Phonological Priming in Children with Hearing Loss:

Effect of Speech Mode, Fidelity, and Lexical Status

Susan Jerger, Ph.D.*#

Nancy Tye-Murray, Ph.D.+

Markus F. Damian, Ph.D.^

Hervé Abdi, Ph.D.*

University of Texas at Dallas*

Callier Center for Communication Disorders#

Central Institute for the Deaf of Washington University School of Medicine+

University of Bristol, UK^

Conflicts of Interest: None.
Source of Funding: This work was supported by the National Institute on Deafness and Other Communication Disorders, grant DC-00421 to the University of Texas at Dallas

Corresponding Author: Susan Jerger, School of Behavioral and Brain Sciences, GR4.1, University of Texas at Dallas, 800 W. Campbell Rd, Richardson, TX 75080, sjerger@utdallas.edu, Phone: +1 972-883-2355; FAX - 972-883-2491.

[1]Present address: 14070 Goodall Rd., Lake Oswego, OR 97034, Phone: +1 512-216-2961

1

Abstract

Objectives. Our research determined 1) how phonological priming of picture naming was affected by the mode (audiovisual [AV] vs auditory), fidelity (intact vs non-intact auditory onsets), and lexical status (words vs nonwords) of speech stimuli in children with prelingual sensorineural hearing impairment (CHI) vs. children with normal hearing (CNH); and 2) how the degree of hearing impairment (HI), auditory word recognition, and age influenced results in CHI. Our AV stimuli were not the traditional bimodal input but instead consisted of an intact consonant/rhyme in the visual track coupled to a non-intact onset/rhyme in the auditory track. Example stimuli for the word *bag* are: 1) AV: intact visual (b/ag) coupled to non-intact auditory (**–***b*/ag) and 2) Auditory: static face coupled to the same non-intact auditory (**–***b*/ag). Our question was whether the intact visual speech would "restore or fill-in" the non-intact auditory speech in which case performance for the *same* auditory stimulus *would differ* depending upon the presence/absence of visual speech.

Design. Participants were 62 CHI and 62 CNH whose ages yielded a group-mean and -distribution akin to that in the CHI group. Ages ranged from 4 to 14 years. All participants: 1) spoke English as a native language, 2) were able to successfully communicate aurally/orally, and 3) were not diagnosed or with disabilities other than HI and its accompanying verbal problems. The phonological priming of picture naming was assessed with the multi-modal picture word task.

Results. Both CHI and CNH showed greater phonological priming from high than low fidelity stimuli and from AV than auditory speech. These overall fidelity and mode effects did not differ in the CHI vs. CNH—thus these CHI appeared to have sufficiently well specified phonological onset representations to support priming, and visual speech did not appear to be a disproportionately important source of the CHI's phonological knowledge. Two exceptions occurred, however. *First*—with regard to lexical status—both the CHI and CNH showed significantly greater phonological priming from the nonwords than words, a pattern consistent with the prediction that children are more aware of phonetics-phonology content for nonwords. This overall pattern of similarity between the groups was qualified by the finding that CHI showed more nearly equal priming by the high vs. low fidelity nonwords than the CNH; in other words, the CHI were less affected by the fidelity of the auditory input for nonwords. *Second*, auditory word recognition—but not degree of HI or age—uniquely influenced phonological priming by the nonwords presented AV.

Conclusions. With minor exceptions, phonological priming in CHI and CNH showed more similarities than differences. Importantly, we documented that the addition of visual speech significantly increased phonological priming in both groups. Clinically these data support intervention programs that view visual speech as a powerful asset for developing spoken language in CHI.

Children learn phonemes—the building blocks of spoken language—through their early communicative experience. Determining which part(s) of this communicative experience is critical for phonetic learning has been the focus of much previous research. Currently, the primary view proposes that children learn phonology via hearing/overhearing a variety of talkers. For this view, access to auditory input is essential for successful phonological development (e.g., Tye-Murray 1992; Tye-Murray et al. 1995; Moeller et al. 2010; Klein & Rapin 2013).

Evidence supporting a disproportionately important role for auditory input is that children with prelingual sensorineural hearing impairment develop poorer phonological skills (e.g., Gilbertson & Kamhi 1995; Briscoe et al. 2001; Nittrouer & Burton 2001; Norbury et al. 2001; Gibbs 2004; Halliday & Bishop 2005; Delage & Tuller 2007). The phonological deficits in these children are widespread and involve *phoneme discrimination* (detect difference between speech sounds), *phonological working memory* (remember speech sound pattern for a few seconds), and *phonological awareness* (analyze or manipulate the sounds of speech). A qualification emphasized by these studies, however, is that performance in these children varies widely.

In distinction to the above view, a less well-established orientation proposes that phonological knowledge is not exclusively auditory in nature and can be established in non-auditory ways such as lipreading, speech production, and reading/orthography (e.g., Dodd & Campbell 1987; Kyle & Harris 2010). Evidence supporting this view is that young adults with prelingual deafness who use sign language (Deaf and Hard-of-Hearing, DHH) demonstrate phonological knowledge (e.g., Hanson & Fowler 1987; Engle et al. 1989; Hanson & McGarr 1989; Hanson et al. 1991; see Treiman & Hirsh-Pasek 1983 and Ormel et al. 2010 for exceptions). As an example (from a lexical decision task): DHH show faster word response times to written rhyming (WAVE-SAVE) than non-rhyming (HAVE-CAVE) pairs (Hanson & Fowler 1987). Also, on a silent reading task (i.e., make judgements about tongue-twister vs control sentences), DHH make more errors for the tongue-twister sentences (Hanson et al. 1991). There is the normal pattern of results, and it indicates that DHH are recoding the written words phonologically.

In addition to the evidence in DHH—and in distinction to the evidence above supporting the auditory viewpoint—children and adolescents with prelingual sensorineural hearing impairment who use aural/oral or total communication approaches (CHI) may also demonstrate phonological knowledge (Dodd & Hermelin 1977; Dodd et al. 1983; Dodd 1987; Leybaert & Alegria 1993; Sterne & Goswami 2000; see Campbell & Wright 1988, 1990 for exceptions). Evidence in this more aural/oral group is that—on a written letter cancellation task (strike out all the g's)—CHI strike out the pronounced g's (e.g., tiger) more often than the unpronounced g's (e.g., night, Dodd 1987). Again, this is the normal pattern of results, which is attributed to CHI children recoding the words phonologically. Other studies in CHI have focused on the recency effect of serial recall tasks. The recency effect occurs when the final item of a to-be-remembered list is recalled better than the mid-list items. The recency effect is prominent for auditory input, but not for visual input (Jerger & Watkins 1988). A recency effect occurs when CHI and children with normal hearing (CNH) try to recall lipread items (Dodd et al. 1983). This pattern of results indicates that lipread items act more like auditory input than visual input in memory.

Overall, these studies indicate that DHH and CHI possess an impressive degree of phonological knowledge. This finding implies that—regardless of the input route: lipreading, speech production, orthography/reading, and/or residual hearing—perceptual and cognitive processes can encode and abstract sufficient phonological knowledge to influence performance on a variety of tasks. A noteworthy qualification of the above research, however, is that CHI generally performed poorer than CNH.

Finally, another view that has also received less attention stresses the importance of visual speech in learning the phonology of spoken language (e.g., Dodd & Campbell 1987; Locke 1993; Weikum et al. 2007, Lewkowicz & Hansen-Tift 2012). From this point of view, auditory and visual speech inputs are integrated complementary essentials. This line of research assesses whether visual speech enhances auditory speech (i.e., AV vs auditory inputs). Example evidence in infants with normal hearing (infantsNH) is that visual speech improves phoneme discrimination (Teinonen et al. 2008). Example evidence in CNH is that performance is better for AV than auditory input for 1) discrimination of visually distinct phonemes (Lalonde & Holt 2015), 2) feature contrast discrimination (Hnath-Chisolm et al. 1998), and 3) vowel phoneme monitoring (Fort et al. 2012; but see Boothroyd et al. 2010, for exceptions). Some age-related variability characterizes these results. Research in CHI is scant, and the results are mixed (Jerger, Lai, et al. 2002; Eisenberg et al. 2003).

The current research was conducted within the latter school of thought: visual and auditory speech (i.e., AV vs. auditory inputs) are complementary essentials in phonological development. We assessed the influence of visual speech on phonological priming by high vs. low fidelity (intact vs. non-intact onsets) auditory speech in CHI and CNH. We use the terms “high fidelity or intact” and (respectively) “low fidelity or non-intact” as synonyms. We selected phonological priming because priming is an *indirect* task that assesses the quality of stored phonological knowledge without requiring children to *directly* access and retrieve their knowledge and formulate a response. We selected non-intact auditory onsets because visual speech is more beneficial to individuals when they process low fidelity speech. A current research question is whether the fidelity of speech matters more to CNH than CHI whose impaired ears degrade all auditory input to a lower fidelity. Below we detail our new stimuli, the phonological priming task, our research questions, and predicted results.

New Distractors: Non-Intact Auditory Onsets

The new stimuli are words and nonwords with an intact consonant/rhyme in the visual track coupled to a non-intact onset/rhyme in the auditory track (our methodological criterion excised—from the auditory onsets—about 50 ms for words and 65 ms for nonwords, see Methods). Stimuli are presented as AV vs. auditory input. Example stimuli for the word *bag* are: 1) AV: intact visual (b/ag) coupled to non-intact auditory (**–***b*/ag) and 2) Auditory: static face coupled to the same non-intact auditory (**–***b*/ag). Our question was whether the intact visual speech would “restore or fill-in” the non-intact auditory speech; in which case, performance for the *same* auditory stimulus *would differ* depending upon the presence/absence of visual speech. Responses illustrating this influence of visual speech on a repetition task (Jerger et al. 2014) are perceiving /bag/ for AV input but /ag/ for auditory input. To study the influence of visual speech on phonological priming, these stimuli were administered via the multi-modal picture word (PW) task.

Phonological Priming: Multi-Modal PW Task.

In the original “cross-modal” PW task (Schriefers et al. 1990), participants name pictures while attempting to ignore nominally irrelevant auditory speech distractors. To study phonological priming, the relation between the [picture]–[distractor] onsets is either congruent (priming condition: e.g., [bug]–[bus]) or neutral (baseline vowel-onset condition: e.g., [bug]–[onion]). The dependent measure is picture naming times and the congruent condition—relative to the baseline condition—speeds up or primes picture naming (Jerger, Martin, et al. 2002; Jerger et al. 2009). The congruent onset is thought to prime picture naming because of crosstalk between the phonological representations that support speech perception and production (Levelt et al. 1991). Congruent distractors activate input phonological representations whose activation spreads to corresponding output phonological representations, which speeds the selection of these speech segments for naming (Roelofs 1997). In our “multi-modal” PW task (Jerger et al. 2009), the speech distractors are presented AV or auditory only (see methods), a manipulation that enables us to study the influence of visual speech on phonological priming.

Orientation and Research Questions.

The literature reviewed above indicates that the development of phonological knowledge in CHI may involve not only residual hearing but also non-auditory factors. An interdependence between auditory and visual speech in phonological development is widely accepted for CHI (Woodhouse et al. 2009) yet the influence of visual speech on phonological processing remains understudied in CHI. Our research addressed two questions in two separate analyses: Analysis 1) Is phonological priming in CHI vs. CNH differentially affected by the characteristics of the stimuli (i.e., mode, fidelity, and lexical status)? Analysis 2) In CHI, is the influence of visual speech on phonological priming uniquely affected by the degree of HI, auditory word recognition, and/or age?

Analysis 1: Predicted Results

Mode and Fidelity. Research indicates that the relative weighting of auditory and visual speech is modulated by the relative quality of each input. To illustrate: when responding to McGurk stimuli with incongruent visual and auditory inputs (visual aka; auditory apa), CHI who are cochlear implant users listening to ear-degraded speech and CNH who are listening to experimentally-degraded auditory speech respond more on the basis of the intact visual input (Huyse et al. 2013). When listening to conflicting inputs such as auditory 'meat' coupled with visual 'street,' CNH and CHI with good auditory word recognition respond on the basis of the auditory input (Seewald et al. 1985). In contrast, CHI with more severe impairment—and more degraded perception of auditory input—respond more on the basis of the visual input. To the extent that these data obtained with conflicting auditory and visual inputs apply to our study, these results predict that both CNH and CHI may weight the intact visual speech more heavily than the non-intact auditory speech—and this should produce a significant influence of visual speech for our low fidelity auditory speech in both groups. When this "fidelity" effect (with its relatively greater weighting of the intact visual speech) is coupled with the relatively greater weighting of the phonetic-phonological content for nonwords (Mattys et al. 2005, see immediately below), a significantly greater influence of visual speech will be observed for nonwords than words.

Words vs. Nonwords (e.g., Bag vs. Baz). The hierarchical model of speech segmentation (Mattys et al. 2005) proposes that listeners assign the greatest weight to 1) lexical-semantic content when listening to words and 2) phonetic-phonological content when listening to nonwords. It is also assumed that familiar monosyllabic words such as our stimuli (bag) may activate their lexical representations without requiring phonological decomposition whereas nonwords (baz) require phonological decomposition (Mattys 2014;

see also Morton 1982). If these ideas generalize to our task, word stimuli should be heavily weighted in terms of lexical-semantic content whereas nonword stimuli should be heavily weighted in terms of phonetic-phonological content for both AV and auditory speech. A greater weight on phonetics-phonology for the nonwords should increase children's awareness of the phonetic-phonological content and produce greater priming for the nonwords.

Previous Results on the PW Task. Previous results in CHI are perplexing: Intact congruent phonological distractors produced significant priming on the cross-modal PW task with nonsense syllable distractors—e.g., [picture]–[distractor]: [pizza]–[pi] (Jerger, Martin, et al., 2002)—but not on the multi-modal PW task with word distractors—e.g., [picture]–[distractor]: [pizza]–[peach] (Jerger et al., 2009). These results predict that the intact auditory nonword distractors, but not the intact word distractors, will significantly prime picture naming in CHI. Previous results in CNH (Jerger et al. 2016) on the multi-modal PW task with the new distractors of this study showed significant priming by all auditory and AV distractors (with a minor exception). Further, overall picture naming times showed greater priming from the intact than non-intact distractors and from the AV than auditory distractors. The addition of visual speech boosted priming significantly more for the non-intact nonwords than non-intact words. If these results in CNH generalize to CHI, we predict that all of the AV and auditory distractors, but particularly the intact onsets, will significantly prime naming. Further, we predict significant greater priming by the AV distractors, particularly for the non-intact nonwords.

Analysis 2: Predicted Results

We analyzed the influence of visual speech as a *unique* function of degree of HI, auditory word recognition, and age. We defined “unique” statistically as the independent contribution of each variable after controlling for the other variables (Abdi et al. 2009). Below we focus on a few relevant points about these child factors.

Degree of HI, Suprathreshold Auditory Word Recognition, and Age. Sensorineural HI impairs not only threshold hearing sensitivity, but also the processing of suprathreshold sounds (e.g., because of reduced spectral and temporal resolution, Moore 1995). The outcome of these dual impairments is that suprathreshold auditory word recognition in CHI with the same threshold hearing sensitivity loss can vary widely (Erber 1974). These findings indicate that the influence of visual speech may be more readily predicted by auditory word recognition (a measure of functional hearing status taking into account suprathreshold distortions) than by degree of HI (a measure of threshold impairment for pure tones). A previous study in cochlear implant users supports this idea, showing that CHI who are better at recognizing auditory words benefit more from visual speech (Lachs et al. 2001). With regard to age, previous results on the multi-modal PW task with the new distractors in CNH (Jerger et al. 2016) showed one significant age effect, namely greater overall priming in younger than older children. By contrast, visual speech influenced performance in all CNH from 4– to 14–years, with greater priming by the AV than auditory distractors. If these results in CNH generalize to CHI, we predict a significant visual speech effect at all ages, with significantly greater overall priming in younger CHI.

Recapitulation

The evidence reviewed above establishes that perceiving everyday speech is an AV event with visual speech serving as an important source of phonological knowledge. Despite this backdrop, phonology and speech perception are typically studied only in the auditory mode—even in CHI. The current research will provide critical new information about how visual speech contributes to phonological priming in CHI. Such data could have clinical implications for current intervention programs that emphasize listening in the auditory mode only as the optimal approach for developing spoken language (e.g., Lew et al. 2014).

Methods

Participants

Groups. Participants were 62 CHI with early-onset sensorineural loss (55% girls) and 62 CNH (55% girls). The CNH group—whose ages yielded a group-mean and -distribution akin to that in the CHI group—was formed from a pool of 132 CNH from associated projects (see Jerger et al. 2014; Jerger et al. 2016). Ages (yr;mo) ranged from 4;1 to 14;9 (*M* = 9;2) in the CHI and 4;2 to 14;5 (*M* = 9;1) in the CNH. The racial distributions were 73% Whites, 22% Blacks, and 5% Asian in CHI and 82% Whites, 6% Blacks, 8% Asian, and 4% Multiracial in CNH, with 9% of CNH reporting Hispanic ethnicity. All participants met the following criteria: 1) English as a native language, 2) ability to communicate successfully aurally/orally, and 3) no diagnosed or suspected disabilities other than HI and its accompanying speech and language problems.

Audiological Characteristics. Hearing sensitivity in the CNH at hearing levels (HLs) of 500, 1000, and 2000 Hz (pure-tone average, PTA; American National Standards Institute, ANSI 2004) averaged 2.83 dB HL (*SD* = 4.66; right ear) and 3.68 dB HL (*SD* = 5.35; left ear). The PTAs in the CHI averaged 49.80 dB HL (*SD* = 21.43; right ear) and 52.39 dB HL (*SD* = 24.34; left ear). Average PTAs on the two ears were distributed as follows: ≤20 dB (5%), 21 to 40 dB (27%), 41 to 60 dB (36%), 61 to 80 dB (27%), and 81 to 102 dB (5%). The children with PTAs of ≤20 dB had losses in restricted frequency regions. Hearing aids were used by 90% of the children. Participants who wore amplification were tested while wearing their devices. The estimated age at which the children who wore amplification received their first device averaged 2.79 yrs (*SD* = 2.09); the estimated duration of device use averaged 6.17 yrs (*SD* = 3.21). Forty-six children were mainstreamed in a public school setting and 16 children were enrolled in an aural/oral school for CHI.

Comparison of Groups. Table 1 compares results in the CNH and CHI on a set of nonverbal and verbal measures. The nonverbal and verbal measures were analyzed with mixed-design analyses of variance with one between-participants factor (Group: CNH vs. CHI) and one within-participants factor (Measure). The nonverbal measures were standardized scores for simple auditory reaction time (RT), simple visual RT, and visual perception. The verbal measures were standardized scores for receptive vocabulary, expressive vocabulary, articulation, phonological awareness, auditory word recognition, and lipreading.

The nonverbal results indicated a significant difference between groups and a significant measures x groups interaction, respectively F (1, 122) = 23.74, MSE = 0.178, p < .0001, partial η^2 = .163 and F (2, 244)

= 11.36, *MSE* = 0.438, $p < .0001$, partial η^2 = .085. Multiple *t*-tests with the problem of multiple comparisons controlled with the False Discovery Rate (FDR) procedure (Benjamini & Hochberg 1995) indicated that the speed at which the CNH and the CHI could detect auditory and visual inputs (simple RT) did not differ. By contrast, visual perception was significantly better in the CNH than the CHI. The reasons for this latter difference are unclear. Performance for both groups, however, was within the average normal range.

The verbal results indicated a significant difference between groups and a significant measure x group interaction, respectively F (1, 122) = 25.48, *MSE* = 1.497, $p < .0001$, partial η^2 = .173 and F (5, 610) = 21.24, *MSE* = 0.721, $p < .0001$, partial η^2 = .148. Multiple *t*-tests controlled with the FDR correction indicated that auditory word recognition, phonological awareness, articulation proficiency, and receptive and expressive vocabulary were significantly better in the CNH whereas lipreading onsets did not differ between the groups. In short, the CHI vs. CNH showed mostly similarities in their non-verbal skills and differences in their verbal skills as expected. Interestingly—to foreshadow our results—these significant differences in verbal skills observed on direct tasks (i.e., children were informed about the targets of interest and were consciously responding during a post-stimulus interval) did not produce significant differences in phonological priming on our indirect PW task (i.e., children were not informed about nor asked to attend to or consciously respond to the distractors, our targets of interest; see Jordan & Bevan 1997 for similar results between direct vs. indirect measures for lipreading). Finally, we should acknowledge that vocabulary scores in the CNH indicated higher than average vocabulary knowledge, and that such performance could potentially affect the generalizability of these results to CNH with more ‘average’ vocabulary abilities.

Materials and Instrumentation: PW Task

Pictures and Distractors. The pictures and word/nonword distractors consisted of experimental items (8 pictures beginning with /b/ or /g/ and 12 distractors beginning with /b/ or /g/ or a vowel) and filler items (16 pictures and 16 distractors beginning with consonants other than /b/ or /g/; see Supplemental Appendix A for items and see Jerger et al. 2016 for details). As noted above, the relation between the experimental [picture]–[distractor] onsets was manipulated to be congruent (priming condition: e.g., [bug]–[bus/buhl]) or neutral (baseline vowel onset condition: e.g., [bug]–[onion/onyit]). The filler items consisted of [picture]–[distractor] pairs *NOT* sharing an onset and *NOT* beginning with /b/ or /g/: e.g., [cookies]–[horse/hork]. To ensure that the subsequent results were reflecting performance for words vs. nonwords, the participants' understanding of the meaning of the experimental word distractors was assessed (see Supplemental Appendix B).

Stimulus Preparation. The distractors were recorded by an 11-yr-old boy actor. He started and ended each utterance with a neutral face and closed mouth. The color video signal was digitized at 30 frames/s with 24-bit resolution at a 720 × 480 pixel size. The auditory signal was digitized at 48 kHz sampling rate with 16-bit amplitude resolution and adjusted to equivalent A-weighted root mean square sound levels. The video track was routed to a high resolution monitor, and the auditory track was routed through a speech audiometer to a loudspeaker. The intensity level of the distractors was approximately 70 dB SPL. The to-be-named colored pictures were scanned into a computer as 8-bit PICT files (see Jerger et al. 2016 for details).

Editing the Auditory Onsets. We edited the auditory track of the phonologically-related distracters by locating the /b/ or /g/ onsets visually and auditorily with Adobe Premiere Pro and Soundbooth (Adobe Systems Inc., San Jose, CA) and loudspeakers. We applied a perceptual criterion to operationally define a non-intact onset. We excised the waveform in 1 ms steps from the identified auditory onset to the point in the later waveforms for which at least 4 of 5 trained listeners heard the vowel—not the consonant—as the onset in the auditory mode. Splice points were always at zero axis crossings. Using this perceptual criterion, we excised on average 52 ms (/b/) and 50 ms (/g/) from the word onsets and 63 ms (/b/) and 72 ms (/g/) from the nonword onsets (see Jerger et al. 2016 for details).

All stimuli were presented as Quicktime movie files, and we next formed AV (dynamic face) and auditory (static face) presentations. In our experimental design, we compare results for the auditory vs. AV non-intact stimuli. Any coarticulatory cues in the auditory input are held constant in the two modes; thus any influence on picture naming due to coarticulatory cues should be controlled and allow us to evaluate whether the addition of visual speech influences performance.

AV and Auditory Inputs. For the AV input, the children saw: 1) 924 ms (experimental trials) or 627 or 1,221 ms (filler-item trials) of the talker's still face and upper chest, followed by 2) an AV utterance of a distractor and presentation of a picture on the talker's T-shirt 165 ms before the auditory onset of the utterance (auditory distractor lags picture), followed by 3) 924 ms of still face and picture. For the auditory mode, the child heard the same event with the video track edited to contain only the talker's still face and upper chest. The onset of the picture occurred in the same frame for the intact and non-intact distracters. The relation between the onsets of the picture and the distractor—called stimulus onset asynchrony (SOA)—is also a key consideration for the PW task.

SOA. Phonologically-related distracters typically produce a maximal effect on naming when the auditory onset of the distractor lags the onset of the picture by about 150 ms (Damian & Martin 1999; Schriefers et al. 1990). Our SOA was 165 ms as used previously (Jerger et al. 2009). We acknowledge, however, that the non-intact auditory onset altered the target SOA of 165 ms. Our experimental design should control for any performance differences produced by the altered SOA, however, because we compare results for the auditory vs. AV non-intact stimuli, with the auditory onset held constant in the two modes. Thus any influence on picture naming produced by the shift in the auditory onset is controlled and this allows us to evaluate whether the addition of visual speech influences performance.

We also acknowledge that the non-intact onset slightly alters the normal temporal asynchrony between the auditory and visual speech onsets. Previous evidence suggests that this shift in the auditory onset of the non-intact stimuli will not affect our results. Specifically, AV speech is normally asynchronous, with visual speech leading auditory speech by variable amounts (Bell-Berti & Harris 1981; ten Oever et al. 2013). Adults with NH synthesize visual and auditory speech—without any detection of the asynchrony and without any effect on intelligibility—even when visual speech leads by as much as 200 ms (Grant et al. 2004). CNH of 10–11-yrs perform like adults for AV asynchrony when visual speech leads (Hillock et al. 2011), but CNH of 4–6-yrs have a wider window and do not detect AV asynchrony until visual speech leads by more than 366 ms (Pons et al. 2013). Below we summarize our final set of materials.

Final Set of Items. Two presentations of each auditory and AV experimental item (i.e., baseline and intact vs. non-intact phonetically-related distractors) were randomly intermixed with the filler items and formed into four lists (presented forward or backward). Each list contained 57% experimental and 43% filler items. The items of a list varied randomly under constraints such as 1) no onset could repeat, 2) the intact and non-intact pairs (e.g., bag and /–*b*/ag or vice versa) could not occur without at least two intervening items, 3) one-half of items must occur first AV and one-half first auditory, and 4) all types of onsets (vowel, intact /b/ and /g/, non-intact /b/ and /g/, and not /b/ or /g/) must be dispersed uniformly throughout the lists (see Jerger et al. 2016 for details).

Naming Responses. To quantify picture naming times, the computer triggered a counter/timer with better than 1 ms resolution at the initiation of each movie file. The timer was stopped by the onset of the naming response into a unidirectional microphone which was fed through a stereo mixing console amplifier and attenuator to a voice-operated relay (VOR). A pulse from the VOR stopped the timing board via a data module board. The counter timer values were corrected for the amount of silence in each movie file before the onset of the picture.

Procedure

The children completed the multi-modal PW task along with other procedures in three sessions occurring 1) on three days for the CNH and 2) on one (16%), two (40%), or three (44%) days for the CHI. The interval between testing days averaged 12 days in each group. The order of presentation of the word vs. nonword conditions was counterbalanced across participants in each group. Results were collapsed across the counterbalancing conditions. In the first session, the children completed three of the word (or nonword) lists; in the second session, the children completed the fourth word (or nonword) list and the first nonword (or word) list; and in the third session, the children completed the remaining three nonword (or word) lists. A variable number of practice trials introduced each list.

At the start of the first session, a tester showed each picture on a card, asking children to name the picture and teaching the target names of any pictures named incorrectly. Next the tester flashed the picture cards quickly and modeled speeded naming. Speeded naming practice trials went back and forth between the tester and child until the child was naming the pictures fluently. Mini-practice trials started the other sessions.

For formal testing, a tester sat at a computer workstation and initiated each trial by pressing a touch pad (out of child's sight). The children, usually with a co-tester alongside, sat at a distance of 71 cm in front of an adjustable height table containing the computer monitor and loudspeaker. Trials that the tester or co-tester judged flawed were deleted online and re-administered after intervening items. The children were told they would see and hear a boy whose mouth would sometimes be moving and sometimes not. For the words, participants were told that they might hear words or nonwords because some of the non-intact words (e.g., /–B/ag) may be perceived as nonwords (e.g., ag), especially in the auditory mode. For the nonwords, participants were told that they would always hear nonwords. Participants were told to focus on 1) watching for the picture that would pop up on the boy's T-shirt and 2) naming it as quickly and as accurately as possible. The participant's view of the picture subtended a visual angle of 5.65° vertically and 10.25° horizontally; the view of the talker's face subtended a visual angle of 7.17° vertically (eyebrow – chin) and 10.71° horizontally (at eye level). This research protocol was approved by the Institutional Review Boards of University of Texas at Dallas and Washington University St. Louis.

Results

Preliminary Analyses

The quality of the PW data (e.g., number of missing trials) is detailed in Supplemental Appendix B. Initially, we also analyzed the PW data to determine whether results could be collapsed across the distractor onsets (/b/ vs. /g/; see Supplemental Appendix C for results). Briefly, separate factorial mixed-design analyses of variance were performed for the baseline distractors and the phonologically-related distractors. Findings indicated that the different onsets influenced results only for the phonologically-related distractors; *overall* picture naming times were facilitated slightly more for the /b/ than /g/ onset (–142 vs. –110 ms). Despite this statistically significant outcome, the difference in performance due to onset was minimal (32 ms), and no other significant onset effects were observed. These results agree with our previous findings (Jerger et al. 2016); phonological priming by /b/ and /g/ onsets does not show the pronounced differences that characterize identifying these phonemes on direct measures of speechreading (see Jordan & Bevan 1997 for similar results). Thus, for the analyses below, naming times were collapsed across the onsets to promote clarity and accessibility.

Analysis 1: Lexical Status, Mode, and Fidelity

Baseline Distractors

Figure 1 shows average picture naming times in the groups for the auditory vs. AV word (left) and nonword (right) baseline distractors. Results were analyzed with a factorial mixed-design analysis of variance with one between-participants factor (Group: CHI vs. CNH) and two within-participants factors (Lexical Status: words vs. nonwords and Mode: auditory vs. AV. Table 2 summarizes the results (significant results are bolded). Overall picture naming times were significantly faster for the CNH than the CHI, respectively 1372 vs 1591 ms. Overall naming times were also significantly faster for 1) auditory than AV speech (1473 ms vs. 1489 ms) and 2) nonwords than words (1464 ms vs. 1499 ms), but these numerical differences were notably small (16 ms and 35 ms). No other significant effect was observed.

The above difference in the baseline times between groups was large enough (about 220 ms) that it could be problematic for our traditional approach of quantifying phonological priming with adjusted naming times (derived by subtracting each participant's baseline naming times from his or her phonologically-related naming times; Jerger, Lai, et al. 2002; Jerger et al. 2009). The different baselines could muddle an unequivocal interpretation of any group differences. To control the differences in baseline performance, we quantified priming proportions—derived by dividing each participant's adjusted naming times by his or her corresponding baseline naming times (see Damian & Dumay 2007]. Greater priming is indicated by a larger negative proportion.

Phonologically-Related Distractors

Figure 2 depicts average priming proportions in the CHI vs. CNH for the high vs low fidelity stimuli presented as auditory and AV speech (left vs. right panels). Results are presented for the words (Figure 2a) and nonwords (Figure 2b). An initial multifactorial analysis with all factors yielded a complex higher order interaction (Lexical Status × Fidelity × Mode, elaborated at the end of the results section). Thus—before considering the effects of lexical status—results were analyzed separately for the words and nonwords with a factorial mixed-design analysis of variance with one between-participants factor (Group: CHI vs CNH) and two within-participant factors (Fidelity: low vs high; Mode: auditory vs AV). Table 3 summarizes the results.

Words (Figure 2a). Overall priming for the words did not differ in the CNH vs. CHI. The other main factors, however, did significantly influence the phonological priming of picture naming, with a significant effect of *fidelity* and *mode*. Both the CNH and CHI showed greater priming by the intact than non-intact distractors (respectively –.074 vs. –.050 when collapsed across mode). Both the CNH and CHI also showed greater priming from the AV than auditory distractors (–.075 vs. –.049 when collapsed across fidelity). This latter result is particularly relevant because this pattern highlights a significant benefit of visual speech on performance for all children. No other significant effect was observed. Results in the CNH and CHI did not differ for the word stimuli.

Nonwords (Figure 2b). Overall priming for the CHI and CNH again did not differ. The other main factors that significantly influenced phonological priming were: 1) *fidelity*, showing greater priming by the intact than non-intact distractors (respectively –.111 vs. –.081 when collapsed across mode) and 2) *mode*, showing greater priming by the AV than auditory distractors (–.114 vs. –.078 when collapsed across fidelity). As seen in Table 3, however, significant interactions between and among all possible factors complicated a simple account of these main effects. To clarify the interactions, we reanalyzed results for each mode separately with a factorial mixed-design analysis of variance with one between-participants factor (Group: CHI vs. CNH) and one within-participant factors (Fidelity: high vs. low). The results are summarized at the bottom of Table 3.

The separate analyses revealed that the previously noted Group × Fidelity × Mode interaction occurred because of the differences in the pattern of results for the two modes. For the AV input, the priming of picture naming in the CNH and CHI did not differ. The only significant effect was that the intact distractors produced greater priming than the non-intact distractors (respectively –.122 vs –.106). For the auditory mode, the fidelity of the stimuli again affected priming, but this time with a fidelity × group interaction. Priming by the intact vs non-intact auditory distractors differed significantly more in the CNH (respectively –.099 vs –.037) than in the CHI (–.100 vs –.076). Stated differently high fidelity auditory input primed picture naming similarly in the CHI and CNH (–.100 vs –.099) whereas low fidelity auditory input primed picture naming to a greater extent in CHI than in CNH (–.076 vs –.037). Priming in the CHI was less affected by the fidelity of the auditory input. Next, we determined whether phonological priming was greater for the nonwords than words as predicted.

Lexical Status. The initial multifactorial analysis with all factors can be used to address the effects of lexical status. This analysis indicated significantly greater overall priming for the nonwords than the words, respectively –.096 vs –.062, $F(1,122) = 42.62$, $MSE = .007$, $p < .0001$, partial $\eta^2 = .259$. As noted above, however, a significant complex interaction was also observed, Lexical Status × Fidelity × Mode: $F(1,122) = 4.06$, $MSE = .002$, $p < .040$, partial $\eta^2 = .034$. Although we did not have sufficient statistical power to detect the higher fourth order interaction (Lexical Status × Fidelity × Mode × Group), the significant third order interaction is most easily understood by the above differences for mode and fidelity that were observed for the nonwords but not for the words (and which also involved the group). No other significant effects or interactions for lexical status were observed. In addition to these results addressing differences between the groups, it is also relevant to consider whether priming within each group differed for auditory vs. AV speech.

Did the Addition of Visual Speech Produce Significantly Greater Priming? To evaluate whether priming by auditory vs. AV speech differed significantly for each condition in each group, we carried out planned orthogonal contrasts (Abdi & Williams 2010). Table 4 summarizes the results. The CNH showed a significant benefit from visual speech—subsequently called the visual speech effect (VSPE)—for both words and nonwords in the high and low fidelity conditions. The CHI showed a significant VSPE for both conditions for the words, but only for the low fidelity condition for the nonwords. That said, results for the high fidelity nonwords in the CHI approached significance. Next we consider whether/how the results were influenced by the children's hearing status and age.

Analysis 2: Unique Effect of Degree of HI, Auditory Word Recognition, and Age in CHI

To understand the unique effects of degree of HI, auditory word recognition, and age on phonological priming, we carried out separate multiple regression analyses for the words and nonwords. We defined "uniquely" statistically as expressed by part correlations, which reveal the independent contribution of a variable after controlling for all the other variables (Abdi et al. 2009). We did not include fidelity in these analyses because stimulus intactness influenced results only minimally in the CHI (for high vs low fidelity respectively, –.087 vs –.069 for CHI and –.098 vs –.062 for CNH). The degree of HI was quantified by the average four-frequency [500–4000 Hz] PTA of the two ears. The dependent variable was the priming proportions; the independent variables were the standardized scores for auditory word recognition, degree of HI, and age. The intercorrelations among this set of variables were as follows: 1) Auditory Word Recognition vs. degree of HI (.271) and age (.194), and 2) Age vs. degree of HI (.113). Results of the multiple regression analysis indicated that degree of HI, auditory word recognition, and age did not influence priming by the word stimuli. Thus only results for the nonwords are elaborated.

Table 5 shows the slopes, part correlation coefficients, and partial *F* statistics evaluating the variation in nonword performance uniquely accounted for by each individual variable (Abdi et al. 2009). The multiple correlation coefficients for all of the variables considered simultaneously are provided below the Table for interested readers. The part correlations indicated that only auditory word recognition uniquely impacted results and only for AV speech. In short—with age and degree of HI controlled—auditory word recognition influenced priming when visual speech was added to the input.

Discussion

This research assessed the influence of visual speech on phonological priming by high (intact) vs. low (non-intact) fidelity auditory speech in CHI. The low fidelity stimuli were words and nonwords with an intact visual consonant + rhyme coupled to a non-intact auditory onset + rhyme. Our research protocol investigated whether phonological priming was 1) differentially affected in CHI vs. CNH by the fidelity (high vs. low), mode (AV vs. auditory), and lexical status (words vs. nonwords) of the stimuli, and 2) uniquely affected by degree of HI, auditory word recognition, and/or age in CHI. Below we consider these issues.

Did Fidelity and Mode Differentially Affect Phonological Priming in CHI vs. CNH?

For words as well as nonwords, both CHI and CNH consistently showed greater phonological priming from high than low fidelity input (when collapsed across mode) and from AV than auditory speech (when collapsed across fidelity). The latter result is particularly relevant because the pattern highlights a significant VSPE in these children. When the VSPE was examined as a function of fidelity, we found many similarities in the CHI and CNH. Both groups showed a significant VSPE for the high and low fidelity words and for the low fidelity nonwords. A seeming contrast between the groups was that CNH showed a significant VSPE for high fidelity nonwords whereas CHI did not; however, the latter result in CHI clearly showed a numerical difference that approached statistical significance.

Results for the VSPE underscore the importance of visual and auditory speech as complementary resources for phonological development in both CHI and CNH. This proposal is bolstered by the findings of delayed and/or different phonology and early expressive language skills in individuals with early-onset blindness (e.g., McConachie & Moore 1994; Menard et al. 2013, Mills 1987). Our children participants—like adults—perceive speech by eye *and* ear. The VSPE findings clearly endorse visual speech as a vital resource for learning spoken language.

Finally, with regard to effects of hearing impairment, results indicate that the priming of picture naming by congruent phonological distractors did not generally differ in CHI vs. CNH. Thus these CHI had sufficiently well specified phonological onset representations to support priming. A possible qualification is that all of the current CHI communicated successfully aurally/orally (see Method).

Did Lexical Status Differentially Affect Phonological Priming in CHI vs. CNH?

Phonological priming in both the CHI and CNH was significantly greater for nonwords than words. This result is consistent with our predictions that children will weigh phonetic-phonological content more heavily when listening to nonwords, and this increased weighting will increase priming (Mattys et al. 2005). This overall pattern of similarity between the groups was qualified, however, by the finding that phonological priming in the CHI vs. CNH differed for nonwords but not for words. The difference occurred because the fidelity of the auditory input affected phonological priming by unfamiliar nonwords more in CNH than CHI. Apparently, CHI can benefit from unfamiliar low fidelity auditory input more than CNH, perhaps because CHI regularly experience ear-degraded auditory speech.

Did Degree of HI, Auditory Word Recognition, and Age Affect Phonological Priming in CHI?

We analyzed the unique contribution of each of these individual variables with the effects of the other variables controlled. Neither the degree of HI nor age uniquely affected phonological priming in CHI. By contrast, the auditory word recognition skills of the CHI did uniquely influence phonological priming of the nonwords by AV speech. As before, we hypothesize that the nonwords were affected because 1) listeners assign the greatest weight to phonetic-phonological content for nonwords and 2) the processing of nonwords requires phonological decomposition (Mattys et al. 2005; Mattys 2014; see also Morton 1982). Evidence from the literature—illuminating the link between auditory word perception by CHI and the priming of picture naming by AV speech—is discussed below.

A close bidirectional link between the developing speech production and speech perception systems is proposed by the Native Language Magnet Theory Expanded (Kuhl et al. 2008). This link between speech production/visual speech and auditory speech perception is clearly supported by previous research in infantsNH, such as Kushnerenko et al. (2013) who demonstrated a significant relation between looking-time patterns to AV speech—to the eyes vs. the mouth—at 6–9 months and auditory speech comprehension at 14–16 months. The evidence in CNH (e.g., Erdener & Burnham 2013) also supports this linkage by showing that perceptual tuning to the phoneme contrasts of the native—as opposed to non-native—language predicts AV speech perception. Stated differently from the viewpoint of this discussion, the Erdener and Burnham data also seem to support the idea that the more highly tuned children are to visual speech, the better they learn the phoneme contrasts of the ambient language and hence the better they learn words. This idea agrees with studies showing that visual speech improves feature contrast discrimination, phoneme monitoring, and/or phoneme discrimination in infantsNH and CNH (Fort et al. 2012; Hnath-Chisolm et al. 1998; Lalonde & Holt 2015; Teinonen et al. 2008).

Overall, results suggest that greater sensitivity to visual speech yields better phoneme discrimination and hence better word comprehension, which supports our finding of a relation between phonological priming by AV speech and auditory word recognition. This type of linkage— between benefit from visual speech and auditory word recognition—has been observed previously in CHI who use cochlear implants

(Lachs et al. 2001). Visual speech clearly seems a vital *enriching* complement to auditory speech for developing phonological knowledge, particularly for CHI.

Conclusions

With minor exceptions, phonological priming in CHI and CNH showed more similarities than differences. Importantly, the addition of visual speech significantly increased phonological priming in both groups. Auditory word recognition also significantly impacted the influence of visual speech on phonological priming. Clinically these data support intervention programs that view visual speech as a powerful asset for developing spoken language in CHI.

References

Abdi, H., Edelman, B., Valentin, D., Dowling, W. (2009). *Experimental design and analysis for psychology*. New York: Oxford University Press.

Abdi, H., & Williams, L. (2010). Contrast analysis. In N. Salkind (Ed.), *Encyclopedia of research design* (pp. 243-251). Thousand Oaks, CA: Sage.

American National Standards Institute (ANSI). (2004). *Specifications for audiometers (ANSI S3.6-2004)*. New York: Author.

Beery, K., & Beery, N. (2004). *The Beery-Buktenica developmental test of visual-motor integration with supplemental developmental tests of visual perception and motor coordination.* (5th ed.). Minneapolis: NCS Pearson, Inc.

Bell-Berti, F., & Harris, K. (1981). A temporal model of speech production. *Phonetica, 38*, 9-20.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol, 57*, 289-300.

Boothroyd, A., Eisenberg, L., Martinez, A. (2010). An on-line imitative test of speech-pattern contrast perception (OlimSpac): Developmental effects in normally hearing children. *J Speech Lang Hear Res, 53*, 531-542.)

Briscoe, J., Bishop, D., Norbury, C. (2001). Phonological processing, language, and literacy: A comparison of children with mild-to-moderate sensorineural hearing loss and those with specific language impairment. *J Child Psychol Psychiatry, 42*, 329-340.

Brownell, R. (2000). *Expressive one-word picture vocabulary test* (Third ed.). Novato, CA: Academic Therapy Publications.

Campbell, R., & Wright, H. (1988). Deafness, spelling and rhyme - how spelling supports written word and picture rhyming skills in deaf subjects. *Q J Exp Psychol A, 40*(4), 771-788.

Campbell, R., & Wright, H. (1990). Deafness and immediate memory for pictures - dissociations between inner speech and the inner ear. *J Exp Child Psychol, 50*(2), 259-286.

Cumming, G., & Finch, S. (2005). Inference by eye: Confidence intervals and how to read pictures of data. *Amer Psychol, 60*, 170-180.

Damian, M., & Dumay, N. (2007). Time pressure and phonological advance planning in spoken production. *J Mem Lang, 57*, 195-209.

Damian, M., & Martin, R. (1999). Semantic and phonological codes interact in single word production.*J Exp Psychol Learn Mem Cogn, 25*, 345-361.

Delage, H., & Tuller, L. (2007). Language development and mild-to-mode rate hearing loss: Does language normalize with age? *J Speech Lang Hear Res, 50*, 1300-1313.

Dodd, B. (1979). Lip reading in infants: Attention to speech presented in- and out-of-synchrony. *Cogn Psychol, 11*, 478-484.

Dodd, B. (1987). Lip-reading, phonological coding, and deafness. In B. Dodd & R. Campbell (Eds.), *Hearing By eye: The psychology of lipreading* (pp. 177-189). London: Lawrence Erlbaum.

Dodd, B., & Campbell, R. (Eds.) (1987). *Hearing by eye: The psychology of lip-reading*. London: Lawrence Erlbaum.

Dodd, B., Crosbie, S., McIntosh, B., et al. (2003). *Pre-reading inventory of phonological awareness*. San Antonio, TX: The Psychological Corporation Ltd.

Dodd, B., & Hermelin, B. (1977). Phonological coding by the prelinguistically deaf. *Percept Psychophys, 21*, 413-417.

Dodd, B., Hobson, P., Brasher, J., et al. (1983). Deaf children's short term memory for lip-read, graphic and signed stimuli. *Br J Dev Psychol, 1*, 353-364.

Dunn, L., & Dunn, D. (2007). *The peabody picture vocabulary test-IV* (Fourth ed.). Minneapolis, MN: NCS Pearson.

Eisenberg, L. S., Martinez, A. S., Boothroyd, A. (2003). Auditory-visual and auditory-only perception of phonetic contrasts in children. *Volta Rev, 103*, 327-346.

Engle, R., Cantor, J., Turner, M. (1989). Modality effects - do they fall on deaf ears. *Q J Exp Psychol A, 41*, 273-292.

Erber, N. (1974). Pure tone thresholds and word recognition abilities of hearing impaired children. *J Speech Lang Hear Res, 17*, 194-202.

Erdener, D., & Burnham, D. (2013). The relationship between auditory-visual speech perception and language-specific speech perception at the onset of reading instruction in English-speaking children *J Exp Child Psychol, 114*, 120-138.

Fort, M., Spinelli, E., Savariaux, C., et al. (2012). Audiovisual vowel monitoring and the word superiority effect in children. *Int J Behav Dev*, *36*, 457-467.

Gibbs, S. (2004). The skills in reading shown by young children with permanent and moderate hearing impairment. *Educ Res, 46*(1), 17-27.

Gilbertson, M., & Kamhi, A. (1995). Novel word learning in children with hearing impairment. *J Speech Lang Hear Res, 38*, 630-642.

Goldman, R., & Fristoe, M. (2000). *Goldman Fristoe 2 test of articulation*. Circle Pines, MN: American Guidance Service.

Grant, K., van Wassenhove, V., Poeppel, D. (2004). Detection of auditory (cross-spectral) and auditory-visual (cross-modal) synchrony. *Speech Commun, 44*, 43-53.

Halliday, L., & Bishop, D. (2005). Frequency discrimination and literacy skills in children with mild to moderate sensorineural hearing loss. *J Speech Lang Hear Res, 48*, 1187-1203.

Hanson, V., & Fowler, C. (1987). Phonological coding in word reading: Evidence from hearing and deaf readers. *Mem Cogn, 15*, 199-207.

Hanson, V., Goodell, E., Perfetti, C. (1991). Tongue twister effects in the silent reading of hearing and deaf college students. *J Mem Lang, 30*, 319-330.

Hanson, V., & McGarr, N. (1989). Rhyme generation of deaf adults. *J Speech Hear Res, 32*, 2-11.

Hillock, A., Powers, A., Wallace, M. (2011). Binding of sights and sounds: Age-related changes in multisensory temporal processing. *Neuropsychologia, 49*(3), 461-467.

Hnath-Chisolm, T., Laipply, E., Boothroyd, A. (1998). Age-related changes on a children's test of sensory-level speech perception capacity. *J Speech Lang Hear Res, 41*, 94-106.

Huyse, A., Berthommier, F., Leybaert, J. (2013). Degradation of labial information modifies audiovisual speech perception in cochlear-implanted children. *Ear Hear, 34*, 110-121.

Jerger, S., Damian, M., Tye-Murray, N., et al. (2014). Children use visual speech to compensate for non-intact auditory speech. *J Exp Child Psychol, 126*, 295-312.

Jerger, S., Damian, M., Tye-Murray, N., Abdi, H. (2016). Children perceive speech onsets by ear and eye. *J Child Lang*,. available on CJO2016. doi:10.1017/S030500091500077X.

Jerger, S., Lai, L., Marchman, V. (2002). Picture naming by children with hearing loss: II. Effect of phonologically-related auditory distractors. *J Am Acad Audiol, 13*, 478-492.

Jerger, S., Martin, R., Damian, M. (2002). Semantic and phonological influences on picture naming by children and teenagers. *J Mem Lang, 47*, 229-249.

Jerger, S., Tye-Murray, N., Abdi, H. (2009). Role of visual speech in phonological processing by children with hearing loss. *J Speech Lang Hear Res, 52*, 412-434.

Jerger, S., & Watkins, M. (1988). Evidence of echoic memory with a multichannel cochlear prosthesis. *Ear Hear, 9*, 231-236.

Jordan, T. & Bevan, K. (1997). Seeing and hearing rotated faces: Influences of facial orientation on visual and audiovisual speech recognition. *J Exp Psychol Hum Percept Perform*, *23*, 388-403.

Klein, S., & Rapin, I. (2013). Intermittent conductive hearing loss and language development. In D. Bishop & K. Mogford (Eds.), *Language development in exceptional circumstances* (pp. 96-109). Hove, East Sussex, UK: Psychology Press, Taylor & Francis.

Kuhl, P., Conboy, B., Coffey-Corina, S., et al. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philos Trans R Soc Lond B Biol Sci*, 363, 979-1000.

Kushnerenko, E., Tomalski, P., Ballieux, H., et al. (2013). Brain responses and looking behavior during audiovisual speech integration in infants predict auditory speech comprehension in the second year of life.

Front Psychol 4, doi: 10.3389/fpsyg.2013.00432

Kyle, F., & Harris, M. (2010). Predictors of reading development in deaf children: A 3-year longitudinal study. *J Exp Child Psychol, 107*, 229-243.

Lachs, L., Pisoni, D., Kirk, K. (2001). Use of audiovisual information in speech perception by prelingually deaf children with cochlear implants: A first report. *Ear Hear, 22*, 236-251.

Lalonde, K., & Holt, R. (2015). Preschoolers Benefit From Visually Salient Speech Cues. *J Speech Lang Hear Res, 58*, 135-150.

Levelt, W., Schriefers, H., Vorberg, D., et al. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychol Rev, 98*, 122-142.

Lew, J., Purcell, A., Doble, M., et al. (2014). Hear here: Children with hearing loss learn words by listening. *Int J Pediatr Otorhinolaryngol, 78*, 1716-1725.

Lewkowicz, D., & Hansen-Tift, A. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proc Natl Acad Sci U S A, 109*, 1431-1436.

Leybaert, J., & Alegria, J. (1993). Is word processing involuntary in deaf children. *Br J Dev Psychol, 11*, 1-29.

Locke, J. (1993). *The child's path to spoken language*. Cambridge, MA: Harvard University Press.

MacLeod, A., & Summerfield, Q. (1987). Quantifying the contribution of vision to speech perception in noise. *Br J Audiol*, 21, 131-141.

Mattys, S. (2014). Speech perception. In D. Reisberg (Ed.), *The Oxford handbook of cognitive psychology* (pp. 391-411). Oxford, UK: Oxford University Press.

Mattys, S. L., White, L., Melhorn, J. F. (2005). Integration of multiple speech segmentation cues: A hierarchical framework. *J Exp Psychol-Gen, 134*, 477-500.

McConachie, H., & Moore, V. (1994). Early expressive language of severely visually impaired children. *Dev Med Child Neurol, 36*, 230-240.

Menard, L., Toupin, C., Baum, S. et al. (2013). Acoustic and articulatory analysis of French vowels produced by congenitally blind adults and sighted adults. *J Acoust Soc Am, 134*, 2975-2987.

Mills, A. (1987). The development of phonology in the blind child. In B. Dodd & R. Campbell (Eds.), *Hearing by eye: The psychology of lipreading* (pp. 145-161). London: Erlbaum.

Moeller, M., McCleary, E., Putman, C., et al. (2010). Longitudinal development of phonology and morphology in children with late-identified mild-moderate sensorineural hearing loss. *Ear Hear, 31*, 625-635.

Moore, B. (1995). Perceptual consequences of cochlear hearing loss and their implications for the design of hearing aids. *Ear Hear, 17*, 133-160.

Morton, J. (1982). Disintegrating the lexicon: An information processing approach. In E. Mehler, M. Garrett (Ed.), *Perspectives on mental representation: Experimental and theoretical studies of cognitive processes* *and capacities* (pp. 89-109). Hillsdale, NJ: Lawrence Erlbaum.

Nittrouer, S., & Burton, L. (2001). The role of early language experience in the development of speech perception and language processing abilities in children with hearing loss. *Volta Rev, 103*, 5-37.

Norbury, C. F., Bishop, D., Briscoe, J. (2001). Production of English finite verb morphology: A comparison of SLI and mild-moderate hearing impairment. *J Speech Lang Hear Res, 44*, 165-178.

Ormel, E., Hermans, D., Knoors, H., et al. (2010). Phonological activation during visual word recognition in deaf and hearing children. *J Speech Lang Hear Res, 53*, 801-820.

Pons, F., Andreu, L., Sanz-Torrent, M., et al. (2013). Perception of audio-visual speech synchrony in Spanish-speaking children with and without specific language impairment. *J Child Lang, 40*, 687-700.

Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cogn, 64*, 249-284.

Ross, M., & Lerman, J. (1971). *Word intelligibility by picture identification*. Pittsburgh: Stanwix House, Inc.

Schriefers, H., Meyer, A., Levelt, W. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *J Mem Lang, 29*, 86-102.

Seewald, R., Ross, M., Giolas, T., et al. (1985). Primary modality for speech perception in children with normal and impaired hearing. *J Speech Hear Res, 28*, 36-46.

Sterne, A., & Goswami, U. (2000). Phonological awareness of syllables, rhymes, and phonemes in deaf children. *J Child Psychol Psychiatry, 41*, 609-625.

Teinonen, T., Aslin, R., Alku, P., et al. (2008). Visual speech contributes to phonetic learning in 6-month-old infants. *Cogn, 108*, 850-855.

ten Oever, S., Sack, A. T., Wheat, K. L., et al. (2013). Audio-visual onset differences are used to determine syllable identity for ambiguous audio-visual stimulus pairs. *Frontiers Psychol, 4*, doi: 10.3389/fpsyg.2013.00331

Treiman, R., & Hirsh-Pasek, K. (1983). Silent reading - Insights from 2nd-generation deaf readers. *Cogn Psychol, 15*, 39-65.

Tye-Murray, N. (1992). Articulatory organizational strategies and the roles of audition. *Volta Rev, 94*, 243-260.

Tye-Murray, N., & Geers, A. (2001). *Children's audio-visual enhancement test*. St. Louis, MO: Central Institute for the Deaf.

Tye-Murray, N., Spencer, L., Gilbert-Bedia, E. (1995). Relationships between speech production and speech perception skills in young cochlear-implant users. *J Acoust Soc Am, 98*, 2454-2460.

Weikum, W., Vouloumanos, A., Navarra, J., et al. (2007). Visual language discrimination in infancy. *Science, 316*, 1159.

Woodhouse, L., Hickson, L., Dodd, B. (2009). Review of visual speech perception by hearing and hearing-impaired people: Clinical implications. *Int J Lang Commun Disord, 44*, 253-270.

Acknowledgements

This research was supported by the National Institute on Deafness and Other Communication Disorders, grant DC-00421. Dr. Abdi would like to acknowledge the support of an EURIAS fellowship at the Paris Institute for Advanced Studies (France), with the support of the European Union's 7th Framework Program for research, and from a funding from the French State managed by the "*Agence Nationale de la Recherche* (program: *Investissements d'avenir, ANR-11-LABX-0027-01 Labex RFIEA+*)." Sincere appreciation also to the speech science colleagues who advised us to adopt a perceptual criterion for editing the non-intact stimuli. We thank Dr. Brent Spar for recording the stimuli. We thank the children and parents who participated and the research staff who assisted, namely Aisha Aguilera, Carissa Dees, Nina Dinh, Nadia Dunkerton, Alycia Elkins, Brittany Hernandez, Cassandra Karl, Demi Krieger, Michelle McNeal, Jeffrey Okonye, Rachel Parra, and Kimberly Periman of UT-Dallas (data collection, analysis, presentation), and Derek Hammons and Scott Hawkins of UT-Dallas and Brent Spehar of CID-Washington University (computer programming).

Figure Legends

Fig. 1. Average picture naming times for the groups in the presence of the baseline vowel-onset distractors for the words (left) and nonwords (right) presented as auditory or AV input. Error bars are one standard error of the mean.

Fig. 2. Average priming proportions in the CHI vs. CNH for the high (intact) and low (non-intact) fidelity stimuli presented as auditory (left) or AV (right) input. Results are presented for the words (Figure 2a) and nonwords (Figure 2b). Error bars are one standard error of the mean. We derived the priming proportions by dividing each participant's adjusted naming times by his or her corresponding baseline naming times. Greater priming is indicated by a greater negative proportion.

Table 1. Average age and performance (standard deviation in parentheses) on set of nonverbal, verbal, and speech perception measures in the CNH vs CHI.

	Groups	
	CNH N = 62	CHI N = 62
Age (yr;mo)	9;1 (3;0)	9;2 (3;3)
Nonverbal Skills		
Simple Reaction Time (ms)		
Auditory	578 (161)	611 (240)
Visual	771 (255)	754 (318)
Visual Perception (standard score)	114.34 (14.08)	99.73 (15.50)
Verbal Skills		
Vocabulary (standard score)		
Receptive	120.50 (11.25)	94.39 (16.23)
Expressive	120.52 (11.91)	87.43 (12.16)
Articulation Proficiency (# errors)	0.73 (1.70)	4.98 (8.39)
Phonological Awareness (%)	78.02 (10.68)	66.16 (26.07)
Word Recognition (%)		
Auditory	99.87 (0.71)	89.35 (11.94)
AV	-------#	95.92 (10.49)
Lipreading Onsets*	66.57 (15.56)	68.87 (21.96)

-------# = test was not administered in the AV mode due to ceiling performance in the auditory only mode. *= lipreading onsets was selected because we are assessing phonological priming by onsets; lipreading words averaged 17.38% in CNH and 24.27% in CHI. ***Note***: Simple reaction times were estimated by a laboratory button push task quantifying the speed of detecting and responding to a predetermined auditory or visual target. Visual perception was estimated by the Beery-Buktenica Developmental Test of Visual Perception (Beery & Beery 2004). Vocabulary skills were estimated with the Peabody Picture Vocabulary Test-III (Dunn & Dunn 1997) and the Expressive One-Word Picture Vocabulary Test (Brownell 2000). Articulation proficiency was estimated with the Goldman Fristoe Test of Articulation (Goldman & Fristoe 2000). Phonological awareness was estimated with subtests of the Pre-Reading Inventory of Phonological Awareness (Dodd et al. 2003). Spoken word recognition at 70 dB SPL was estimated with the Word Intelligibility by Picture Identification (WIPI) Test (auditory mode; Ross & Lerman 1971) and the Children's Audiovisual Enhancement Test (CAVET; auditory, AV, and visual only (lipreading) modes; Tye-Murray & Geers 2001).

Table 2. *Summary of Statistical Results for Baseline Distractors*

Factors	***Mean Square Error***	***F value***	***p value***	***partial η^2***
Group	**691843.11**	**8.60**	**.004**	**.066**
Mode	**6116.18**	**5.65**	**.019**	**.044**
Lexical Status	**28495.62**	**5.34**	**.023**	**.042**
Group *x* Mode	6116.18	1.49	ns	--
Group *x* Lexical Status	28495.62	0.06	ns	--
Mode *x* Lexical Status	4063.39	0.18	ns	--
Group x Mode *x* Lexical Status	4063.39	0.08	ns	--

Note: Results of a mixed-design analysis of variance with one between-participants factor (Group: CHI vs. CNH) and two within-participants factors (Mode: auditory vs. AV; Lexical Status: words vs. nonwords). The dependent variable is the baseline naming times in ms. The degrees of freedom are 1,122 for all factors.

Table 3. *Summary of Statistical Results for Phonologically-Related Distractors*

A. Words

Factors	***Mean Square Error***	***F value***	***p value***	***partial*** η^2
Group	.016	0.54	ns	--
Fidelity	.072	**32.04**	**< .0001**	**.209**
Mode	.087	**22.25**	**< .0001**	**.155**
Group *x* Fidelity	.004	1.92	ns	--
Group *x* Mode	.007	1.83	ns	--
Fidelity *x* Mode	.002	1.52	ns	--
Group x Fidelity x Mode	.003	2.42	ns	--

B. Nonwords

Factors	***Mean Square Error***	***F value***	***p value***	***partial*** η^2
Group	.002	0.15	ns	--
Fidelity	.107	**63.95**	**< .0001**	**.347**
Mode	.156	**33.01**	**< .0001**	**.215**
Group *x* Fidelity	.018	**10.63**	**.001**	**.082**
Group *x* Mode	.028	**5.87**	**.017**	**.047**
Fidelity *x* Mode	.027	**15.74**	**.0001**	**.117**
Group x Fidelity x Mode	.007	**4.13**	**.044**	**.033**

B1. Nonwords for Each Mode Separately

AV

Factors	***Mean Square Error***	***F value***	***p value***	***partial*** η^2
Group	.010	0.69	ns	--
Fidelity	.001	**10.43**	**.002**	**.078**
Group *x* Fidelity	.001	1.04	ns	--

Auditory

Factors	***Mean Square Error***	***F value***	***p value***	***partial*** η^2
Group	.009	2.75	ns	--
Fidelity	.002	**57.33**	**< .0001**	**.319**
Group *x* Fidelity	.002	**10.67**	**.001**	**.081**

Note: Results of a mixed-design analysis of variance with one between-participants factor (Group: CHI vs CNH) and two within-participants factors (Fidelity: intact vs non-intact; Mode: auditory vs AV) followed by analyses of the nonwords for each mode separately. The dependent variable is the priming proportions ([mean time in the phonologically-related condition minus mean time in the baseline condition] divided by [mean time in the baseline condition]). The degrees of freedom are 1,122 for all factors.

Table 3. *Summary of F Contrast Results*

	High (intact) Fidelity				Low (non-intact) Fidelity			
	Mean Square Error	$F_{contrast}$	***p value***	***partial*** η^2	***Mean Square Error***	$F_{contrast}$	***p value***	***partial*** η^2
Group				***Words***				
CNH	.001	18.13	< **.0001**	.129	.001	47.32	< **.0001**	.279
CHI	.001	10.73	**.001**	.081	.001	7.75	**.006**	.060
				Nonwords				
CNH	.002	16.68	< **.0001**	.120	.002	96.10	< **.0001**	.441
CHI	.002	3.63	.06	.029	.002	14.53	**.0002**	.106

Note: Significant p values are bolded. *df*'s = 1,122.

Table 5. The part correlation coefficients and *p*-values evaluating the variation in the priming proportions for the nonwords uniquely accounted for (after removing the influence of the other variables) by auditory word recognition, age, and degree of hearing loss. Results are presented for auditory and AV speech (collapsed across fidelity).

Mode	***Auditory***				***AV***			
Variables	**Slope**	**Part *r***	***Partial F***	***p***	**Slope**	**Part *r***	***Partial F***	***p***
Auditory Word Recognition	.000	.000	.162	.987	**.024**	***.285****	***5.062***	***.028***
Age	–.004	.055	.000	.689	.009	.110	.731	.396
Degree of Loss	.011	.122	.001	.351	–.004	.055	.161	.689

Note: Significant results are bolded and starred. The multiple correlation coefficients for all of the variables considered simultaneously were 0.133 for Auditory and 0.327 for AV.
df's = 1,58 for Part *r and* 3,58 for Multiple *R*.

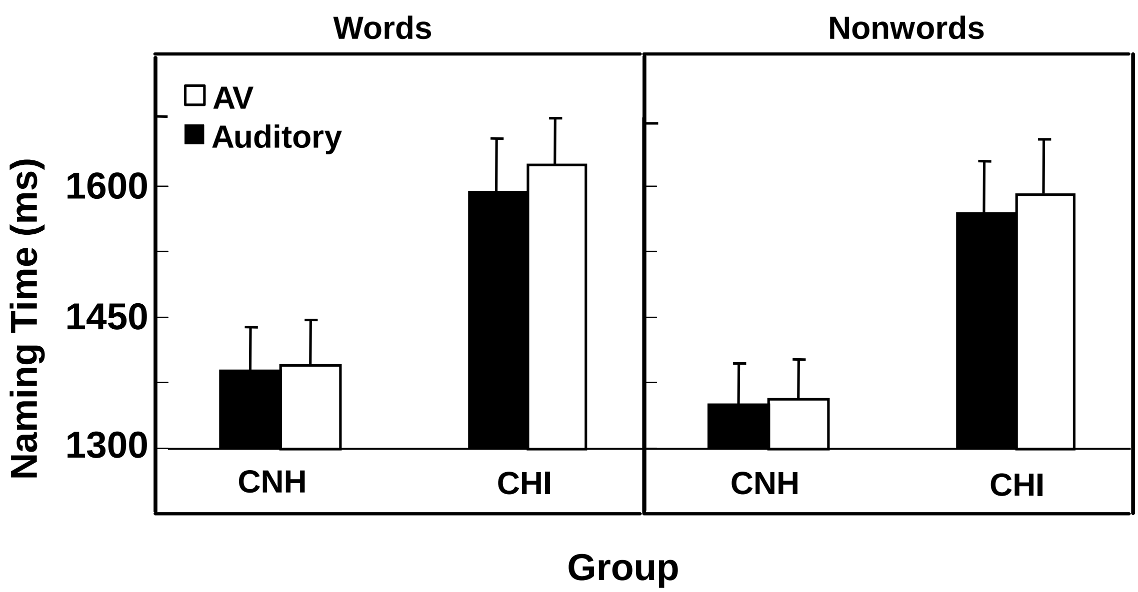

Fig. 1. Average picture naming times for the groups in the presence of the baseline vowel-onset distractors for the words (left) and nonwords (right) presented as auditory or AV input. Error bars are one standard error of the mean.

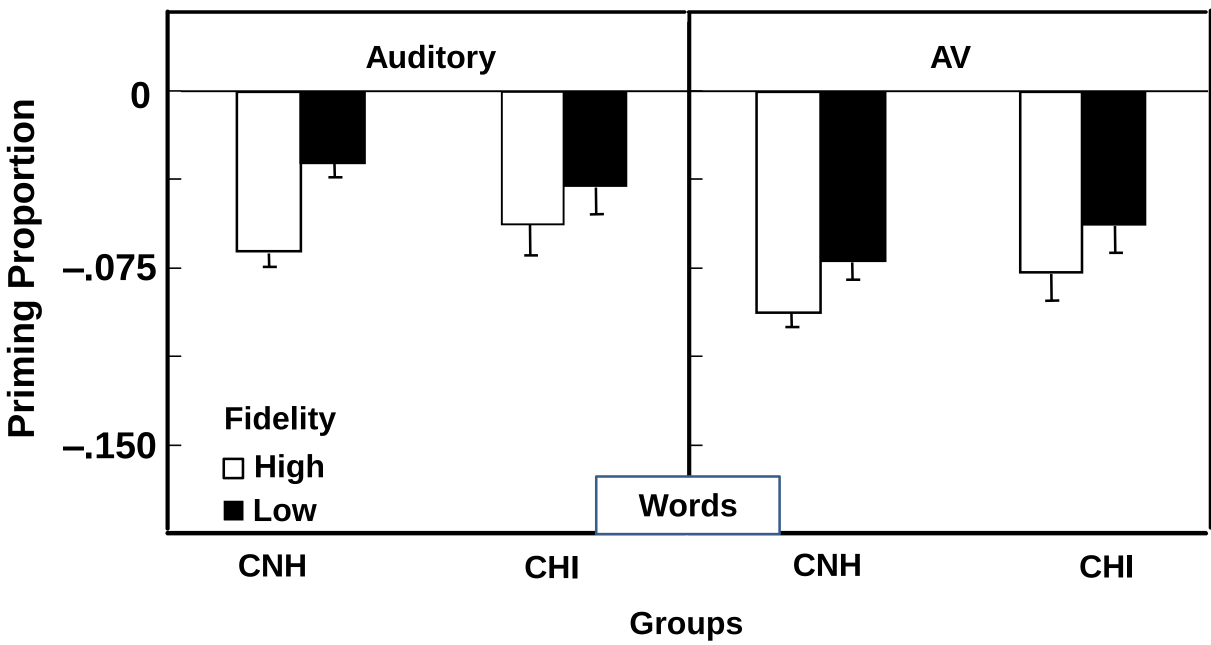

Fig. 2. Average priming proportions in the CHI vs. CNH for the high (intact) and low (non-intact) fidelity stimuli presented as auditory (left) or AV (right) input. Results are presented for the words (Figure 2a) and nonwords (Figure 2b). Error bars are one standard error of the mean. We derived the priming proportions by dividing each participant's adjusted naming times by his or her corresponding baseline naming times. Greater priming is indicated by a greater negative proportion.

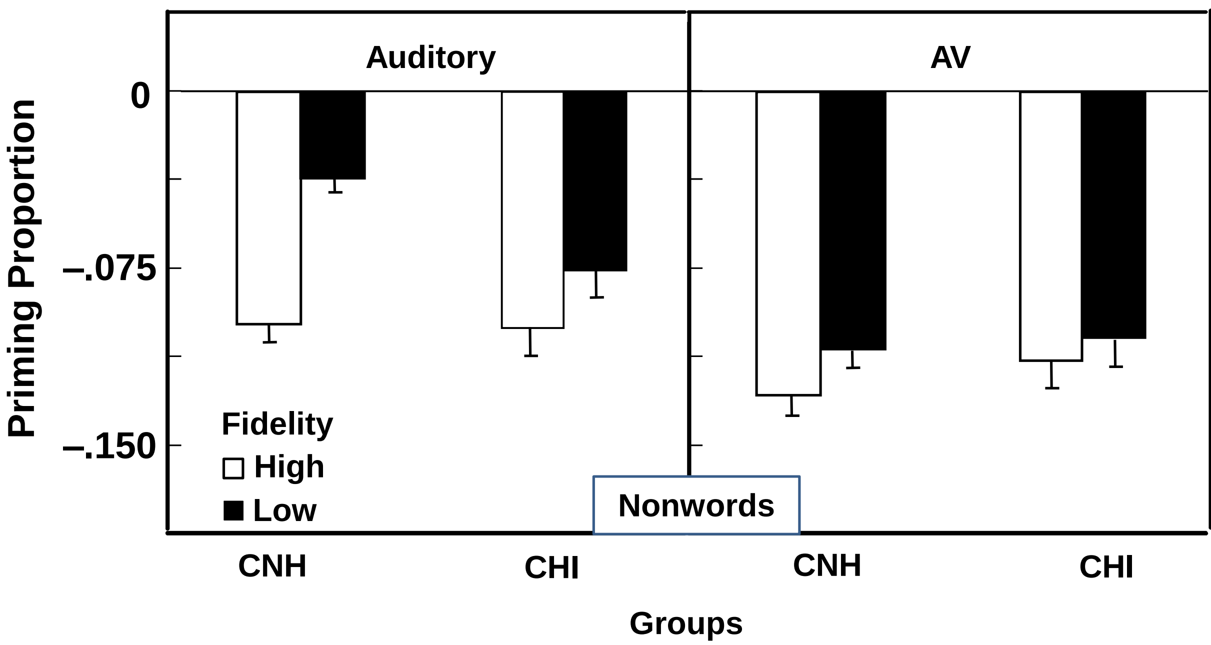

Fig. 2. Average priming proportions in the CHI vs. CNH for the high (intact) and low (non-intact) fidelity stimuli presented as auditory (left) or AV (right) input. Results are presented for the words (Figure 2a) and nonwords (Figure 2b). Error bars are one standard error of the mean. We derived the priming proportions by dividing each participant's adjusted naming times by his or her corresponding baseline naming times. Greater priming is indicated by a greater negative proportion.

Appendix A: Items

Table 1A.

The Pictures and Word and Nonword Distractors

Along with Example Filler Items

Pictures		Distractors	
		Word	***Nonword***
Experimental Items			
		Phonologically-Related	
bat		bag	*baz*
beads		bean	*beece*
boat		bone	*bohs*
bug		bus	*buhl*
gas		gap	*gak*
geese		gear	*geen*
ghost		gold	*guks*
gun		guts	*guks*
		Vowel-Onset Baseline	
bat	**gas**	apple	*apper*
beads	**geese**	eagle	*eeble*
boat	**ghost**	ocean	*oshuck*
bug	**gun**	onion	*onyit*
Example Filler Items			
dog		cheese	*cheeg*
mouth		tiger	*tyfer*
shoes		finger	*fihver*
sun		fox	*foms*

Note. The word and nonword distractors were constructed to have as comparable phonotactic probabilities as possible *(see Jerger et al., 2014, for details).* If a picture-filler item pair was used for words, it was not used for nonwords (and vice versa) in order to emphasize the distinction between the words and nonwords.

Appendix B: Characteristics of the Data

Replacement and Missing Trials. Naming responses that were flawed (e.g., lapses of attention; squirming out of position; triggering the VOR with a nonspeech sound, etc) were deleted on-line and re-administered after intervening items. The total number of trials deleted on-line (with replacement) for the phonologically-related distractors averaged 2.1 (intact)and 2.4 (non-intact) for the CNH and 1.6 (intact)and 1.8 (non-intact) for the CHI. The number of missing trials remaining at the end because the replacement trial was also flawed averaged 0.4 (intact)and 0.5 (non-intact) for the CNH and 0.3 (intact & non-intact) for the CHI.

Words vs Nonwords. To ensure that the experimental results were reflecting performance for words vs nonwords, the participants' understanding of the meaning of the word distractors was tested by parental report and a picture-pointing task. The number of word distractors whose meaning had to be taught averaged 0.18 in the CNH and 2.23 in the CHI. With regard to the word distractors, 15% of the CNH had to be taught, on average, 1.4 distractors and 59% of the CHI had to be taught, on average, 3.4 distractors. With regard to the pictures, the number of pictures whose names had to be taught averaged 0.03 in the CNH and 0.71 in the CHI. With regard to the pictures, 2% of CNH had to be taught, on average, 1 picture-name and 32% of the CHI had to be taught, on average, 2 picture-names. Mean naming times with the taught vs previously known word distractors or pictures did not differ; no trials were eliminated.

Pronunciation. The onsets of the pictures' names were accurately pronounced by all CNH and by 95% of the CHI. The 3 CHI—who mispronounced, on average, 2.6 picture onsets—*correctly* pronounced some of the corresponding onsets for the distractor repetition task, the remaining picture-names, and/or the Goldman-Fristoe Test of Articulation (Goldman & Fristoe 2000). Because these few children were inconsistent in their mispronunciations, no pictures were deleted on this basis.

Distractor Repetition Task. To control for mishearing (operationally defined as misrepeating) the onset of the intact distractors, we deleted all distractors whose onsets were not correct on a repetition task. This constraint required deletion of 1 picture-distractor pair for the nonwords, auditory mode, in the CNH. In the CHL, we deleted, on average, from 1.1 to 1.3 picture-distractor pairs (out of 8 pairs) in 11% (AV) to 18% (auditory) of children for words and in 5% (AV) to 21% (auditory) of children for nonwords. Incorrect picture-distractor pairs were deleted for both the intact and non-intact stimuli to keep the intact/non-intact comparison to the same base.

Appendix C: Preliminary Analyses of Picture Word Data

To determine whether results could be collapsed across the different distractor onsets (/b/ vs /g/), we examined the effect of the onset on the baseline and experimental conditions.

Baseline Conditions. A factorial mixed-design analysis of variance was carried out with one between-participants factor (group: CHL vs CNH) and three within-participant factors representing lexical status (words vs nonwords), mode (auditory vs AV), and onset (/b/ vs /g/). Results indicated that the baseline picture naming times did not differ as a function of the onset.

Experimental Conditions. We quantified the priming of picture naming produced by the phonologically-related distractors with adjusted picture naming times derived by subtracting each participant's baseline naming times from his or her phonologically-related naming times as done previously (e.g., Jerger et al., 2002; Jerger et al., 2009). Analysis of the adjusted picture naming times consisted of one between-participants factor (group: CHL vs CNH) and four within-participant factors representing lexical status (words vs nonwords), fidelity (intact vs non-intact), mode (auditory vs AV), and onset (/b/ vs /g/). Results revealed that the phonologically-related distractors facilitated overall picture naming significantly more for the /b/ than /g/ onsets (–142 vs –110 ms; i.e., a difference of 32 ms), F (1,122) = 15.74, MSE =33355.31, $p < .0001$, partial $\eta^2 = .114$. No other significant onsets effects were observed. Despite this statistically significant outcome, the difference in performance between the two onsets was small and did not interact with group (CHL vs CNH), lexical status (words vs nonwords), mode (AV vs auditory), or fidelity (intact vs non-intact). Thus, for the primary analyses, all naming times were collapsed across the onsets for simplicity.