s/.

The excitation drop trajectories for the word-initial /s/ and /ʃ/ productions for both groups of children are shown in Figure 4. In the productions by children with NH, the excitation drop trajectories for both /s/ and /ʃ/ generally followed a rising-then-falling trajectory, with a relatively steeper rise across the first half and a relatively shallower fall across the second. This trajectory shape was indicated by a positive simple effect of linear time ($\beta = 8.36$, SE = 1.28, CI = [5.73, 10.94]) and a negative effect of quadratic time ($\beta = -19.38$, SE = 0.78, CI = [−21.01, −17.88]). The simple effect of cubic time was also significant ($\beta = -1.26$, SE = 0.57, CI = [−2.47, −0.21]), but smaller in magnitude than those of the lower powers of time. In the productions by children with CIs, the excitation drop trajectories for /s/ and /ʃ/ also followed rising-then-falling trajectories. However, the extent of curvature was decreased, as indicated by a positive interaction between group and quadratic time ($\beta = 4.84$, SE = 1.13, CI = [2.76, 7.00]). Furthermore, negative interactions of group with linear ($\beta = -8.66$, SE = 1.83, CI = [−12.30, −4.90]) and with cubic time ($\beta = -2.16$, SE = 0.83, CI = [−3.75, −0.60]) indicated that for the children with CIs, the excitation drop trajectories for /s/ and /ʃ/ decreased more steeply across the second half of the fricative than they rose across the first half. Finally, a negative simple effect of consonant ($\beta = -3.58$, SE = 0.82, CI = [−5.23, −1.88]) and a positive interaction between consonant and quadratic time ($\beta = 0.45$, SE = 1.04, CI = [1.46, 5.59]) indicated that within each group, excitation drop trajectories for /ʃ/ were lower in level and shallower in curvature.

DISCUSSION

Experiment 1 revealed that the subphonemic properties of correct productions of /s/ and /ʃ/ by children with CIs differed from those of their NH peers along multiple psychoacoustic dimensions. Foremost among these differences was a reduction of contrast in terms of peak frequency. This finding has previously been reported by Todd et al. (2011), who computed peak frequency from spectra estimated at fricative midpoint only. The current analyses, thus, replicated their finding with a psychoacoustic measure of peak frequency and extended it to a dynamic representation of fricative peak frequency. Post hoc analyses revealed that the productions of /s/ by children with CIs were lower in peak frequency than the productions by
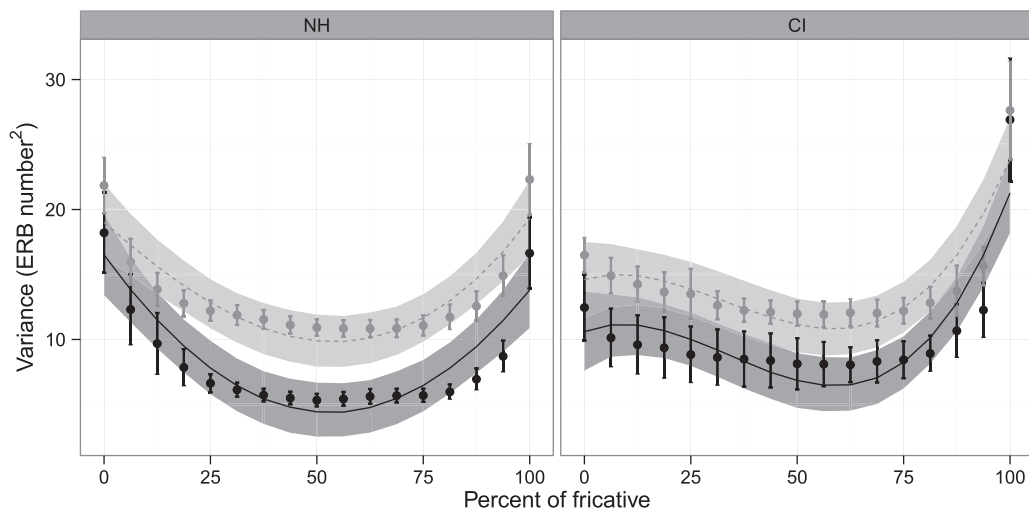

Fig. 3. Variance trajectories of /s/ (solid, black) and /ʃ/ (dashed, gray). Points and error bars denote the mean and standard errors of the data, respectively. The solid lines and ribbons denote the median predicted trajectory and the 95% prediction interval, bootstrapped from 1000 replicates. CI indicates cochlear implant; NH, normal hearing.
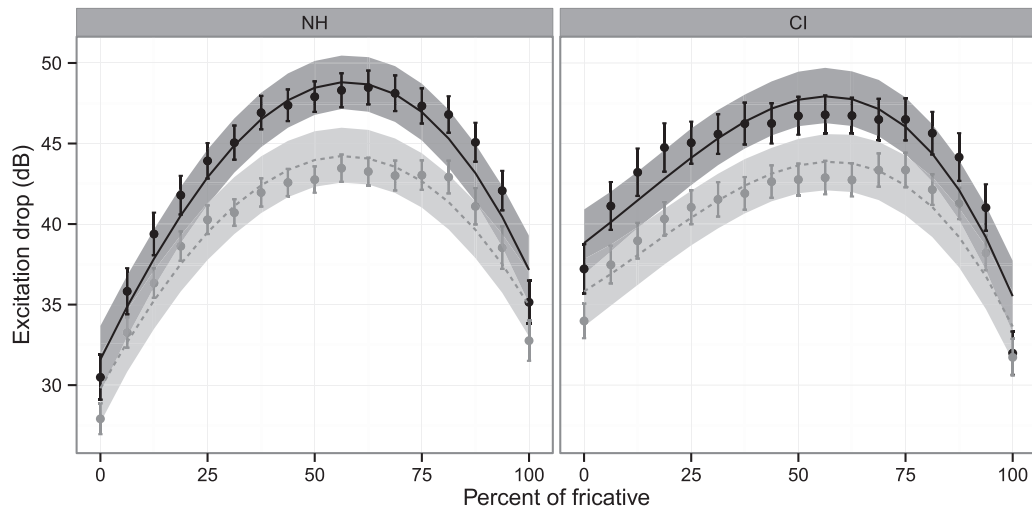
Fig. 4. Excitation drop trajectories of /s/ (solid, black) and /ʃ/ (dashed, gray). Points and error bars denote the mean and standard errors of the data, respectively. The solid lines and ribbons denote the median predicted trajectory and the 95% prediction interval, bootstrapped from 1000 replicates. CI indicates cochlear implant; NH, normal hearing.

children with NH, which is plausibly explained on perceptual grounds because the spectrum of /s/ has energy concentrated at high frequencies, which are not encoded by the CI speech processor. That is, across the middle half of productions of /s/ by children with NH, mean peak frequency was 34.91 ERB numbers, or approximately 9.68 kHz; the upper limit of CI frequency analysis is roughly 8 kHz. Conversely, the post hoc analysis of the peak frequency trajectory of /ʃ/ found no significant difference between the two groups, but the children with CIs showed a tendency to produce /ʃ/ with slightly higher peak frequency than their NH peers.

Additional group differences were revealed in terms of the dynamic aspects of the two measures that were included to index the degree of sibilance of the production. The variance trajectories exhibited the greatest qualitative difference in shape across the two groups toward the end of the fricative, where the productions by children with CIs became much more spectrally diffuse (i.e., less sibilant-like). This greater spectral variance in the second half of the productions by children with CIs suggests possible different coarticulatory strategies between the two groups; the children with CIs may have released the lingua-palatal constriction earlier in the consonant production, thus, coupling the back cavity resonances and increasing the spectral variance of the turbulent noise. For the excitation drop trajectories, dissimilarities between the two groups' productions were found at both the beginning and the end of the frication. Across the middle half of both /s/ and /ʃ/, the variance and excitation drop trajectories appeared comparable between the two groups of children. Taken together, these differences suggest that the productions by the two groups of children were comparably sibilant across the middle half of frication, but that the groups differed in the gestures used to form and then release the articulatory posture that is responsible for generating strong sibilant noise across the middle half of the fricative. These observations highlight the added value of dynamic measurements of fricative spectra above the more conventional midpoint analysis.

Previous analyses of the spectral dynamics of sibilant fricatives have focused exclusively on peak frequency trajectories (e.g., Iskarous et al. 2011; Reidy 2015). In Reidy's (2015)

analysis of the peak frequency trajectories of /s/ and /ʃ/, 4- and 5-year-old children differentiated these two consonants in terms of both the level and shape of the trajectory. However, in the present analysis, none of the interactions involving consonant and higher powers of time were significant, suggesting that neither the children with CIs nor the children with NH differentiated these sibilants in terms of trajectory shape, despite these children being the same age or older as the children reported by Reidy (2015). It is possible that this discrepancy is simply due to the smaller sample size of the present study. Reidy (2015) included 15 productions of each target consonant from 20 children from each age group. Previous analyses of the peak frequency dynamics in adults' productions of /s/ and /ʃ/ have found their trajectory shape to be asymmetric, with the rise in peak frequency exhibiting a greater extent than its fall; hence, the peak ERB trajectories for the children with NH (Fig. 2, left panel) seem to be more adult-like in shape than those for the children with CIs (right panel).

While Experiment 1 revealed a number of group differences in the subphonemic psychoacoustic properties of children's sibilant fricative productions, it is unclear whether they should be of concern to clinicians; however, this would be the case if such subphonemic differences also coincided with differences in intelligibility between the two groups. This question was the focus of Experiment 2.

EXPERIMENT 2

Motivation

A number of studies have found that the speech of children with CIs is less intelligible than that of their chronological age peers with NH. This relationship is consistent across studies that used different methods to estimate intelligibility (e.g., Chin et al. 2003; Peng et al. 2004; Chin et al. 2012; Chuang et al. 2012). Chin et al. (2003) evaluated speech intelligibility in English-speaking children using the Beginners' Intelligibility Test (Osberger et al. 1994). Children with CIs were judged to be 35% correct on average at the level of connected speech compared to children with NH who were judged to be 87%

correct on average. In more recent work, Chin et al. (2012) also found intelligibility scores to be higher for children with NH compared to children with CIs at the level of connected speech as measured by the Beginners' Intelligibility Test (i.e., listeners judging samples from children with NH were near ceiling and were about 80% correct for samples from children with CIs). Some studies, such as that by Baudonck et al. (2011), found that the lower intelligibility for children with CIs compared to children with NH did not reach significance. Overall, decreased speech intelligibility can affect the ability of children with CIs to communicate and socialize, especially in the classroom where learning needs must be communicated to teachers with little exposure to children with hearing loss. Listening environments that often contain multiple speakers further compound this decreased speech intelligibility. Furthermore, there appears to be a limited time window during which children with CIs refine their speech production skills. Tomblin et al. (2008) found that the development of speech sound production in prelingually deaf children stabilizes after 6 years of CI experience and, on average, approaches a plateau by 8 years of device use. Further understanding of the factors that account for the variability in speech intelligibility in children with CIs will allow clinicians and caregivers to maximize the impact of therapy during these time windows.

There are many factors that might lead to decreased speech intelligibility for children with CIs relative to their NH peers, including both segmental and suprasegmental differences in speech sound production between the two groups of children. The question in experiment 2 was whether subphonemic acoustic differences, such as a less robust acoustic contrast between /s/ and /ʃ/, could lead to reduced speech intelligibility at the word level. There is some evidence of a relation between a less distinct acoustic contrast between /s/ and /ʃ/ and poorer speech intelligibility in research on adults with NH. For example, Newman et al. (2001) observed longer reaction times for participants identifying /s/ versus /ʃ/ in quiet from talkers with more between-category overlap. Hazan and Baker (2011) represented adult talkers' productions of /s/ and /ʃ/ as samples of centroid frequency values and measured the distance between the mean of each sample (cross-category distance) and the mean standard deviation of the two samples (category dispersion) for each speaker. When these productions were then used as stimuli in a perception experiment, listeners were slowest to respond to productions from talkers with low cross-category distance and low-category dispersion.

There is a paucity of work on the relationship between subphonemic acoustic differences such as those observed in experiment 1 and speech intelligibility ratings in children with CIs relative to their NH peers. However, a relation between subphonemic acoustic measures and perceptual goodness ratings has been found across these two groups of children by Bernstein et al. (2013), who asked adults to rate the productions of /s/- and /ʃ/-initial consonant–vowel syllables from Todd et al. (2011). When adult listeners rated these syllables, the productions by children with CIs were rated as less good exemplars of the target sounds, relative to the productions by an age-matched group of children with NH. Furthermore, Bernstein et al. (2013) found that listeners responded more slowly to /s/ productions with lower spectral peaks, regardless of whether the syllables were produced by children with CIs or children with NH.

The purpose of experiment 2 was to determine whether the reduced acoustic contrast between productions of /s/ and /ʃ/ for children with CIs relative to children with NH resulted in decreased intelligibility of words containing these sounds in initial position. As in experiment 1, only productions that were judged to be phonemically correct in quiet were included. This decision was made because, as reported earlier, the phonemic accuracy of both target consonants was lower for the children with CIs than for their NH peers. Hence, we did not want any potential group differences in intelligibility to be confounded by their differences in phonemic accuracy. To prevent ceiling effects due to only accurate productions being included, a challenging listening environment was created by embedding these words within multi-talker babble. Adult listeners' accuracy in repeating the embedded words was measured. There were two main hypotheses. First, given that productions by children with CIs exhibited reduced acoustic contrast due to /s/ productions having lower peak frequency relative to their NH peers, we predicted that /s/-initial, but not /ʃ/-initial, words produced by children with CIs would be less intelligible than those produced by their NH peers. Second, we predicted that there would be a relationship between the psychoacoustic measures derived in experiment 1 and word-level intelligibility. The experiment was conducted with approval from the Institutional Review Board at the University of Wisconsin–Madison.

MATERIALS AND METHODS

Participants

Speakers • The stimuli came from a subset of children from experiment 1. All children included as speakers in experiment 2 produced at least 8 of the stimulus words from experiment 1 correctly. Supplemental Appendix A (http://links.lww.com/EANDH/A296) provides information on the number of /s/-initial, /ʃ/-initial, and filler words for each child; Table 4 provides descriptive information for the two groups of children.

Listeners • Eighty adults (40 males and 40 females) with an average age of 21 years (SD = 3.68 years, range = 18–35 years) participated. The participants were self-reported native English speakers with NH, no diagnoses or history of articulation disorders, and no training in phonetic transcription and no significant experience listening to the speech of children with

TABLE 4. Demographic characteristics of the two groups of children who participated in experiment 2

Group	Number of Speakers	Mean Chronological Age in Years;Months (SD, Range)	Mean Hearing Age (SD, Range)	Number of Male:Female Speakers	Average Receptive Vocabulary Standard Score (SD, Range)
CI	10	5;3 (1;1, 4;1–7;8)	4;0 (1;0, 2;9–6;5)	4:6	106 (15, 82–123)
NH	10	5;3 (1;0, 4;3–7;9)	—	4:6	110 (16, 85–145)

CI indicates cochlear implant; NH, normal hearing.

CIs (as evidenced by a short questionnaire before participation). Listeners were recruited from Madison, Wisconsin, via posting on a student job site and class announcements. To avoid familiarity effects due to word repetition, we used a between-subjects design so that each listener was randomly assigned to a single speaker. Therefore, each listener only heard a single production of each target word. The productions from each speaker were presented to 4 different listeners (20 Speakers × 4 listeners/speaker = 80 listeners).

Materials

Speech Stimuli • The speech stimuli were a subset of children's productions of the sibilant-initial target and stop-initial filler words recorded during experiment 1. The stimuli comprised only productions for which all segments in the word were transcribed as correct by a trained phonetician; productions in which a noninitial segment was not transcribed as accurate were excluded so that the listener's perception of the full word would not be biased. Furthermore, correct productions that included distortion due to signal clipping were excluded. Therefore, not all speakers listened to the same number of tokens. The target words were segmented from a larger recording using Praat and were root-mean-square-amplitude normalized.

Multi-Talker Babble • The four-talker babble was generated from recordings of four female adults producing sentences from various corpora: one talker produced sentences from the IEEE corpus (IEEE 1969), one talker produced sentences from the BKB corpus (Bench et al. 1979), and two talkers produced sentences from the AzBio corpus (Spahr et al. 2012). The use of four-talker babble helps to avoid effects of masking release from amplitude modulation where the signal may be presented in a randomly low amplitude portion of the babble, which would confound listener performance. Four-talker babble also reduces informational masking, which occurs when a listener can decode individual words from within the babble. Including a larger number of talkers in the babble would have further decreased the likelihood of informational masking; however, four-talker babble

offered a better ecological validity with the aims of the present study. While target word productions from the children were root-mean-square-amplitude normalized, the multi-talker babble was not standardized. After pilot testing to identify a challenging signal to noise ratio, 0 dB signal to noise ratio was selected so that listeners correctly identified about 60% of single words produced by children with NH. Random selections of the babble were added off-line to the individual speech samples using MAT-LAB. For each target or filler word produced by the children, the interval of babble was 2 seconds longer than the word production, so that once mixed, the stimuli comprised 1 second of babble, the masked word production, and then 1 second of babble.

Procedure

Participants were tested in a quiet room using a laptop, headphones, a serial response box with voice key, and microphone. The experiment was presented in E-Prime. Listeners completed two phases: a practice and a test phase. Before the practice and test phases, listeners were instructed as follows: they would first hear overlapping speech from several adult talkers; they would hear a single word spoken by a child embedded within the babble; and their job was to repeat this embedded word as quickly as possible once the babble had ceased. After the practice phase, the listeners judged words from the target talker, which were presented in randomized order. To help avoid a learning effect confound, the practice and test phases included different talkers. Listeners were asked to avoid starting a response with a filler, such as "umm, pizza." Listeners' oral responses were scored online by the second author using an informational word semantic match (Hustad 2006). Responses were scored as correct if the semantic intent of the word was preserved (i.e., morphological modifications were accepted).

RESULTS

Figure 5 shows the mean intelligibility scores for both groups and target consonants. These scores were modeled with
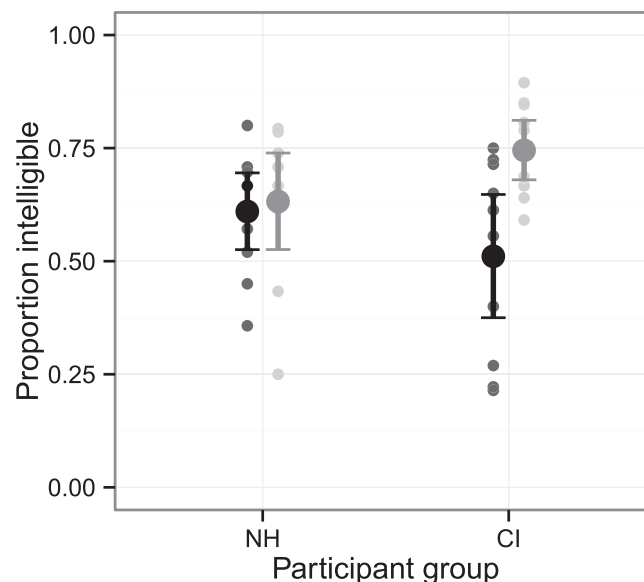

Fig. 5. Mean intelligibility scores for each group of children (large circles), shown with ±2 standard error bars, for /s/-initial words (black) and /ʃ/-initial words (gray). Mean intelligibility scores for individual children are shown as backgrounded smaller circles. CI indicates cochlear implant; NH, normal hearing.

logistic mixed-effects regression to test the significance of the differences between groups and target consonants. The model included simple fixed effects of group and of target consonant and their binary interaction; random effects of intercept by child were also included. The levels of the consonant and group effects were mapped to indicator variables in two different ways to explore all binary contrasts. An adjusted alpha level of 0.025 was used to account for multiple models. The first model, with CI = 1 and /ʃ/ = 1, indicated that the two groups did not differ in their accuracy on /s/ productions ($\beta = -0.43$, SE = 0.27, $p > 0.11$) and that the children with NH did not differ in their accuracy of /s/ versus /ʃ/ productions ($\beta = 0.07$, SE = 0.19, $p > 0.70$). The second model, with NH = 1 and /s/ = 1, indicated that the two groups did not differ in their accuracy on /ʃ/ productions ($\beta = -0.52$, SE = 0.29, $p > 0.07$) and that the children with CIs were significantly less accurate on /s/ productions than on /ʃ/ productions ($\beta = -1.03$, SE = 0.2, $p < 0.001$). In both models, the group-by-consonant interaction was significant ($\beta = 0.95$, SE = 0.29, $p < 0.005$), indicating that the difference in accuracy between the two consonants was greater for the children with CIs than for their NH peers.

To test whether word-level intelligibility scores were predicted by the acoustic properties of the initial consonant, the productions were pooled across group, but within consonant. One model was, thus, fitted to the /s/-initial words and one to the /ʃ/-initial words. To characterize the acoustics of the initial consonant of each word token, an orthogonal cubic polynomial model was fitted to each of the three psychoacoustic trajectories computed for each word-initial consonant. The coefficients of the fitted models (intercept and linear, quadratic, and cubic time), which summarized the shape of the trajectories, were used as predictor variables of intelligibility. Thus, each production had 12 possible acoustic predictors (4 Coefficients × 3 Models), but the cubic coefficient for the peak frequency trajectory was excluded a priori because neither its simple effect nor any of its interactions with group or consonant were significant in the peak frequency model reported in experiment 1. Each child's age and accuracy rate (from the phonemic transcriptions in experiment 1) were also included as predictors. A simple effect of group was not included as a predictor because the models above did not reveal a significant group difference in the intelligibility of either consonant. However, binary interactions between group and each of the 13 acoustic or child-level predictors were included to test group differences in the relationship between acoustics and intelligibility. No interactions between acoustic or child-level predictors were included. For space considerations, only significant effects are reported. Because intelligibility of each target consonant was modeled separately, an adjusted alpha level of 0.025 was used to account for multiple models.

The model fitted to the /s/-initial words had a Cox & Snell pseudo-$R^2 = 0.210$. There was a significant simple effect of accuracy rate ($\beta = 0.73$, SE = 0.29, $p < 0.025$), indicating that children who had higher phonemic accuracy were also more intelligible in noise. There was a significant simple effect of the quadratic coefficient of the peak frequency trajectory ($\beta = 0.39$, SE = 0.17, $p < 0.025$) and a significant interaction between it and the group ($\beta = -0.61$, SE = 0.27, $p < 0.025$). Because the mean quadratic coefficient for the peak frequency trajectories of the /s/ productions was negative (−2.60), these effects indicate

that the productions of /s/-initial words by children with NH were more intelligible as the curvature of the peak frequency trajectory flattened out, but that the productions by children with CIs were more intelligible as the peak frequency trajectory of the initial /s/ became more curved. The model also included a significant simple effect of the linear coefficient of the variance trajectory ($\beta = -1.53$, SE = 0.53, $p < 0.005$), indicating that intelligibility increased as the linear trend of the variance trajectory decreased. The mean linear coefficient for the variance trajectories of the /s/ productions was positive (2.55); hence, the negative effect of this coefficient indicates that intelligibility increased as the linear trend of the variance trajectory tends toward zero. Finally, the fitted model included significant simple effects of the linear ($\beta = -0.86$, SE = 0.29, $p < 0.005$) and cubic coefficients ($\beta = 0.57$, SE = 0.21, $p < 0.01$) of the excitation drop trajectory and a significant interaction between this trajectory's linear coefficient and group ($\beta = 1.19$, SE = 0.44, $p < 0.01$). For the excitation drop trajectories of the initial /s/ productions, the mean linear coefficient was positive (4.43); hence, the productions by the children with NH were more intelligible because the excitation drop trajectory of their initial /s/ productions were flatter in slope, but the productions by children with CIs were more intelligible when this trajectory increased more across the initial /s/. The mean cubic coefficient of the variance trajectories was negative (−3.15); hence, for both groups, intelligibility was greater in productions where the cubic trend was less present, and the trajectory exhibited less asymmetry across its tails.

For the /ʃ/-initial words, the fitted model's Cox & Snell pseudo-R^2 was 0.190. There was a significant simple effect of accuracy rate ($\beta = 0.57$, SE = 0.23, $p < 0.025$), which together indicated that intelligibility increased with judged accuracy rate. Regarding peak frequency, there were significant simple effects of its trajectory's linear ($\beta = 0.76$, SE = 0.25, $p < 0.01$) and quadratic trends ($\beta = 0.47$, SE = 0.21, $p < 0.025$), as well as a significant interaction between its linear trend and group ($\beta = -0.79$, SE = 0.32, $p < 0.025$). The mean linear coefficient of the peak frequency trajectory was positive (1.69). Hence, the simple effect of this coefficient indicated that the productions by children with NH were more intelligible as the linear trend in the trajectory became even more exaggerated; however, the interaction between this term and group indicated that this was not the case in the productions by children with CIs. The mean quadratic coefficient of the peak frequency trajectory was negative (−3.21), indicating downward concave curvature; hence, the positive effect of this coefficient indicates that intelligibility increased as the trajectory became flatter in curvature. There was a significant simple effect of the mean level of the variance ($\beta = 1.88$, SE = 0.40, $p < 0.001$) trajectory and a significant interaction between it and group ($\beta = -1.34$, SE = 0.51, $p < 0.01$), indicating that intelligibility increased significantly as mean spectral variance increased, but that this increase in intelligibility was significantly smaller for productions by the children with CIs. Recalling from experiment 1 that /ʃ/ spectra exhibit greater variance than /s/ spectra, the effect of variance level may indicate that tokens with greater variance are more intelligible because they are less like /s/. Finally, there was a significant simple effect of the mean level of the excitation drop trajectory ($\beta = 0.95$, SE = 0.35, $p < 0.01$), indicating that productions became more intelligible as the relative height of the spectral peak increased.

DISCUSSION

Our first prediction that there would be group differences in intelligibility of /s/-initial words, but not in intelligibility of /ʃ/-initial words, was not fully confirmed. Across groups, the intelligibility of /s/-initial words was lower for children with CIs than it was for their NH peers; however, this difference was not statistically significant. Conversely, the intelligibility of /ʃ/-initial words was greater for children with CIs than for their NH peers; again this difference was not significant. The absence of group differences could have been due to the fact that only tokens that were judged phonemically accurate in quiet were used as stimuli in the listening experiment. If inaccurate productions had also been used, then group differences may have emerged because children with CIs are often less accurate at a segmental level than their NH peers (see Smit et al. 1990; Chin 2003); however, in such a situation, intelligibility in noise would be confounded with phonemic accuracy. Despite the absence of group differences for either initial consonant individually, the intelligibility of sibilant-initial words by children with CIs was more sensitive to the place of articulation of the initial consonant. Transcription analyses have shown that children with CIs are typically more accurate on target /ʃ/ than target /s/ (e.g., Serry & Blamey 1999; Blamey et al. 2001; Reidy et al. 2015). However, in these previous studies, the accuracy judgments were made in quiet. The current finding suggests that even when tokens that were judged to be phonemically correct in quiet are embedded in noise, this same segmental deficit is reflected in lower intelligibility of /s/-initial versus /ʃ/-initial tokens.

One possible explanation for the intelligibility results is that the articulatory gestures that children with CIs learn for a given sound reflect the effect of a CI processor on the auditory properties of that sound, resulting in poorer speech intelligibility scores for sounds whose spectral information is more degraded by the CI processor. The auditory representation that a child learns for a given speech sound reflects the auditory properties of the tokens of that sound to which the child has been exposed (e.g., Cristià 2011). For listeners with NH, the perceived auditory signal of a sound is very similar to its produced acoustic signal. However, for listeners with a CI, there is greater dissimilarity between the auditory and acoustic properties of a sound due to the CI processor's reduced spectral resolution and limited analysis bandwidth. In productions of sibilant fricatives by talkers with NH, energy is most concentrated between 7 and 10 kHz for /s/ and between 4 and 6 kHz for /ʃ/ (Jongman et al. 2000; Li 2012); hence, a CI transmits less of the salient spectral content for /s/ than for /ʃ/. Consequently, the dissimilarity between CI-transmitted auditory properties and produced acoustic properties is greater for /s/ than for /ʃ/. During language acquisition, a child uses auditory feedback to learn an articulatory gesture for producing a given speech sound, comparing the auditory signal of their production to their auditory representation learned from exposure to tokens produced by caregivers (Plummer 2014). In the case where the CI user is a prelingually deaf child, the dissimilarity between the child's auditory representation and the acoustic properties of ambient productions of a given sound may, thus, propagate to the child's articulatory gesture for the sound because this articulatory gesture is learned through the articulatory–auditory feedback loop. As a result, the fact that the CI processor preserves the acoustic–auditory properties of some sounds better than others (i.e., /ʃ/ better than /s/) may

likely explain why children with CIs are more intelligible on /ʃ/-initial words than on /s/-initial words.

Our second prediction was that intelligibility would be partially predicted from the acoustic properties of the initial consonant. This prediction was confirmed. However, it must be emphasized that the pseudo-R^2 of the fitted models were not very large (0.210 for the /s/-initial words and 0.190 for the /ʃ/-initial words), suggesting that intelligibility depends on a number of other factors not considered here. This result suggests a limitation of the present study: that intelligibility depends on a number of other factors not considered here. These factors may include children's speech production skills more generally as well as factors related to the acoustic properties of speech in noise. Regarding the former, future work should investigate whether a child's scores on clinical instruments such as norm-referenced articulation tests further predict the child's speech intelligibility. Regarding the latter, future work should consider acoustic measures of consonant contrast that span multiple segments—for example, the relative difference in the spectral envelope of the fricative and vowel across a consonant–vowel boundary (cf. Hedrick 1997; Hedrick & Carney 1997)—or acoustic measures of the total proportion of "spectral glimpses" that were not masked by the babble (cf. Cooke 2006).

The models wherein intelligibility was regressed against acoustic properties of the initial consonant indicated that the significant predictors primarily reflected dynamic rather than static properties of the acoustics. For the /s/-initial words, all significant acoustic effects indexed either the linear, quadratic, or cubic trend of some psychoacoustic trajectory. Likewise, intelligibility of /ʃ/-initial words was significantly related in some way to five different acoustic measures, three of which characterized either the linear or quadratic trend of a psychoacoustic trajectory. That dynamic measures were preponderant over static measures in predicting intelligibility in noise suggests that future work should characterize the acoustics of children's /s/ and /ʃ/ productions with dynamic measures of multiple acoustic features.

GENERAL CONCLUSIONS

This article presented two experiments related to the properties of sibilant fricatives produced by children with CIs. The first experiment analyzed these productions in terms of one feature—peak frequency—that indexed the place of articulation difference between /s/ and /ʃ/ and two features—spectral variance and excitation drop—that indicated the degree of sibilance of a production. It was found that, in terms of peak frequency, the children with CIs produced less acoustic contrast than their NH peers between /s/ and /ʃ/. This difference in acoustic contrast was likely due to the children with CIs producing these two sibilants with more similar places of articulation. Additional group differences were found in the dynamic aspects of the variance and excitation drop trajectories, especially near the beginnings and ends of the fricatives, suggesting differences between groups in the articulatory gestures executed to produce these sibilants. A direction for future research is to investigate the underlying articulatory correlates of these group differences and to relate the acoustic properties of children with CIs' productions to their perceptual abilities.

The second experiment investigated the intelligibility in noise of sibilant-initial words that were judged to be

phonemically correct in quiet. Here, it was found that productions by children with NH were equally intelligible regardless of the initial consonant; however, the /s/-initial word productions by children with CIs were less intelligible than their productions of /ʃ/-initial words. Furthermore, the intelligibility of the word productions was partially predicted by the acoustic properties of the initial consonant, suggesting that the subphonemic differences between groups observed in experiment 1 have consequences for the intelligibility of whole word productions in background noise.

While group differences were found in both the subphonemic acoustic properties and intelligibility of word-initial sibilant productions, these findings should be interpreted in the context of the study's limitations. First, the participants in the present study were all young children aged between 4 and 7 years; however, the acoustic properties of a child's productions of the sibilant fricatives /s/ and /ʃ/ continue to develop toward adult-like acoustics into the adolescent years. Romeo et al (2013) found that, in productions of /s/ and /ʃ/ by children between 9 and 14 years, the cross-category distance and category dispersion developed toward adult-like norms, but there were still significant differences between the adolescents and the adults in terms of these category properties. The present study should, thus, be understood as providing just a snapshot of one point in time during the children's development of adult-like categories. This is especially true for the children with CIs, who had on average just less than 4 years of experience with at least one CI. Tomblin et al. (2008) found that the speech production skills of children with CIs continued to improve up through 8 years after implantation; hence, it is plausible that the performance of the children in the present study will continue to improve toward that of their NH peers in subsequent years.

A second, possible, limitation of the present study concerns the methodology used to elicit productions from the children. The acoustic properties of adults' speech is known to vary according to whether an audio prompt is used to elicit that speech or not (e.g., D'Imperio & German, 2015 and D'Imperio et al, 2015, report this for schwa and f0). Hence, eliciting speech with an audio prompt, rather than having t.he participants read or spontaneously name the pictures, could have led to the children using a more careful speech register than they typically use. Consequently, for each group, the acoustic differences observed between the consonants may be greater than would be expected in conversational speech. It is not immediately apparent that the group differences in acoustics and intelligibility are dependent upon the elicitation methodology. We leave it as a question for future research to determine whether children with NH and children with CIs are differently sensitive to how their speech is elicited.

Despite these limitations, the current findings are of importance for researchers and clinicians. The findings that group differences were observed on the dynamic aspects of the acoustic measures and that a greater number of acoustic predictors of intelligibility were dynamic rather than static in nature suggest that researchers should consider time-varying spectral representations of sibilants and other consonants when studying children's speech (cf. Nossair & Zahorian 1991; Assmann & Katz 2000). Likewise, these results invite clinicians to focus therapy on speech gestures rather than static articulatory targets. In addition, the dissociation between accuracy in quiet and intelligibility in noise (experiment 2) underscores the importance both for clinicians to consider the intelligibility of child in real-world settings, rather than just in quiet, and for speech–language therapy to continue even after children with CIs are able to produce sounds correctly.

REFERENCES

Assmann, P. F., & Katz, W. F. (2000). Time-varying spectral change in the vowels of children and adults. *J Acoust Soc Am*, *108*, 1856–1866.

Baudonck, N., Van Lierde, K., D'haeseleer, E., et al. (2011). A comparison of the perceptual evaluation of speech production between bilaterally implanted children, unilaterally implanted children, children using hearing aids, and normal-hearing children. *Int J Audiol*, *50*, 912–919.

Bench, J., Kowal, A., Bamford, J. (1979). The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *Br J Audiol*, *13*, 108–112.

Bernstein, S., Todd, A., Edwards, J. (2013). How do adults perceive the speech of children with cochlear implants? Poster Presented at the Annual Conference of the American Speech-Language-Hearing Association, Chicago, IL, 14–16.

Bird, J., Bishop, D. V., Freeman, N. H. (1995). Phonological awareness and literacy development in children with expressive phonological impairments. *J Speech Hear Res*, *38*, 446–462.

Blamey, P. J., Barry, J. G., Jacq, P. (2001). Phonetic inventory development in young cochlear implant users 6 years postoperation. *J Speech Lang Hear Res*, *44*, 73–79.

Boersma, P. & Weenink, D. (2015). *Praat: Doing phonetics by computer [Computer program]*. Version 5.4.22.

Brownell, R. (Ed.). (2000). Receptive One Word Picture Vocabulary Test (2nd ed.). Novato, CA: Academic Therapy Publication, Inc.

Burkholder, R. A., & Pisoni, D. B. (2003). Speech timing and working memory in profoundly deaf children after cochlear implantation. *J Exp Child Psychol*, *85*, 63–88.

Chin, S. B. (2003). Children's consonant inventories after extended cochlear implant use. *J Speech Lang Hear Res*, *46*, 849–862.

Chin, S. B., Bergeson, T. R., Phan, J. (2012). Speech intelligibility and prosody production in children with cochlear implants. *J Commun Disord*, *45*, 355–366.

Chin, S. B., Tsai, P. L., Gao, S. (2003). Connected speech intelligibility of children with cochlear implants and children with normal hearing. *Am J Speech Lang Pathol*, *12*, 440–451.

Chuang, H. F., Yang, C. C., Chi, L. Y., et al. (2012). Speech intelligibility, speaking rate, and vowel formant characteristics in Mandarin-speaking children with cochlear implant. *Int J Speech Lang Pathol*, *14*, 119–129.

Connor, C. M., Craig, H. K., Raudenbush, S. W., et al. (2006). The age at which young deaf children receive cochlear implants and their vocabulary and speech-production growth: is there an added value for early implantation? *Ear Hear*, *27*, 628–644.

Cooke, M. (2006). A glimpsing model of speech perception in noise. *J Acoust Soc Am*, *119*, 1562–1573.

Cristià, A. (2011). Fine-grained variation in caregivers' /s/ predicts their infants' /s/ category. *J Acoust Soc Am*, *129*, 3271–3280.

Cupples, L., Ching, T. Y., Crowe, K., et al. (2014). Predictors of early reading skill in 5-year-old children with hearing loss who use spoken language. *Read Res Q*, *49*, 85–104.

D'Imperio, M., & German, J. S. (2015). Phonetic detail and the role of exposure in dialect imitation. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th ICPhS*. Glasgow, UK: The University of Glasgow. Paper no. 1009.

D'Imperio, M., Petrone, C., Graux-Czachor, C. (2015). The influence of metrical constraints on direct imitation across French varieties. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th ICPhS*. Glasgow, UK: The University of Glasgow. Paper no. 0626.

Dunn, L. M. & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test*, (4th ed.). San Antonio, TX: Pearson Assessments.

Ertmer, D. J., & Goffman, L. (2011). Speech production accuracy and variability in young cochlear implant recipients: comparisons with typically developing age-peers. *J Speech Lang Hear Res*, *54*, 177–189.

Forrest, K., Weismer, G., Milenkovic, P., et al. (1988). Statistical analysis of word-initial voiceless obstruents: preliminary data. *J Acoust Soc Am*, *84*, 115–123.

Friesen, L. M., Shannon, R. V., Baskent, D., et al. (2001). Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants. *J Acoust Soc Am*, *110*, 1150–1163.

Fuddle, J. B. (1974). *Arizona Articulation Proficiency Scale: Revised*. Los Angeles: Western Psychological Services.

Geers, A. E., Nicholas, J. G., Sedey, A. L. (2003). Language skills of children with early cochlear implantation. *Ear Hear*, *24*(1 Suppl), 46S–58S.

Goldman, R. & Fristoe, M. (2000). *Goldman-Fristoe Test of Articulation*, (2nd ed.). San Antonio, TX: Pearson.

Greenwood, D. D. (1990). A cochlear frequency-position function for several species–29 years later. *J Acoust Soc Am*, *87*, 2592–2605.

Hazan, V. & Baker, R. (2011). Is consonant perception linked to within-category dispersion or across-category distance? In W.-S. Lee & E. Zee (Eds.), *Proceedings of the 17th ICPhS* (pp. 839–842). Hong Kong: City University of Hong Kong.

Hedrick, M. (1997). Effect of acoustic cues on labeling fricatives and affricates. *J Speech Lang Hear Res*, *40*, 925–938.

Hedrick, M. S., & Carney, A. E. (1997). Effect of relative amplitude and formant transitions on perception of place of articulation by adult listeners with cochlear implants. *J Speech Lang Hear Res*, *40*, 1445–1457.

Hess, C., Zettler-Greeley, C., Godar, S. P., et al. (2014). The effect of differential listening experience on the development of expressive and receptive language in children with bilateral cochlear implants. *Ear Hear*, *35*, 387–395.

Hughes, G. W., & Halle, M. (1956). Spectral properties of fricative consonants. *J Acoust Soc Am*, *28*(2), 303–310.

Hustad, K. C. (2006). A closer look at transcription intelligibility for speakers with dysarthria: evaluation of scoring paradigms and linguistic errors made by listeners. *Am J Speech Lang Pathol*, *15*, 268–277.

IEEE. (1969). IEEE recommended practice for speech quality measurements. *IEEE Trans. Audio Electroacoust.*, *17*, 225–246.

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

Iverson, P. (2003). Evaluating the function of phonetic perceptual phenomena within speech recognition: an examination of the perception of /d/-/t/ by adult cochlear implant users. *J Acoust Soc Am*, *113*, 1056–1064.

James, D., Rajput, K., Brown, T., et al. (2005). Phonological awareness in deaf children who use cochlear implants. *J Speech Lang Hear Res*, *48*, 1511–1528.

Jongman, A., Wayland, R., Wong, S. (2000). Acoustic characteristics of English fricatives. *J Acoust Soc Am*, *108*(3 Pt 1), 1252–1263.

Koenig, L. L., Shadle, C. H., Preston, J. L., et al. (2013). Toward improved spectral measures of /s/: results from adolescents. *J Speech Lang Hear Res*, *56*, 1175–1189.

Li, F. (2012). Language-specific developmental differences in speech production: a cross-language acoustic study. *Child Dev*, *83*, 1303–1315.

Li, F., Edwards, J., Beckman, M. E. (2009). Contrast and covert contrast: The phonetic development of voiceless sibilant fricatives in English and Japanese toddlers. *J Phon*, *37*, 111–124.

Loizou, P. C. (2006). Speech processing in vocoder-centric cochlear implants. *Adv Otorhinolaryngol*, *64*, 109–143.

Lonigan, C. J., Burgess, S. R., Anthony, J. L. (2000). Development of emergent literacy and early reading skills in preschool children: evidence from a latent-variable longitudinal study. *Dev Psychol*, *36*, 596–613.

Melby-Lervåg, M., Lyster, S. A., Hulme, C. (2012). Phonological skills and their role in learning to read: a meta-analytic review. *Psychol Bull*, *138*, 322–352.

Middlebrooks, J. C., Bierer, J. A., Snyder, R. L. (2005). Cochlear implants: the view from the brain. *Curr Opin Neurobiol*, *15*, 488–493.

Misurelli, S. M., & Litovsky, R. Y. (2012). Spatial release from masking in children with normal hearing and with bilateral cochlear implants: effect of interferer asymmetry. *J Acoust Soc Am*, *132*, 380–391.

Misurelli, S. M., & Litovsky, R. Y. (2015). Spatial release from masking in children with bilateral cochlear implants and with normal hearing: Effect of target-interferer similarity. *J Acoust Soc Am*, *138*, 319–331.

Moore, B. C., & Glasberg, B. R. (1983). Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. *J Acoust Soc Am*, *74*, 750–753.

Munson, B. (2004). Variability in /s/ production in children and adults: evidence from dynamic measures of spectral mean. *J Speech Lang Hear Res*, *47*, 58–69.

Munson, B., Donaldson, G. S., Allen, S. L., et al. (2003). Patterns of phoneme perception errors by listeners with cochlear implants as a function of overall speech perception ability. *J Acoust Soc Am*, *113*, 925–935.

Narayanan, S. S. & Alwan, A. A. (2000). Noise source models for fricative consonants. *IEEE T Speech Audi P*, *8*(2), 328–344.

Newman, R. S., Clouse, S. A., Burnham, J. L. (2001). The perceptual consequences of within-talker variability in fricative production. *J Acoust Soc Am*, *109*, 1181–1196.

Nittrouer, S. (1995). Children learn separate aspects of speech production at different rates: evidence from spectral moments. *J Acoust Soc Am*, *97*, 520–530.

Nittrouer, S., Caldwell, A., Lowenstein, J. H., et al. (2012). Emergent literacy in kindergartners with cochlear implants. *Ear Hear*, *33*, 683–697.

Nittrouer, S., Studdert-Kennedy, M., McGowan, R. S. (1989). The emergence of phonetic segments: evidence from the spectral structure of fricative-vowel syllables spoken by children and adults. *J Speech Hear Res*, *32*, 120–132.

Nossair, Z. B., & Zahorian, S. A. (1991). Dynamic spectral shape features as acoustic correlates for initial stop consonants. *J Acoust Soc Am*, *89*(6), 2978–2991.

Osberger, M. J., Maso, M., Sam, L. K. (1993). Speech intelligibility of children with cochlear implants, tactile aids, or hearing aids. *J Speech Hear Res*, *36*, 186–203.

Osberger, M. J. Robbins, A. M., Todd, S. L. et al. (1994). Speech intelligibility of children with cochlear implants. *Volta Rev*, *96*, 169–180.

Peng, S. C., Spencer, L. J., Tomblin, J. B. (2004). Speech intelligibility of pediatric cochlear implant recipients with 7 years of device experience. *J Speech Lang Hear Res*, *47*, 1227–1236.

Plummer, A. R. (2014). The acquisition of vowel normalization during early infancy: theory and computational framework. Unpublished Ph.D. dissertation, The Ohio State University, Columbus, OH.

Raggio, M. W., & Schreiner, C. E. (2003). Neuronal responses in cat primary auditory cortex to electrical cochlear stimulation: IV. Activation pattern for sinusoidal stimulation. *J Neurophysiol*, *89*, 3190–3204.

Reidy, P. F. (2015). The spectral dynamics of voiceless sibilant fricatives in English and Japanese. Unpublished doctoral dissertation. Ohio State University, Columbus, OH.

Reidy, P. F., Beckman, M. E., Litovsky, R. Y., et al. (2015). The acquisition of English sibilant fricatives by children with bilateral cochlear implants. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th ICPhS*. Glasgow, UK: The University of Glasgow. Paper no. 0219.

Romeo, R., Hazan, V., Pettinato, M. (2013). Developmental and gender-related trends of intra-talker variability in consonant production. *J Acoust Soc Am*, *134*, 3781–3792.

Rvachew, S., & Jamieson, D. G. (1989). Perception of voiceless fricatives by children with a functional articulation disorder. *J Speech Hear Disord*, *54*, 193–208.

Rvachew, S., Nowak, M., Cloutier, G. (2004). Effect of phonemic perception training on the speech production and phonological awareness skills of children with expressive phonological delay. *Am J Speech Lang Pathol*, *13*, 250–263.

Serry, T. A., & Blamey, P. J. (1999). A 4-year investigation into phonetic inventory development in young cochlear implant users. *J Speech Lang Hear Res*, *42*, 141–154.

Shadle, C. H. (1985). *The Acoustics of Fricative Consonants*. Technical Report 506, Cambridge, MA: MIT Research Laboratory of Electronics.

Shadle, C. H. (1990). Articulatory-acoustic relationships in fricative consonants. In W.J. Hardcastle & A. Marchal (Eds.),*Speech Production and Speech Modeling* (pp. 187–209). Netherlands: Springer.

Smit, A. B., Hand, L., Freilinger, J. J., et al. (1990). The Iowa articulation norms project and its Nebraska replication. *J Speech Hear Disord*, *55*, 779–798.

Soli, S. D. (1981). Second formants in fricatives: Acoustic consequences of fricative-vowel coarticulation. *J Acoust Soc Am*, *70*(4), 976–984.

Spahr, A. J., Dorman, M. F., Litvak, L. M., et al. (2012). Development and validation of the Pediatric AzBio sentence lists. *Ear Hear*, *33*, 418–422.

Spencer, L., Tomblin, J. B., Gantz, B. J. (1997). Reading Skills in Children with Multichannel Cochlear-lmplant Experience. *Volta Rev*, *99*, 193–202.

Spencer, L. J., Gantz, B. J., Knutson, J. F. (2004). Outcomes and achievement of students who grew up with access to cochlear implants. *Laryngoscope*, *114*, 1576–1581.

Spencer, L. J., Tye-Murray, N., Tomblin, J. B. (1998). The production of English inflectional morphology, speech production and listening performance in children with cochlear implants. *Ear Hear*, *19*, 310–318.

Thomson, D. J. (1982). Spectrum estimation and harmonic analysis. *Proc IEEE*, *70*, 1055–1096.

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*, 3969–3979.

Tomblin, J. B., Peng, S. C., Spencer, L. J., et al. (2008). Long-term trajectories of the development of speech sound production in pediatric cochlear implant recipients. *J Speech Lang Hear Res*, *51*, 1353–1368.

Tomblin, J. B., Spencer, L., Flock, S., et al. (1999). A comparison of language achievement in children with cochlear implants and children using hearing aids. *J Speech Lang Hear Res*, *42*, 497–509.

Uchanski, R. M., & Geers, A. E. (2003). Acoustic characteristics of the speech of young cochlear implant users: a comparison with normal-hearing age-mates. *Ear Hear*, *24*(1 Suppl), 90S–105S.

Warner-Czyz, A. D., & Davis, B. L. (2008). The emergence of segmental accuracy in young cochlear implant recipients. *Cochlear Implants Int*, *9*, 143–166.

Webb, M. Y., & Lederberg, A. R. (2014). Measuring phonological awareness in deaf and hard-of-hearing children. *J Speech Lang Hear Res*, *57*, 131–142.

Zharkova, N., Hewlett, N., Hardcastle, W. J., et al. (2014). Spatial and temporal lingual coarticulation and motor control in preadolescents. *J Speech Lang Hear Res*, *57*, 374–388.