Role of Visual Speech in Phonological Processing by Children With Hearing Loss

Susan Jerger
University of Texas at Dallas, Callier Center
for Communication Disorders and
Central Institute for the Deaf at
Washington University School
of Medicine, St. Louis, MO

Nancy Tye-Murray
Central Institute for the Deaf and Washington
University School of Medicine,
University of Texas at Dallas

Hervé Abdi
University of Texas at Dallas

Purpose: This research assessed the influence of visual speech on phonological processing by children with hearing loss (HL).
Method: Children with HL and children with normal hearing (NH) named pictures while attempting to ignore auditory or audiovisual speech distractors whose onsets relative to the pictures were either congruent, conflicting in place of articulation, or conflicting in voicing—for example, the picture "pizza" coupled with the distractors "peach," "teacher," or "beast," respectively. Speed of picture naming was measured.
Results: The conflicting conditions slowed naming, and phonological processing by children with HL displayed the age-related shift in sensitivity to visual speech seen in children with NH, although with developmental delay. Younger children with HL exhibited a disproportionately large influence of visual speech and a negligible influence of auditory speech, whereas older children with HL showed a robust influence of auditory speech with no benefit to performance from adding visual speech.
The congruent conditions did not speed naming in children with HL, nor did the addition of visual speech influence performance. Unexpectedly, the /ʌ/-vowel congruent distractors slowed naming in children with HL and decreased articulatory proficiency.
Conclusions: Results for the conflicting conditions are consistent with the hypothesis that speech representations in children with HL (a) are initially disproportionally structured in terms of visual speech and (b) become better specified with age in terms of auditorily encoded information.

KEY WORDS: phonological processing, lipreading, picture–word task, multimodal speech perception

For decades, evidence has suggested that visual speech may play an important role in learning the phonological structure of spoken language (Dodd, 1979, 1987; Locke, 1993; Mills, 1987; Weikum et al., 2007). A proposed link between visual speech and the development of phonology may be especially essential for children with prelingual hearing loss. The importance of this link is challenged, however, by the observation that visual speech has less influence on speech perception in typically developing children than in adults. The McGurk effect illustrates this finding (McGurk & MacDonald, 1976). In a McGurk task, individuals hear a syllable whose onset represents one place of articulation while seeing a talker simultaneously mouthing a syllable whose onset represents a different place of articulation (e.g., auditory /ba/ and visual /ga/). Adults typically experience the illusion of perceiving a blend of the auditory and visual places of articulation (e.g., /da/ or /ða/). Significantly fewer children than adults experience this illusion. As an example, in response to one type of McGurk stimulus (auditory /ba/ – visual

/ga/), McGurk and MacDonald (1976) reported auditory capture (individuals heard /ba/) in 40%–60% of children but in only 10% of adults. This pattern of results has been replicated and extended to other tasks by several investigators. Visual speech has less influence on performance by children for (a) identification of nonsense syllables representing congruent and/or incongruent audiovisual pairings such as auditory /ba/ paired with visual /ba/, /va/, /da/, /ga/, /ða/, and/or /θa/ (Desjardins, Rogers, & Werker, 1997; Hockley & Polka, 1994); auditory /aba/ paired with visual /aba/, /ada/, /aga/, and /ava/ (Dupont, Aubin, & Menard, 2005); or auditory /ba/ in noise paired with visual /ba/, /da/, and /ga/ (Sekiyama & Burnham, 2004); (b) identification of synthesized speech stimuli ranging from /ba/ to /da/ paired with visual /ba/ or /da/ (Massaro, 1984; Massaro, Thompson, Barron, & Laren, 1986); and (c) the release from informational masking due to the addition of visual speech for recognizing speech targets (e.g., colored numbers such as red eight) in the presence of competing speech (Wightman, Kistler, & Brungart, 2006). Overall results are consistent with the conclusion that audiovisual speech perception is dominated by auditory input in typically developing children and visual input in adults. The course of development of audiovisual speech perception to adultlike performance is not well understood. A few studies report an influence of visual speech on performance by the pre-teen/teenage years (Conrad, 1977; Dodd, 1977, 1980; Hockley & Polka, 1994), with one report citing an earlier age of 8 years (Sekiyama & Burnham, 2004).

Recently, we modified the Children's Cross-Modal Picture-Word Task (Jerger, Martin, & Damian, 2002) into a multimodal procedure for assessing indirectly the influence of visual speech on phonological processing (Jerger, Damian, Spence, Tye-Murray, & Abdi, 2009). Indirect versus direct tasks may be demarcated on the basis of task instructions as recommended by Merikle and Reingold (1991). An indirect task does not direct participants' attention to the experimental manipulation of interest, whereas a direct task unambiguously instructs participants to respond to the experimental manipulation. Both the Cross-Modal and Multimodal Picture-Word tasks qualify as indirect measures. Children are instructed to name pictures and to ignore auditory or audiovisual speech distractors that are nominally irrelevant to the task. The participants are not informed of, nor do they consciously try to respond to, the manipulation. An advantage of indirect tasks is that performance is less influenced by developmental differences in higher level cognitive processes, such as the need to consciously access and retrieve task-relevant knowledge (Bertelson & de Gelder, 2004). The extent to which age-related differences in multimodal speech processing are reflecting development change versus the varying demands of tasks remains an important unresolved issue. In the next section, we briefly describe our original task and the recent adaptation.

Children's Multimodal Picture-Word Task

In the Cross-Modal Picture-Word Task (Jerger, Martin, & Damian, 2002), the dependent measure is the speed of picture naming. Children name pictures while attempting to ignore auditory distractors that are either phonologically onset related, semantically related, or unrelated to the picture. The phonological distractors contain onsets that are congruent, conflicting in place of articulation, or conflicting in voicing (e.g., the picture "pizza" coupled with the distractors "peach," "teacher," or "beast," respectively). The semantic distractors comprise categorically related pairs (e.g., the picture "pizza" coupled with the distractor "hot dog"). The unrelated distractors are composed of vowel nucleus onsets (e.g., the picture "pizza" coupled with the distractor "eagle"). The unrelated distractors usually form a baseline condition, and the goal is to determine whether phonological or semantic distractors speed up or slow down naming relative to the baseline condition. Picture-word data may be gathered simultaneously with the entire set of distractors, particularly in children, due to (a) the difficulties in recruiting and testing the participants and (b) the idea that an inconsistent relationship between the distractors and pictures may aid participants' attempts to disregard the distractors. Research questions about the different types of distractors, however, are typically reported in separate articles. In this article aimed at explicating the influence of visual speech on phonological processing, we focus only on the phonological distractors.

The onset of the phonological distractors may be varied to be before or after the onset of the pictures, referred to as the *stimulus onset asynchrony* (SOA). Whether a distractor influences picture naming depends upon the type of distractor and the SOA. First, with regard to the SOA, a phonologically related distractor typically produces a maximal effect on naming when the onset of the distractor lags behind the onset of the picture (Damian & Martin, 1999; Schriefers, Meyer, & Levelt, 1990). With regard to the type of distractor, a phonologically related distractor speeds naming when the onset is congruent and slows naming when the onset is conflicting in place or voicing relative to the unrelated or baseline distractor. The basis of any facilitation or interference is assumed to be interactions between the phonological representations supporting speech production and perception.

The interaction between speech production and perception is illustrated in Figure 1, which portrays the speech chain (see Denes & Pinson, 1993) with an auditory distractor. The speaker is supposed to be naming the

Figure 1. The speech chain (adapted from Jerger, 2007) for auditory-only input representing the stages of processing for speech production and perception. The speech chain has been augmented to portray a speaker naming the picture, "duck," while hearing an auditory phonologically related distractor conflicting in place of articulation, "buckle" (see Jerger, 2007).
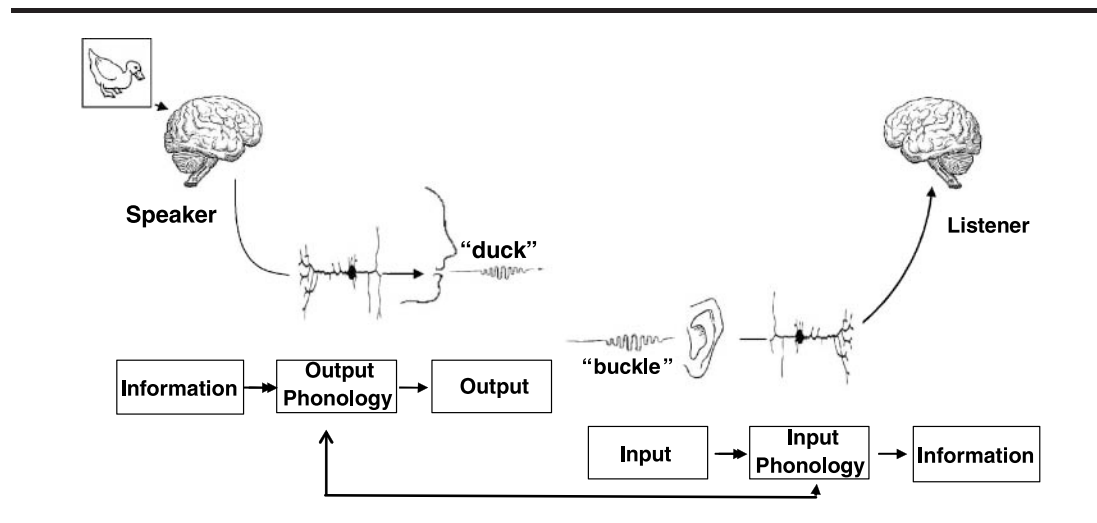
picture "duck" while hearing a phonologically onset-related distractor conflicting in place of articulation, "buckle." The boxes in Figure 1 illustrate the stages of processing for speech production and perception, which proceed in opposite directions. With regard to producing speech, the picture to be named activates conceptual linguistic information that undergoes transformation to derive its pronunciation, termed *output phonology,* followed by articulation, termed *output.* With regard to perceiving speech, the acoustic waveform of the distractor enters the listener's ear, where it undergoes input phonological processing to derive the word's phonological pattern, which activates the word's meaning (i.e., conceptual information).

When congruent or conflicting auditory distractors speed up or slow down naming, performance is assumed to reflect an interaction between production and perception (Levelt et al., 1991). The interaction is assumed to occur when the picture-naming process is occupied with output phonology and the distractor perceptual process is occupied with input phonology. Congruent distractors are assumed to speed picture naming by activating input phonological representations whose activation spreads to output phonological representations, thus allowing the shared speech segments to be selected more rapidly during the naming process. Conflicting distractors are assumed to slow naming by activating conflicting output phonological representations that compete with the picture's output phonology for control of the response. A novel contribution of our new methodology is that the distractors are presented both auditorily and audiovisually. Thus, we can gain new knowledge about

the contribution of visual speech to phonological processing by children with HL.

Results in Typically Developing Children

Originally, we hypothesized that results on our indirect Multimodal Picture-Word Task would consistently reveal an influence of visual speech on performance, in contrast to results of the studies reviewed previously using direct tasks (Desjardins et al., 1997; Dupont et al., 2005; Hockley & Polka, 1994; Massaro, 1984; Massaro et al., 1986; Sekiyama & Burnham, 2004; Wightman et al., 2006). Disparate results on indirect versus direct visual speech tasks have been reported previously by Jordan and Bevan (1997). Results did not support our hypothesis (see Jerger et al., 2009). Although preschoolers (4-year-olds) and preteens/teenagers (10- to 14-year-olds) showed an influence of visual speech on phonological processing (termed *positive results*), young elementary school–aged children did not (termed *negative results*). Positive results in younger and older children coupled with negative results in children of in-between ages yielded a U-shaped developmental function.

U-shaped functions have been carefully scrutinized by dynamic systems theorists, who propose that multiple interactive factors—rather than one single factor—typically form the basis of developmental change (Gershkoff-Stowe & Thelen, 2004; Smith & Thelen, 2003). Furthermore the components of early skills are viewed as "softly assembled" behaviors that reorganize into more mature and stable forms in response to internal and environmental forces (Gershkoff-Stowe & Thelen, 2004).

From this perspective, the plateau of the U-shaped trajectory, which seems to reflect the loss or disappearance of a behavior, is instead viewed as reflecting a period of transition and instability. Applying knowledge systems that are in a period of significant growth may require more resources and may overload the processing system, resulting in temporary decreases in processing efficiency. A proposal, in concert with dynamic systems theory, is that the temporary loss of sensitivity to visual speech in the children of in-between ages is reflecting reorganization and adaptive growth in the knowledge systems underpinning performance on our task, particularly phonology.

In short, the apparent loss of the influence of visual speech on performance in elementary school–aged children may be viewed as a positive sign, an indication that relevant knowledge systems—particularly phonology— are reorganizing and transitioning into more highly specified, stable, and robust forms. The reorganization and transition may be in response to environmental and internal forces such as (a) formal instruction in the reading and spelling of an alphabetic language and (b) developmental changes in multimodal processing as well as in auditory perceptual, linguistic, and cognitive skills. An important point is that the ends of the U-shaped trajectory, which seemed to reflect identical performance in the 4-year-olds and 10- to 14-year-olds, are not necessarily reflecting identical substructures for the performance.

Results in the typically developing children with normal hearing (NH) raise questions about whether children with prelingual hearing loss (HL), who rely more on visual speech to learn to communicate, will show a lack of influence of visual speech on phonological processing during the years characterizing reorganization in the children with NH. To the extent that the apparent loss of the influence of visual speech on performance reflects developmental evolution within the relevant knowledge systems, particularly phonology, then perhaps children with HL and no other disabilities, excluding delayed speech and language, also undergo this pivotal transformation. In the next section, we summarize general primary and secondary research questions, predict some of the possible results, and interpret some of the possible implications for speech representations in children with HL. Our model for predicting and interpreting results proposes that performance on the Multimodal Picture-Word Task reflects cross-talk between the output and input phonological representations supporting speech production and perception. The patterns of cross-talk provide a basis for theorizing about the nature of the knowledge representations underlying performance (see Rayner & Springer, 1986, for a discussion). Visual speech is viewed as an extra phonetic resource (perhaps adding another type of phonetic feature such as mouth shape)

that can enhance performance relative to auditory speech only (see Campbell, 1988, for a discussion).

Research Questions and Predicted Results in Children With HL

Our general research aim was to explain whether and how visual speech may enhance phonological processing by children with HL relative to auditory speech alone. We addressed the following research questions by comparing performance in NH versus HL groups or in subgroups of children with HL.

Question 1. Does performance on the picture-word task differ in the NH versus HL groups for the auditory and audiovisual distractors?

Question 2. Does performance on the picture-word task differ in HL subgroups representing less mature versus more mature age-related competencies?

With regard to the possible influence of prelingual childhood HL on phonology, auditory input is assumed to play a disproportionately important role in building phonological knowledge (Levelt, Roelofs, & Meyer, 1999; Ruffin-Simon, 1983; Tye-Murray, 1992; Tye-Murray, Spencer, & Gilbert-Bedia, 1995). The presence of HL may degrade and filter auditory input, resulting in less well-specified auditory phonological knowledge. Visual speech may become a more important source of phonological knowledge, with representations more closely tied to articulatory or speechread codes than auditory codes. Relatively impoverished phonological knowledge in children with moderate HL is supported by previous research demonstrating abnormally poor phoneme discrimination and phonological awareness (Briscoe, Bishop, & Norbury, 2001).

With regard to the first research question, the data of Briscoe and colleagues predict that the auditory distractors should affect performance significantly less in the HL group than in the NH group. This outcome would support the contention that speech representations in the children with HL are less well-structured in terms of auditorily based linguistic information. With regard to the audiovisual distractors, the data suggest that the children with HL should show a stronger influence of visual speech on performance relative to the NH group. This outcome would be consistent with the contention that speech representations in the children with HL are encoded disproportionally in terms of visual speech gestural and/or articulatory codes, as previously suggested by Jerger, Lai, and Marchman (2002).

In distinction to this vantage point, there is evidence indicating extensive phoneme repertoires and reasonably normal phonological processes in children with moderate HL, although results are characterized by significant developmental delay and individual

variability (Dodd & Burnham, 1988; Dodd & So, 1994; Oller & Eilers, 1981; see Moeller, Tomblin, Yoshinaga-Itano, Connor, & Jerger, 2007, for a discussion). With regard to the first research question, these data predict that results for the auditory distractors in the children with HL will not differ from those in the NH group. This outcome would support the idea that speech representations in these children with HL are sufficiently specified in terms of auditorily encoded linguistic information to support performance without relying on visual speech. Scant evidence with the Cross-Modal Picture-Word Task and auditory nonsense syllable distractors suggests that phonological distractors produce the expected pattern of results, facilitation from congruent distractors and interference from conflicting distractors, in children with moderate HL and good phoneme discrimination but not in children with moderate HL and poor phoneme discrimination (Jerger, Lai, & Marchman, 2002).

With regard to possible results for the second of the primary research questions, one predicted outcome for the auditory distractors from the evidence in the literature about developmental delay is that performance in the children with HL may change with increasing age. To the extent that auditorily encoded linguistic information becomes sufficiently strong with age to support performance without relying on visual speech, we may see a developmental shift in the influence of visual speech on performance as seen in a previous study with children with NH (Jerger et al., 2009). Results would echo the normal pattern of results if there is an influence of visual speech in younger children that decreases with increasing age. This outcome would be consistent with the idea of adaptive growth and reorganization in the knowledge systems underpinning performance on our task, particularly phonology.

In addition to the two primary questions, a secondary question addressed whether performance on the picture-word task differs in HL groups representing poorer versus better hearing statuses as measured by auditory speech recognition ability. We should note that most of the current participants with HL had good hearing capacities as operationally defined by auditory word recognition scores due to stringent auditory criteria for participation (see Method section). Thus, results examining poorer versus better hearing status may be limited, especially if age must also be controlled when assessing the influence of hearing status. To the extent that speech representations may be better structured in terms of auditorily based linguistic information in children with better hearing status, a predicted outcome is that auditory distractors will affect performance significantly more in children with better hearing status versus poorer hearing status. A corollary to this idea is that speech representations in children with poorer hearing

status may be encoded disproportionally in terms of visual speech gestural and/or articulatory codes. In this case, a predicted outcome is that audiovisual distractors may influence performance significantly more in children with poorer hearing status due to their greater dependence on visual speech for spoken communication.

These data should provide new insights into the phonological representations underlying the production and perception of words by children with HL. Studying onset effects may be particularly revealing to the extent that incremental speech processing skills are important to spoken word recognition and production and that word-initial input is speechread more accurately than word-noninitial input (Fernald, Swingley, & Pinto, 2001; Greenberg & Bode, 1968; Marslen-Wilson & Zwitserlood, 1989; Gow, Melvold, & Manuel, 1996).

Method
Participants

HL group. Participants were 31 children (20 boys and 11 girls) with prelingual sensorineural hearing loss (SNHL) ranging in age from 5;0 (years;months) to 12;2. The racial/ethnic distribution was 74% White, 16% Black, 6% Asian, and 3% multiracial; of the total group, 6% were of Hispanic ethnicity. We initially screened 63 children with a wide range of hearing impairments to attain the current pool of participants who met the following criteria: (a) prelingual SNHL; (b) English as a native language; (c) ability to communicate successfully aurally/orally; (d) ability to hear accurately—on auditory-only, open-set testing—100% of the baseline distractor onsets and at least 50% of all other distractor onsets; (e) ability to discriminate accurately—on auditory-only two-alternative forced-choice testing with a 50% chance level—at least 85% of the voicing (p-b, t-d) and place of articulation (p-t, b-d) phoneme contrasts; and (f) no diagnosed or suspected disabilities, excluding the speech and language problems that accompany prelingual childhood HL.

Unaided hearing sensitivity on the better ear, estimated by pure-tone average (PTA) hearing threshold levels (HTLs) at 500, 1000, and 2000 Hz (American National Standards Institute [ANSI], 2004), averaged 50.13 dB HTL and was distributed as follows: ≤20 dB ($n = 7$), 21–40 dB ($n = 5$), 41–60 dB ($n = 9$), 61–80 dB ($n = 4$), 81–100 dB ($n = 2$), and >101 dB ($n = 4$). An amplification device was worn by 24 of the 31 children—18 were hearing aid users and 6 were cochlear implant or cochlear implant plus hearing aid users. Of the children who wore amplification, the average age at which they received their first listening device was 34.65 months ($SD = 19.67$ months). Duration of listening device use was 60.74 months ($SD = 20.87$ months). Participants who wore amplification were tested while wearing their devices. Most devices were

digital aids that were self-adjusting with the volume control either turned off or nonexistent. The children were recruited from cooperating educational programs. The type of program was a mainstreamed setting for 25 children with some assistance from special education services for 1 child, deaf education for 5 children, and total communication for 1 child. Again, all children communicated successfully aurally/orally. With regard to non-auditory status, all children passed measures establishing normalcy of visual acuity (including corrected-to-normal) and oral motor function. The average Hollingshead (1975) social strata score, 2.16, was consistent with a minor professional, medium business, or technical socioeconomic status (SES).

NH comparison group. Individuals were 62 children with normal hearing (33 boys and 29 girls) selected from a pool of 100 typically developing children who participated in an associated project explicating normal results (see Jerger et al., 2009). Ages ranged from 5;3 to 12;1. The racial/ethnic distribution was 76% White, 5% Asian, 2% Black, 2% Native American, and 6% multiracial; of the group, 15% reported a Hispanic ethnicity. All children passed measures establishing normalcy of hearing sensitivity, visual acuity (including corrected-to-normal), gross neurodevelopmental history, and oral motor function. The average Hollingshead social strata score, 1.52, was consistent with a major business and professional SES. A comparison of the groups on the verbal and nonverbal demographic measures, described in the section that follows, is detailed in the Results section.

Demographic Measures
Materials, Instrumentation, and Procedure

Sensory-motor function and SES. All of the standardized measures in this and the following sections were administered and scored according to the recommended techniques. Hearing sensitivity was assessed with a standard pure-tone audiometer. Normal hearing sensitivity was defined as bilaterally symmetrical thresholds of ≤20 dB HTL at all test frequencies between 500 and 4000 Hz (ANSI, 2004). Visual acuity was screened with the Rader Near Point Vision Test (Rader, 1977). Normal visual acuity was defined as 7 out of 8 targets correct at 20/30 Snellen Acuity (including participants with corrected vision). Oral-motor function was screened with a questionnaire designed by an otolaryngologist who is also a speech pathologist (Peltzer, 1997). The questionnaire contained items concerning eating, swallowing, and drooling. Normal oral-motor function was assumed if the child passed all items on the questionnaire. SES was estimated with the Hollingshead Four-Factor Index (Hollingshead, 1975).

Nonverbal Measures

Visual-motor integration. Visual-motor integration was estimated with the Beery-Buktenica Developmental Test of Visual-Motor Integration (Beery, Buktenica, & Beery, 2004). *Visual-motor integration* refers to the capacity to integrate visual and motor activities. Visual-motor integration was assessed by children's ability to reproduce shapes ranging from a straight line to complex three-dimensional geometric forms.

Visual perception. Visual perception was estimated with the subtest of the Beery-Buktenica Developmental Test of Visual-Motor Integration (Beery et al., 2004). Visual perception was assessed by children's ability to visually identify and point to an exact match for geometric shapes.

Visual simple reaction time (RT). Simple visual RT was assessed by a laboratory task that quantified children's speed of processing in terms of detecting and responding to a predetermined target. The stimulus was always a picture of a basketball. Children were instructed to push the response key as fast and as accurately as possible. Each run consisted of practice trials until performance stabilized, followed by a sufficient number of trials to yield eight good reaction times.

Verbal Measures

Vocabulary. Receptive and expressive knowledge were estimated with the Peabody Picture Vocabulary Test, Third Edition (PPVT-III; Dunn & Dunn, 1997) and the Expressive One-Word Picture Vocabulary Test (Brownell, 2000).

Output phonology. Output phonology was estimated with the Goldman–Fristoe Test of Articulation (Goldman & Fristoe, 2000). Articulation of the picture-names of the Picture-Word Task (Jerger et al., 2009) was also assessed. Pronunciation of the pictures' names was scored in terms of both onsets and offsets. Dialectic variations in pronunciation were not scored as incorrect.

Input phonology. Input phonological knowledge was estimated by laboratory measures of auditory and audiovisual onset and rhyme skills. These skills were chosen because they are usually operational in young children and are assumed to be more independent of reading skill than other phonological measures (Stanovich, Cunningham, & Cramer, 1984; Bird, Bishop, & Freeman, 1995). The onset test consisted of 10 words beginning with the stop consonants of the Picture-Word Task. Each word had a corresponding picture response card with four alternatives. The alternatives had a CV nucleus and represented consonant onsets with (a) correct voicing and correct place of articulation, (b) correct voicing and incorrect place of articulation, (c) incorrect voicing and correct place of articulation, and (d) incorrect voicing and

incorrect place of articulation. For example, for the target onset segment /dʌ/, the pictured alternatives were "duck," "bus," "tongue," and "puppy." The rhyme judgment task also consisted of 10 words, each with a corresponding picture response card containing four alternatives. The alternatives represented the following relations to the test item ("boat"): the rhyming word ("goat") and words with the test item's initial consonant ("bag"), final consonant ("kite"), and vowel ("toad"). Children were asked to indicate "Which one begins with /dʌ/?" or "Which one rhymes with boat?" Scores for both measures were percentage correct. Individual items making up the tests were administered randomly. The items were recorded and edited as described in the *Preparation of stimuli* section. In brief, all auditory and audiovisual items were recorded digitally by the same talker who recorded the stimuli for the experimental task.

In addition to the previously mentioned measures used to compare skills in the HL versus NH groups, another input phonological task, phoneme discrimination, was used to screen participants. Phoneme discrimination was assessed with auditory and audiovisual same–different tasks, which were administered in a counterbalanced manner across participants. Test stimuli were the stop consonants of the picture-word test (p/, /b/, /t/, /d/) coupled with the vowels (/i/ and /ʌ/). The test comprised 40 trials representing a priori probabilities of 60% "different" and 40% "same" responses. The same stimuli consisted of two different utterances; thus, the same pairs had phonemic constancy with acoustic variability. The intersignal interval from the onset of the first syllable to the onset of the second syllable was 750 ms. The response board contained two telegraph keys labeled "same" and "different." The labels were two circles (or two blocks for younger children) of the same color and shape or of different colors and shapes. A blue star approximately 3 cm to the outside of each key designated the start position for each hand, assumed before each trial. Children were instructed to push the correct response key as fast and as accurately as possible. Children were informed that about one-half of trials would be the same and about one-half would be different. Individual items were administered randomly. Again, the items were recorded by the same talker and were edited as described in the *Preparation of stimuli* section.

Word recognition. Auditory performance for words was quantified by averaging results on a test with a closed-response set, the Word Intelligibility by Picture Identification (WIPI) test (Ross & Lerman, 1971), and one with an open-response set, the auditory-only condition of the Children's Audio-Visual Enhancement Test (CAVET; Tye-Murray & Geers, 2001). Visual-only and audiovisual word performance were also estimated with the CAVET. Finally, recognition of the distractor words of the Picture-Word Task was assessed, both auditory

only and audiovisually. The Distractor Recognition Task was scored in terms of both words and onsets.

Experimental Picture-Word Task
Materials and Instrumentation

Pictures and distractors. Specific test items and conditions comprising the Children's Cross-Modal Picture-Word Task have been detailed previously (Jerger, Martin, & Damian, 2002). The pictured objects in this study are the same items. In brief, the pictures' names always began with /b/, /p/, /t/, or /d/, coupled with the vowels /i/ or /ʌ/. These onsets represent developmentally early phonetic achievements and reduced articulatory demands and tend to be produced accurately by young children, both NH and HL (Abraham, 1989; Dodd, Holm, Hua, & Crosbie, 2003; Smit, Hand, Freilinger, Bernthal, & Bird, 1990; Stoel-Gammon & Dunn, 1985; Tye-Murray, Spencer, & Woodworth, 1995). To the extent that phonological development is a dynamic process, with knowledge improving from (a) unstable, minimally specified, and harder-to-access/retrieve representations to (b) stable, robustly detailed, and easier-to-access/retrieve representations, it seems important to assess early-acquired phonemes that children are more likely to have mastered in an initial study (see McGregor, Friedman, Reilly, & Newman, 2002, for similar reasoning about semantic knowledge). The onsets represent variations in place of articulation (/b/ – /d/ vs. /p/ – /t/) and voicing (/b/ – /p/ vs. /d/ – /t/), two phonetic features that are traditionally thought to be differentially dependent on auditory versus visual speech (Tye-Murray, 1998).

Each picture was administered in the presence of word distractors whose onsets relative to the pictures' onsets represented either all features congruent, one feature conflicting in place of articulation, or one feature conflicting in voicing. The vowel of all distractors was always congruent. A baseline condition for each picture was provided by a vowel-onset distractor, "eagle" for /i/ vowel-nucleus pictures and "onion" for /ʌ/ vowel-nucleus pictures. Appendix Table A-1 details all of the individual picture and distractor items. Linguistic statistics, the ratings or probabilities of occurrence of various attributes of words, indicated that the test materials were of high familiarity (Coltheart, 1981; Morrison, Chappell, & Ellis, 1997; Nusbaum, Pisoni, & Davis, 1984; Snodgrass & Vanderwart, 1980), high concreteness (Coltheart, 1981; Gilhooly & Logie, 1980), high imagery (Coltheart, 1981; Cortese & Fugett, 2004; Morrison et al., 1997; Paivio, Yuille, & Madigan, 1968), high phonotactics probabilities (Vitevitch & Luce, 2004), low word frequency (Kucera & Francis as cited in Coltheart, 1981), and an early age of acquisition (Carroll & White, 1973; Dale & Fenson, 1996; Gilhooly & Logie, 1980; Morrison et al.). At least 90% of preschool/elementary school children

recognized the distractor words (Jerger, Bessonette, Davies, & Battenfield, 2007). We should note, however, that phonological effects are not affected by whether the distractor is a known word (Levelt, 2001). In fact, non-word auditory distractors produce robust phonological effects on picture-word naming (Jerger, Lai, & Marchman, 2002; Jerger, Martin, & Damian, 2002; Starreveld, 2000).

Preparation of stimuli. The distractors—and the stimuli for assessing phoneme discrimination and input phonological skills—were recorded by an 11-year-old child actor in the Audiovisual Stimulus Preparation Laboratory of the University of Texas at Dallas with recording equipment, soundproofing, and supplemental lighting and reflectors. The child was a native speaker of English with general American dialect. His full facial image and upper chest were recorded. Full facial images have been shown to yield more accurate lipreading performance (Greenberg & Bode, 1968), suggesting that facial movements other than the mouth area contribute to lipreading (Munhall & Vatikiotis-Bateson, 1998). The audiovisual recordings were digitized via a Macintosh G4 computer with Apple Fire Wire, Final Cut Pro, and Quicktime software. Color video was digitized at 30 frames/second with 24-bit resolution at 720 × 480 pixel size. Auditory input was digitized at a 22-kHz sampling rate with 16-bit amplitude resolution. The pool of utterances was edited to an average RMS level of –14 dB.

The colored pictures were pasted onto the talker's chest twice to form SOAs of –165 ms (the onset of the distractor is 165 ms or 5 frames before the onset of the picture) and +165 ms (the onset of the distractor is 165 ms or 5 frames after the picture). To be consistent with results in the literature, we defined a distractor's onset on the basis of its auditory onset. In the results reported herein, only results at the lagging SOA are considered. The rationale for selecting only the lagging SOA is that a lag of roughly 100–200 ms maximizes the interaction between output and input phonological representations (Damian & Martin, 1999; Schriefers et al., 1990). When a phonological distractor is presented about 100–200 ms before the onset of the picture, the activation of its input phonological representation is hypothesized to have decayed prior to the output phonological encoding of the picture, and the interaction is lessened (Schriefers et al., 1990). For the present study, the semantic items were also viewed as filler items. A rationale for including the semantic items is that an inconsistent relationship between the picture–distractor pairs helps participants disregard the distractors.

Experimental setup. To administer picture-word items, the video track of the Quicktime movie file was presented via a Dell Precision Workstation to a high-resolution 457-mm monitor. The outer borders of the monitor contained a colorful frame covering control buttons

for the tester (e.g., delete trial and re-administer item later), yielding an effective monitor size of about 356 mm. The inner facial image was approximately 90 mm in height with a width of about 80 mm at eye level. Previous research has shown that there are no detrimental effects on lipreading for head heights varying from 210 mm to only 21 mm, when viewed at a distance of 1 m (Jordan & Sergeant, 1998). The dimensions of each picture on the talker's chest were about 65 mm in height and 85 mm in width. The auditory track of the Quicktime movie file was routed through a speech audiometer to a loudspeaker. For audiovisual trials, each trial contained 1,000 ms of the talker's still neutral face, followed by an audiovisual utterance of one distractor word with one colored picture introduced on the chest in relation to the auditory onset of the utterance, followed by 1,000 ms of still neutral face and colored picture. For auditory-only trials, the auditory track and the picture were exactly the same as previously mentioned, but the visual track contained a still neutral face for the entire trial. Thus, the only difference between the auditory and audiovisual conditions was that the auditory items had a neutral static face and the audiovisual items had a dynamic face.

Both the computer monitor and the loudspeaker were mounted on an adjustable height table directly in front of the child at a distance of approximately 90 cm. To name each picture, children spoke into a unidirectional microphone mounted on an adjustable stand. To obtain naming times, the computer triggered a counter/timer with better than 1 ms resolution at the initiation of a movie file. The timer was stopped by the onset of the child's vocal response into the microphone, which was fed through a stereo mixing console amplifier and 1-dB step 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 distractor's auditory onset and picture onset.

Procedure

Participants were tested in two separate sessions, one for auditory testing and one for audiovisual testing. The sessions were separated by approximately 13 days for the NH group and 5 days for the HL group. The modality for first and second sessions was counterbalanced across participants for the NH group. For the HL group, however, the first session was always the audiovisual modality because pilot results indicated that recognition of the auditory distractor words was noticeably better in the children with HL who had previously undergone audiovisual testing. Word recognition for the auditory distractors remained at ceiling in the NH group regardless of the modality of the initial test session. Children sat at

a child-sized table in a double-walled sound-treated booth. A tester sat at the computer workstation, and a co-tester sat alongside the child, keeping her on task. Each trial was initiated by the tester's pushing the space bar (out of participant's sight). Children were instructed to name each picture and disregard the speech distractor. They were told that "Andy" (pseudonym) was wearing a picture on his chest, and he wanted to know what it was. They were to say the name correctly and as quickly as possible. The microphone was placed approximately 30 cm from the child's mouth without blocking her view of the monitor. Children were encouraged to speak into the microphone at a constant volume representing a clear conversational speech level. If necessary, the child's speaking level, the position of the microphone or child, and/or the setting on the 1-dB step attenuator between the microphone and VOR were adjusted to ensure that the VOR was triggering reliably. The intensity level of the distractors was approximately 70 dB SPL, as measured at the imagined center of the participant's head with a sound level meter.

Prior to beginning, picture naming was practiced. A tester showed each picture on a 5″ × 5″ card, asking children to name the picture and teaching them the target names of any pictures named incorrectly. Next, the tester flashed some picture cards quickly and modeled speeded naming. The child was asked to copy the tester for another few pictures, emphasizing that we are practicing naming the pictures as fast as we can to say them correctly. Speeded naming practice trials went back and forth between tester and child until the child was naming pictures fluently, particularly without saying "a" before names. For experimental trials, each picture was presented with each type of speech distractor at each of the two SOAs. Test items and SOAs were presented randomly within one unblocked condition (see Starreveld, 2000, for a discussion). Each test began with two practice trials. All trials judged to be flawed (e.g., lapses of attention, squirming out of position, triggering the microphone in a flawed manner) were deleted online and were re-administered after intervening items.

Data Analysis

With regard to picture naming results, the total number of trials deleted online (with replacement) represented 12% of overall trials in the HL group and 17% of overall trials in the NH group. The percentage of missing trials remaining at the end because the replacement trial was also flawed was about 3% of overall trials in both groups. Naming responses that were more than 3 *SD*s from an item's conditional mean were also discarded. This procedure resulted in the exclusion of about 1.5% of trials in both groups. In the HL group, the number of trials deleted due to mishearing the distractor represented

about 3% of overall trials. In sum, about 4.5% of trials in the NH group and 7.5% of trials in the HL group were missing or excluded. Individual data for each experimental condition were naming times averaged across the picture–distractor pairs for which the distractor onset was heard correctly. In other words, if a child with HL misheard the onset of one distractor of an experimental condition, his or her average performance for that condition was based on seven picture–distractor pairs rather than the traditional eight picture–distractor pairs. Performance on the auditory distractor onset recognition task was at ceiling in all children with NH and averaged 96.06% correct (range = 78%–100% correct) in the HL group. We should clarify that an incorrect response on the distractor repetition task could be attributed to mishearing because all of the children pronounced the onsets of our pictures and distractors correctly. In addition to controlling for mishearing distractor onsets, we also controlled for phoneme discrimination abilities by requiring that all participants pass (better than 85% correct) a phoneme discrimination task, as described previously. Performance on the auditory phoneme discrimination task was at ceiling in all children with NH and averaged 98.33% correct (range = 88%–100% correct) in the children with HL.

Data for each condition were analyzed with a mixed-design analysis of variance (ANOVA) by regression and multiple *t* tests. The problem of multiple comparisons was controlled with the false discovery rate (FDR) procedure (Benjamini & Hochberg, 1995; Benjamini, Krieger, & Yekutieli, 2006). The FDR approach controls the expected proportion of false positive findings among rejected hypotheses. A value of the approach is its demonstrated applicability to repeated-measures designs. Finally, to form a single composite variable representing age-related competencies, we analyzed the dependencies among the demographic variables in the HL group with principal component analysis (PCA; Pett, Lackey, & Sullivan, 2003). We decided to use composite variables because they provide a more precise and reliable estimate of developmental capacities than age alone, particularly in nontypically developing children (Allen, 2004). PCA creates composite scores for participants by computing a linear combination (i.e., a weighted mean) of the original variables. Standard scores for simple visual reaction time and articulatory proficiency (number of errors) were multiplied by –1 so that good performance would be positive. Variables in the PCA analysis were visual motor integration, visual perception, visual simple RT, receptive and expressive vocabulary, articulation proficiency, auditory onset, auditory rhyme, and visual-only lipreading. Age and auditory word recognition scores were not entered into the analysis in order to allow assessment of the underlying construct represented by the component by its correlation with age and auditory word recognition (i.e., hearing status).

Results

In this section, we analyze results for the demographic and picture-word tasks. The focus of the demographic analyses was (a) to compare the NH and HL groups in terms of verbal and nonverbal abilities and age and (b) to detail results of the PCA analysis in the HL group. The focus of the picture-word analyses was (a) to compare the groups in terms of the influence of hearing impairment, the type of distractor, and the modality of the distractor on performance, and (b) to detail results in the HL group in terms of age-related competencies and hearing status.

Demographic Results

Comparison between groups. Table 1 compares results on a set of nonverbal and verbal measures in the NH and HL groups. For the nonverbal measures, we attempted to select NH children whose ages and skills were comparable to those in the HL group. Chronological age averaged approximately 7;9 and ranged from 5 to 12 years of age in both groups. Visual skills were quantified by raw scores or the number of correct responses.

Results averaged 19–20 for visual motor integration and 20–24 for visual perception. The raw scores represented comparable percentile scores across groups for visual motor integration, about the 50th percentile, but not for visual perception. For the latter measure, average performance represented about the 40th percentile in the HL group and the 75th percentile in the NH group. Finally, speed of visual processing as indexed by simple RT averaged about 725 ms and ranged from approximately 485 ms to approximately 1,065 ms in both groups.

Multiple regression analyses indicated no significant difference in overall performance between groups (i.e., group membership could not be predicted from knowledge of the set of nonverbal measures). However, at least one of the nonverbal measures appeared to exhibit a difference between groups, resulting in a Group × Nonverbal Measures interaction that approached significance, $F(3, 273) = 2.453, p = .064$. Pairwise comparisons carried out with the FDR method, controlling for multiple comparisons, indicated that visual perception was significantly better in the NH group than in the HL group. Age and all other nonverbal skills did not differ significantly between groups.

Table 1. Average performance on sets of demographic variables in the groups with normal hearing (NH) versus the groups with hearing loss (HL).

Demographic variable	NH groups	HL groups
Nonverbal measures[a]		
Age (years;months)	7;8 (1;9) (5;3–12;1)	8;0 (1;8) (5;0–12;2)
Visual motor integration (raw score)	20.39 (3.98) (11–28)	19.42 (3.93) (12–26)
Visual perception** (raw score)	23.71 (3.62) (15–30)	20.16 (4.09) (9–28)
Visual simple RT (ms)	700 (132.08) (508–1045)	755 (193.37) (464–1088)
Verbal measures[b]		
Vocabulary (raw score)		
Receptive**	120.27 (24.77)	97.52 (28.82)
Expressive**	91.18 (18.81)	76.74 (21.60)
Output phonology		
Articulation (number of errors)**	1.13 (2.60)	4.64 (7.44)
Input phonology (%)		
Onset—Auditory**	98.55 (4.38)	90.32 (19.06)
Onset—Audiovisual	—	94.84 (9.62)
Rhyme—Auditory	91.45 (15.87)	89.03 (21.35)
Rhyme—Audiovisual	—	91.61 (14.85)
Word recognition (%)		
Auditory**	99.48 (1.20)	87.34 (12.23)
Audiovisual	—	93.23 (10.21)
Visual-only lipreading**	11.37 (10.68)	23.06 (15.09)

Note. Dashes in the NH group column indicate that this was not administered due to ceiling performance on auditory-only task. RT = reaction time.

[a]Parentheses show *SD* and range, respectively. [b]Parentheses show *SD*.

**Adjusted *p* < .05.

The lower part of Table 1 summarizes results on the set of verbal measures. We did not attempt to form similar verbal skills in the groups. Average raw scores for vocabulary skills differed between the NH versus HL groups, about 120 versus 97 for receptive abilities and 91 versus 77 for expressive abilities. Average performance represented about the 78th percentile in the NH group and the 40th percentile in the HL group. Output phonological skills, as estimated by articulatory proficiency, differed between groups, with an average number of errors of about 1 in the NH group and 5 in the HL group. Input phonological skills for auditory-only stimuli were slightly better in the NH group than in the HL group, averaging about 96% versus 91%, respectively. Results in the HL group may have been influenced by our entry criterion requiring all participants with HL to have at least 85% phoneme discrimination. Input phonological skills for audiovisual stimuli were slightly improved in the HL group; about 92%–95%; the audiovisual condition was not administered in the NH group due to ceiling performance in the auditory-only condition. Finally, performance for auditory word recognition in the NH and HL groups averaged about 99% and 87%, respectively. Results in the HL group yielded the following distribution: ≥90% (n = 16), 80%–89% (n = 9), 70%–79% (n = 3), 60%–69% (n = 2), and 50%–59% (n = 1). Audiovisual word recognition on the CAVET was improved in the HL group, about 93%; again, the audiovisual condition was not administered in the NH group. Visual-only lipreading of words was better in the HL group than in the NH group, about 23% versus 11%, respectively.

Multiple regression analysis of the set of verbal measures, excluding audiovisual measures which were not obtained in the NH group, indicated significantly different abilities in the groups (i.e., group membership could be predicted from knowledge of the set of measures), $F(1, 91) = 9.439, p = .003$. The pattern of results between groups was not consistent, however, resulting in a significant Group × Verbal Measures interaction, $F(6, 546) = 14.877, p < .0001$. Pairwise comparisons carried out with the FDR correction for multiple comparisons indicated that receptive and expressive vocabulary, articulation proficiency, auditory onset judgments, and auditory word recognition were significantly better in the NH group than in the HL group. Auditory rhyming skills did not differ significantly between groups. In contrast, visual-only lipreading was significantly better in the HL group than in the NH group. This latter finding in children is consistent with previous observations of enhanced lipreading ability in adults with early-onset hearing loss (Auer & Bernstein, 2007).

Finally, we should note that output phonological skills and auditory word recognition were also assessed for the pictures and distracters, respectively, of the picture-word test (not shown in Table 1). All children,

both NH and HL, pronounced all of the onsets of the pictures' names accurately. The offsets of the pictures' names were pronounced correctly by 89% of the NH group and 55% of the HL group. Of the children who mispronounced an offset, typical errors in both groups involved the /th/ in "teeth," the /mp/ in "pumpkin," the /r/ in "deer," and/or the /z/ in "pizza." Performance for auditory-only distractor recognition scored in terms of the word was consistently at ceiling in the NH group and was distributed as follows in the HL group: ≥90% (n = 21), 80%–89% (n = 7), 70%–79% (n = 2), and 60% (n = 1).

PCA analysis in the HL group. Results of a PCA analysis identified two significant (i.e., eigenvalues larger than 1) principal components. The proportion of variance extracted by these two components was 67.842%, with the first principal component accounting for the majority of the variance, 53.018%. We focused only on the first component. This factor reflected relatively high positive loadings on all of the demographic measures except visual speech, as detailed in Table 2. To investigate the underlying construct represented by the component, we assessed the correlation between the participants' composite scores for the component versus age and hearing status. The composite scores were significantly correlated with age, $r = .683, F(1, 29) = 13.977, p < .001$, but not with hearing status. The underlying construct for this component appeared to be age-related competencies. The influence of age-related competencies and hearing status in the HL subgroups was assessed after we studied the general effect of hearing impairment by comparing overall picture-word results in the NH and HL groups. Data are presented separately for the baseline, conflicting, and congruent conditions. An important preliminary step prior to addressing our research questions is to determine whether baseline naming times differed significantly as a function of the type of vowel onset, the modality of the distractor, or the status of hearing.

Picture-Word Results

Baseline condition. Figure 2 shows average naming times in the NH and HL groups for the vowel onset baseline distractors presented in the auditory and audiovisual modalities. The data were collapsed across vowels. Results in Figure 2 were analyzed to assess the following: (a) Do absolute auditory and audiovisual naming times differ for the two vowel onsets, /i/ versus /ʌ/, of the baseline distractors in the groups? (b) Do absolute baseline naming times differ for the auditory versus audiovisual distractors in either group? (c) Do absolute auditory and audiovisual baseline naming times differ in the NH versus HL groups? These issues were addressed with a mixed-design ANOVA by regression with one between-subjects factor, group (NH vs. HL), and two

Table 2. Explanation of the age-related competencies component yielded by the principal components analysis with the set of demographic variables.

Age-related competencies component	Component loadings	Less mature (n =13)	More mature (n = 18)
		Age subgroups	
Vocabulary (raw score)			
Receptive**	.183*	60.46 (14.99)	88.50 (17.77)
Expressive**	.173*	77.31 (21.30)	112.11 (24.65)
Input phonology (% correct)			
Auditory onset**	.141*	79.23 (25.32)	98.33 (5.15)
Auditory rhyme**	.152*	77.69 (29.48)	97.22 (4.61)
Output phonology			
Articulation (number errors)**	.148*	9.00 (9.45)	1.50 (3.09)
Visual skills			
Visual perception (raw score)**	.155*	17.00 (3.39)	22.44 (2.87)
Visual motor integration (raw score)**	.159*	16.38 (2.57)	21.61 (3.24)
Visual simple RT (ms) **	.169*	935.48 (116.03)	625.71 (117.78)
Mixed visual/phonology			
Visual-only lipreading (% correct)	.059	16.92 (14.07)	27.50 (14.58)
Variables not in component analysis			
Age (years;months)**		6;9 (1;3)	8;10 (1;9)
Auditory word recognition (% correct)		86.04 (14.28)	88.28 (10.85)

Note. In this table, the first two columns detail the variables in the PCA analysis and their loadings for the age-related competencies principal component. The last two columns compare average performance and *SD*s (in parentheses) on the variables in HL subgroups representing less mature or negative composite scores ($M = -0.372$) versus more mature or positive composite scores ($M = 0.269$). Average performance and *SD*s are shown at the bottom of the table for chronological age and auditory word recognition, variables that were not included in the analysis.

*Relatively high loading. **Adjusted $p < .05$.

Figure 2. Average absolute naming times in normal hearing (NH) and hearing loss (HL) groups for vowel onset, baseline distractors, presented in the auditory versus audiovisual modalities.
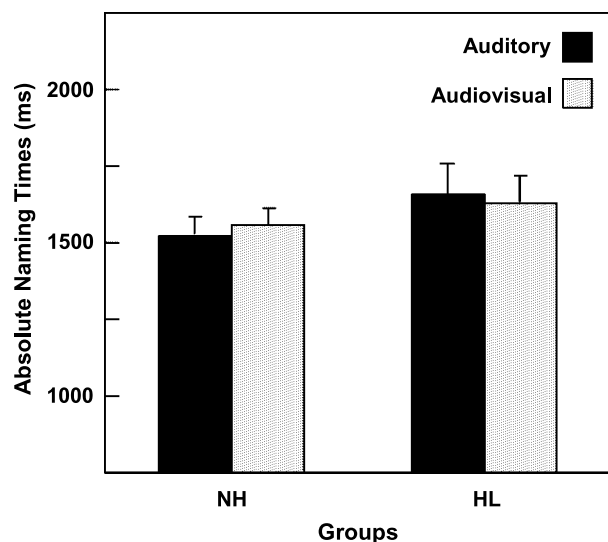
within-subjects factors, type of vowel onset (/i/ vs. /ʌ/) and modality of distractor (auditory vs. audiovisual).

Results indicated that naming times for the /i/ versus /ʌ/ onsets did not differ significantly in the groups for either modality. Naming times collapsed across modality for the /i/ and /ʌ/ onsets, respectively, averaged about 1,555 and 1,520 ms in the NH group and 1,665 and 1,620 ms in the HL group. Naming times for the auditory versus audiovisual distractors also did not differ significantly in the groups. The difference between modalities averaged only about 32 ms for each group. Finally, overall naming times did not differ significantly between groups, $p = .269$, although on average, overall performance was about 100 ms slower in the HL group. This result agrees with our previous findings for baseline verbal distractors in children with NH versus HL (Jerger, Lai, & Marchman, 2002).

In short, baseline naming times did not differ significantly as a function of the type of vowel onset, the modality of the distractor, or the status of hearing. In the subsequent results, we quantified the degree of facilitation and interference from congruent and conflicting onsets, respectively, with adjusted naming times. These data were derived by subtracting each participant's

vowel baseline naming times from his or her congruent and conflicting times, as done in our previous studies (Jerger, Lai, & Marchman, 2002; Jerger, Martin, & Damian, 2002; Jerger et al., 2009). This approach controls for developmental differences in detecting and responding to stimuli and allows each picture to serve as its own control without affecting the differences between the types of distractors.

The adjusted naming times were used to quantify picture-word performance for the conflicting and congruent conditions. Prior to proceeding, we conducted preliminary analyses to determine whether the adjusted naming times differed significantly in the groups as a function of the vowel nucleus (/i/ vs. /ʌ/) of the experimental distractors. Results indicated that adjusted naming times did not differ significantly for the two vowel nuclei in either the NH or HL groups for the conflicting condition. In contrast, naming times for the vowel nuclei did differ significantly for the congruent condition but only in the HL group. In the sections that follow, the data for the conflicting condition are collapsed across the vowel nuclei for all analyses, whereas the data for the congruent condition are analyzed separately for each vowel nucleus. To address our first primary research question concerning whether performance differed in the NH versus HL groups for the auditory and audiovisual distractors, we formulated three a priori queries for the conflicting and congruent conditions: (a) Do overall adjusted naming times differ in the NH versus HL groups? (b) Do adjusted naming times differ for the audiovisual versus auditory distractors in the groups? (c) Do all adjusted naming times differ from zero (i.e., show significant interference or facilitation)? An additional query for the

conflicting condition was whether adjusted naming times differ in the groups for the different types of conflicting distractors.

Conflicting conditions. Figure 3 shows the degree of interference as quantified by adjusted naming times in the NH and HL groups for the auditory and audiovisual distractors conflicting in place of articulation or in voicing. We addressed the issues described previously with a mixed-design ANOVA by regression with one between-subjects factor, group (NH vs. HL) and two within-subjects factors, modality of distractor (auditory vs. audiovisual) and type of distractor (conflicting in place vs. in voicing) and with t tests. The dependent variable was the degree of interference as quantified by adjusted naming times. Results indicated that the overall degree of interference differed significantly in the NH and HL groups, $F(1, 91) = 7.983$, $p = .006$. The degree of interference collapsed across modality and type of distractor was significantly greater in the NH group than in the HL group, about 147 ms versus 82 ms, respectively. The type of distractor, conflicting in place versus in voicing, did not significantly influence results in the groups. However, the modality of the distractor produced noticeably different patterns of interference in the groups, with a modality of Distractor × Group interaction that achieved borderline significance, $F(1, 91) = 3.725$, $p = .057$. As seen in Figure 3, the degree of interference for the audiovisual distractors was similar, about 128 ms, in the NH and HL groups. In contrast, the degree of interference for the auditory distractors was significantly greater in the NH group, about 154 ms versus only 50 ms in the HL group. Because the interaction between the modality of the distractor and the groups was only of borderline significance, we probed the strength of

Figure 3. Distractors conflicting in place of articulation or in voicing. Degree of interference for auditory versus audiovisual modalities as quantified by adjusted naming times in NH and HL groups. The zero baseline of the ordinate represents naming times for vowel onset baseline distractors (see Figure 2). A larger positive value indicates more interference.
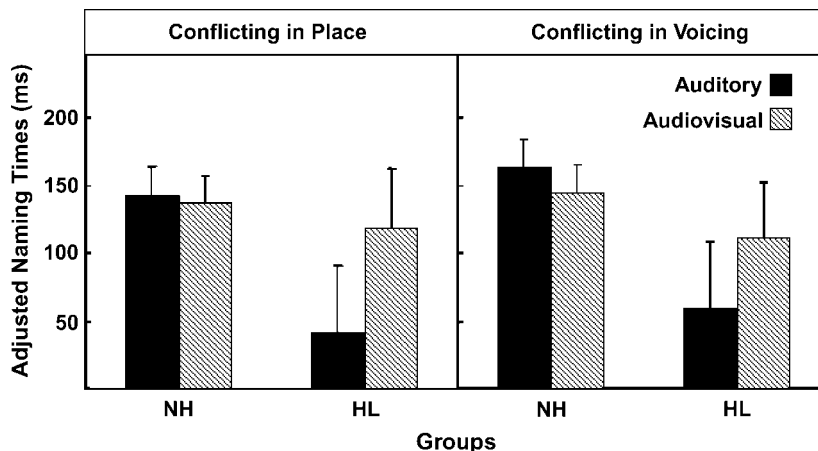
the effect with multiple *t* tests. Results of the FRD approach indicated that the auditory modality produced significantly more interference in the NH group than in the HL group for both types of conflicting distractors. The pairwise comparisons for the audiovisual distractors were not significant. Finally, we carried out multiple *t* tests to determine whether each mean differed significantly from zero, indicating significant interference. Results of the FDR method indicated that the audiovisual distractors produced significant interference in all groups, but the auditory distractors produced significant interference in the NH group only.

The pattern of results in the HL group (i.e., significant interference from the audiovisual conflicting distractors only) suggests a relatively greater influence of visual speech on phonological processing by children with HL. As we noted earlier, however, results from 4 to 14 years of age in children with NH showed a U-shaped developmental function with a period of transition from about 5 years to 9 years that did not show any influence of visual speech on performance. Developmental shifts in the influence of visual speech on performance in the NH group broach the possibility of developmental shifts in the pattern of results seen in Figure 3 in the children with HL. To explore this possibility, we formed HL subgroups with less mature and more mature age-related competencies from the composite scores expressing the first principal component.

Age-related competencies versus performance for conflicting conditions in children with HL. Our second primary research question concerned whether performance differs in HL subgroups representing different levels of age-related competencies. The first two columns of Table 2 summarize the variables in the PCA analysis and their loadings for the age-related competencies principal component. The last two columns compare results on the variables and on chronological age and auditory word recognition in HL subgroups representing less mature and more mature age-related competencies. The less mature subgroup ($n = 13$) included the 12 children with negative composite scores and 1 child whose composite score was essentially zero. The more mature subgroup ($n = 18$) included the 16 children with positive composite scores and 2 children with near-zero composite scores. The children with near-zero composite scores were assigned to the subgroup whose regression equation minimized the deviation of their composite scores from the regression line. Results of *t* tests controlling for multiplicity with the FDR method indicated that the subgroups differed significantly on all measures in Table 2 excepting visual-only lipreading and auditory word recognition. It seems important to stress the latter result indicating that the HL subgroups did not differ on our proxy variable for hearing status.

Figure 4 shows the degree of interference as quantified by adjusted naming times in the less mature and more mature HL subgroups for auditory and audiovisual distractors conflicting in place of articulation or in voicing. We again analyzed results with a mixed-design ANOVA by regression with one between-subjects factor, group (less mature vs. more mature) and two within-subjects factors, modality of distractor (auditory vs.

Figure 4. Distractors conflicting in place of articulation or in voicing: Age-related competency. Degree of interference for auditory versus audiovisual modalities as quantified by adjusted naming times in HL subgroups representing less mature (M = 6;9 [years;months]) and more mature (M = 8;10) age-related competencies. The zero baseline of the ordinate represents naming times for vowel onset baseline distractors (see Figure 2). A larger positive value indicates more interference. Hearing status, as defined by auditory word recognition, did not differ in the subgroups.
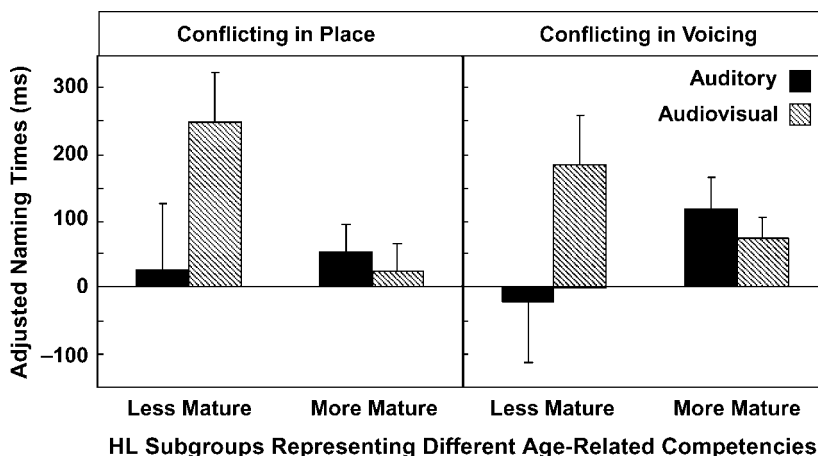
audiovisual) and type of distractor (conflicting in place vs. in voicing) and with t tests. The dependent variable was the degree of interference as quantified by adjusted naming times. Results indicated that the modality of the distractor produced significantly different patterns of interference in the groups, with a significant Modality of Distractor × Group interaction, $F(1, 29) = 4.657, p = .039$. With regard to the less mature subgroup, results showed an influence of visual speech on performance with significantly greater interference from the audiovisual distractors than the auditory distractors. In contrast, the more mature subgroup did not show a greater influence of visual speech on performance. No other significant effects were observed. Multiple t tests were conducted to determine whether each mean differed significantly from zero, which would mean significant interference. When controlling for multiplicity with the FDR method, only performance for the audiovisual distractors conflicting in place in the less mature subgroup achieved significance. Previous results on the Children's Cross-Modal Picture-Word Task with auditory distractors have consistently shown a greater degree of interference and/or facilitation for younger children than for older children (Brooks & MacWhinney, 2000; Jerger, Lai, & Marchman, 2002; Jerger, Martin, & Damian, 2002; Jerger et al., 2009; Jescheniak, Hahne, Hoffmann, & Wagner, 2006).

The pattern of results in Figure 3 indicates developmental shifts in performance in children with significant differences in age-related competencies and no significant differences in hearing status, as documented in Table 2. Results support the idea that speech representations are structured disproportionally in terms of visual speech in less mature children with HL. With increasing maturity, however, speech representations appear to become better structured in terms of auditorily encoded information. It also seems important to consider the converse pattern of results in these children. Thus, we posed an additional research question addressing whether performance differed in HL subgroups with significant differences in hearing status but no significant differences in age. Toward this end, we formed HL subgroups with significantly poorer and better auditory word recognition but comparable age-related competencies. A caveat in this regard is that due to our stringent auditory criteria for participation in this study, most children had good auditory word recognition scores.

Hearing status versus performance for conflicting conditions in children with HL. Table 3 compares performance in the HL subgroups representing poorer ($n = 7$) and better ($n = 14$) hearing statuses as operationally defined by auditory word recognition scores. The poorer subgroup contained all children with auditory word recognition scores of less than 85% correct. The better subgroup contained children whose auditory word recognition scores were ≥85% and whose age-related competencies

corresponded to those in the poorer subgroup. Data in Table 3 compare results in the subgroups for (a) the two grouping variables, auditory word recognition scores and composite scores expressing the age-related competencies component; (b) the individual variables making up the age-related competencies component; and (c) chronological age. Results of t tests controlled for multiple comparisons with the FDR method indicated that the subgroups differed significantly only on auditory word recognition scores. In particular, the HL subgroups did not differ on any of the age-related competencies measures nor on chronological age, per se.

Figure 5 compares the degree of interference, as quantified by adjusted naming times, for the auditory and audiovisual distractors conflicting in place of articulation or in voicing in the subgroups with poorer and better hearing status. Methodical statistical analyses were not attempted due to the small number of children in the HL subgroups and the large variability. Results of t tests controlling for multiplicity indicated that the poorer and better hearing subgroups differed significantly on the audiovisual distractors conflicting in voicing, with results for the auditory distractors conflicting in voicing approaching significance. No other pairwise comparisons achieved significance when controlling for multiple comparisons. Relative to children with HL who had poorer hearing status, children with HL who had better hearing status showed significantly more interference for a relatively strong auditory cue, voicing (Miller & Nicely, 1955). As we turn to analyzing our first primary research question concerning the general effect of hearing impairment on performance for the congruent condition in the total HL versus NH groups, we should stress that results for the conflicting distractors did not differ as a function of the vowel nucleus.

Congruent condition. Figure 6 summarizes adjusted naming times in the NH and HL groups for auditory and audiovisual congruent distractors. Results are plotted separately for each vowel nucleus. We posed four a priori queries: (a) Do overall adjusted naming times differ in the NH versus HL groups? (b) Do adjusted naming times differ for congruent onset distractors with /i/ versus /ʌ/ vowel nuclei for both modalities in the groups? (c) Do adjusted naming times differ for audiovisual versus auditory distractors in the groups? (d) Do all adjusted naming times differ from zero? We again conducted a mixed-design ANOVA by regression with one between-subjects factor, group (NH vs. HL), and two within-subjects factors, modality of distractor (auditory vs. audiovisual) and type of vowel nucleus (/i/ vs. /ʌ/), and t tests. Results indicated that overall adjusted naming times differed significantly between groups, $F(1, 91) = 16.079, p = .0001$, and between vowel nuclei, $F(1, 91) = 39.256, p < .0001$. Further, the vowel nuclei produced different patterns of results in the groups, with a significant Vowel

Table 3. Explanation of subgroup formation and performance in subgroups of children with HL representing better and poorer hearing status.

	Hearing status subgroup	
	Poorer (*n* = 7)	Better (*n* = 14)
Grouping variables		
Auditory word recognition (% correct)**	68.71 (11.76)	92.96 (4.86)
Range	50 to 81	85 to 100
Age-related competencies (composite score)	–0.14 (0.607)	–0.06 (0.241)
Range	–4.24 to –1.11	–4.01 to –0.51
Age-related competencies component		
Vocabulary		
Receptive (raw score)	87.43 (33.07)	90.64 (20.14)
Expressive (raw score)	74.86 (29.83)	69.29 (14.26)
Input phonology		
Auditory onset (% correct)	77.14 (26.90)	94.29 (11.58)
Auditory rhyme (% correct)	67.14 (36.84)	94.29 (7.56)
Output phonology		
Articulation (number errors)	5.14 (9.24)	3.29 (5.58)
Visual skills		
Visual perception (raw score)	20.43 (3.95)	18.79 (3.81)
Visual motor integration (raw score)	19.00 (3.74)	18.07 (3.22)
Visual simple RT (ms)	791.33 (239.37)	776.96 (174.84)
Mixed visual/phonology		
Lipreading	28.57 (18.19)	19.29 (9.58)
Variables not in component analysis		
Chronological age (years;months)	7;10 (1;5)	7;5 (1;2)
Range	(5;7 to 9;7)	(5;6 to 9;3)

Note. This table shows average performance and *SD*s (shown in parentheses) in HL subgroups representing poorer and better hearing status, as quantified by auditory word recognition but comparable age-related competencies. Results are detailed for the two grouping variables and the demographic measures comprising the age-related competencies component. The range is also provided for the two grouping variables as well as for chronological age.

**Adjusted *p* < .05.

Figure 5. Distractors conflicting in place of articulation or in voicing: Hearing status. Degree of interference for auditory versus audiovisual modalities as quantified by adjusted naming times in HL subgroups representing poorer (<85% correct) and better (≥85% correct) hearing status as defined by auditory word recognition. The zero baseline of the ordinate represents naming times for vowel onset baseline distractors (see Figure 2). A larger positive value indicates more interference. Age-related competencies did not differ in the subgroups.
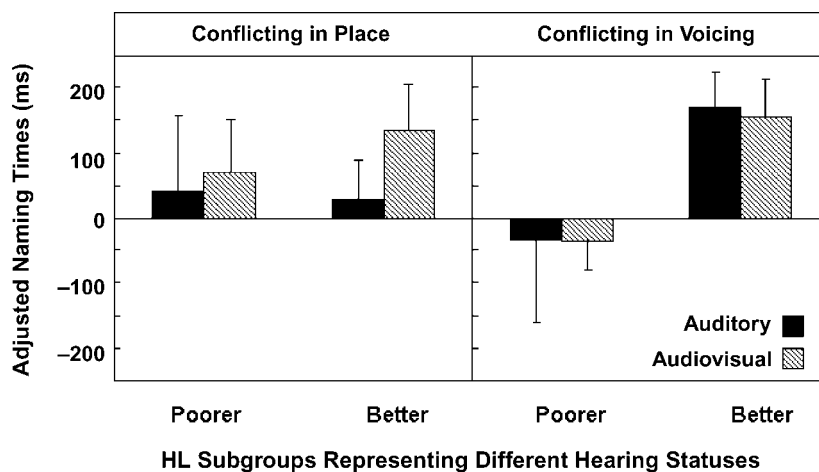
Figure 6. Congruent distractors. Degree of facilitation or interference for auditory versus audiovisual modalities as quantified by adjusted naming times in NH and HL groups. Data are presented separately for congruent onsets coupled with /i/ versus /ʌ/ vowel nuclei. The zero baseline of the ordinate represents naming times for vowel onset baseline distractors (see Figure 2). A larger negative value indicates more facilitation. A larger positive value indicates more interference.
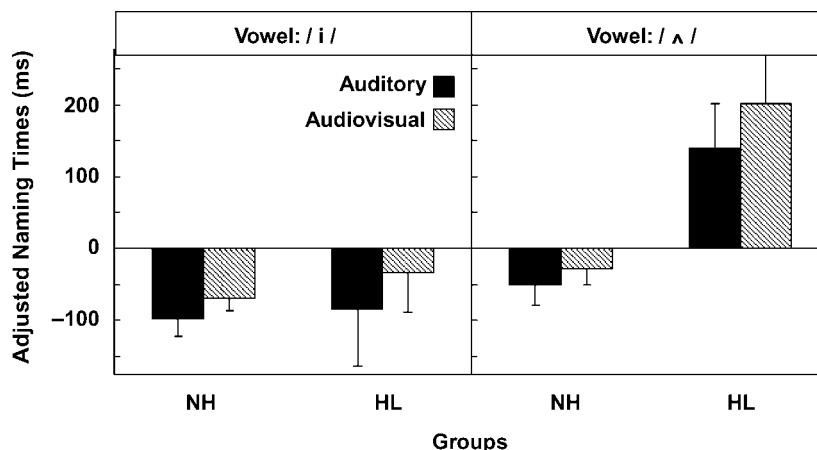
Nucleus × Group interaction, $F(1, 91) = 19.109, p < .0001$. In the NH group, congruent onset distractors with differing vowel nuclei consistently showed negative adjusted naming times of about −40 to −80 ms. Findings in the HL group, however, produced a mixed pattern of results. Average performance showed negative adjusted naming times of about −60 ms for distractors with an /i/-vowel nucleus and positive naming times of about 170 ms for distractors with an /ʌ/-vowel nucleus.

With regard to the modality of the distractor, overall results differed significantly for the auditory versus audiovisual distractors, $F(1, 91) = 3.995, p = .049$. No significant interactions with group membership or type of vowel nucleus were observed. The difference between overall performance for the auditory versus audiovisual inputs, ignoring the sign of the difference, was about 25–30 ms in the NH group and about 50–60 ms in the HL group. The opposing directions of average naming times in the HL group complicated the interpretation of the statistical results for the modality of the distractor. To probe this effect, we conducted multiple t tests for each condition in each group. None of the pairwise comparisons achieved statistical significance. Finally, we carried out multiple t tests to determine whether each mean differed significantly from zero. Results of the FDR method indicated the following: (a) for the /i/ vowel, results for the NH group showed significant facilitation for both auditory and audiovisual distracters, whereas results in the HL group did not achieve statistical significance for either modality; (b) for the /ʌ/ vowel, results for the HL group showed significant interference for the audiovisual distractor with results for the auditory

distractor approaching significance, whereas results in the NH group did not achieve statistical significance for either modality. The vowel nucleus interactions reduced the number of distractors to only four for each type of vowel condition, clearly lessening the resolution and stability of individual results for forming small subgroups of children. Thus, subgroups were not formed to investigate age-related competencies and hearing status.

Congruency interference effect. In an effort to probe the unexpected interference effect for the /ʌ/-vowel nucleus in the HL group, we collapsed results across the auditory and audiovisual distractors. HL subgroups were formed to represent children with the fastest naming times ($n = 15$) versus children with the slowest naming times ($n = 16$). Average performance was −56.55 ms in the fastest subgroup and 384.24 ms in the slowest subgroup. Exploratory data analyses suggested an association between performance and output phonological skills as quantified by the number of articulation errors on the Goldman–Fristoe Test of Articulation (Goldman & Fristoe, 2000). Whereas the fastest subgroup had only two errors, the slowest subgroup had seven errors, $t(29) = 1.73, p = .094$. Inspection of the individual data in the fastest subgroup indicated an outlier child who had an unusually high number of articulation errors for either subgroup ($n = 20$). Elimination of this child reduced the average number of errors in the fastest subgroup to one and resulted in a significant difference between the subgroups, $t(29) = 2.43, p = .022$. Finally, a caveat with regard to these results is that the Goldman–Fristoe Test does not assess all possible articulation errors. Thus, the number of errors mentioned previously is specific to the

subset of consonants and consonant blends tested on the Goldman–Fristoe.

Discussion

This research compared the influence of visual speech on phonological processing by children with NH or HL. Relative to the NH group, the HL group had similar nonverbal skills (excepting visual perception) and poorer verbal skills (excepting auditory rhyming judgments and visual-only lipreading). Our original goal of developing two groups with consistently equivalent nonverbal skills proved difficult. More specifically, the lowest visual perceptual score in the NH group was the 23rd percentile, with only 15% of children below the 50th percentile. In contrast, 32% of the children in the HL group were below the 23rd percentile, including 16% below the 10th percentile. In our battery of tests, we had both dynamic and static measures of visual perception, respectively a lipreading test and a task requiring children to match a target geometric shape from a closed set of alternatives. Relative to the NH group, the HL group showed advanced dynamic visual perception coupled with poorer static visual perception. This finding contrasts with results of our previous studies in children with similar audiological characteristics to the current participants excluding the profound losses. Those studies repeatedly observed normal static visual perception (e.g., Jerger, Grimes, Tran, Chen, & Martin, 1997; Jerger, Lai, & Marchman, 2002; Jerger et al., 1993). A difference between the current and previous studies, however, is that our earlier studies assessed visual perception with different measures, namely the Southern California Figure-Ground Visual Perception Test (Ayres, 1978), the Block Design subtest of the Wechsler Intelligence Scale for Children–Revised (Wechsler, 1974), and the Wechsler Adult Intelligence Scale–Revised (Wechsler, 1981). Further research is needed to explore this thought-provoking difference between static and dynamic visual perception in the current children with HL. Our finding of superior lipreading ability in the children with HL agrees with Auer and Bernstein's (2007) observations in adults with prelingual HL. The present research extends their findings by suggesting that the degree of hearing loss may be less than profound and auditory-only word recognition may be good (recall that the HL group averaged 87% correct on our measure of auditory word recognition), and this performance advantage may still be present. Moreover, the present findings suggest that the performance advantage emerges in childhood.

Our first primary research question addressed whether picture-word performance in the NH versus HL groups differed for auditory and audiovisual distractors whose onsets relative to the pictures' onset were either congruent or conflicting in place of articulation or voicing. Results showed both similarities and differences between the groups. With regard to the congruent onset distractors, results between the groups differed as a function of the vowel nucleus. For the /i/ vowel nucleus, both the NH and HL groups showed negative average naming times for both the audiovisual and auditory modalities. Only performance in the NH group, however, exhibited a significant degree of facilitation. For the /ʌ/ vowel nucleus, the NH group showed negative average naming times, whereas the HL group showed positive average naming times. The degree of interference in the HL group achieved significance for the audiovisual distractors and approached significance for the auditory distractors. Exploratory data analyses suggested that the unexpected "congruency interference effect" may be associated with output phonological skills. Children with faster naming times showed fewer errors of articulation on the Goldman–Fristoe Test than children with slower naming times.

The interference effect for congruent onset distractors with a /ʌ/ vowel nucleus in the children with HL, particularly those with poorer articulatory proficiency, surprised us. Why would the onset of naming (e.g., the picture "bus") be slowed for a congruent onset distractor with a /ʌ/ vowel nucleus (e.g., "buckle") relative to the baseline distractor with a /ʌ/ vowel onset (e.g., "onion")? Why would our set of congruent onsets interfere with naming when coupled with /ʌ/ but not when coupled with /i/? Interestingly, the McGurk effect is also sensitive to vowel context, with an /i/ vowel producing strong McGurk effects and an /u/ vowel producing only weak McGurk effects (Green, 1996; Green, Kuhl, & Meltzoff, 1988). The unusual pattern of interference produced by congruent onset related distractors with a /ʌ/ vowel nucleus stands in stark contrast to other results in the literature (Brooks & MacWhinney, 2000; Jerger, Lai, & Marchman, 2002; Jerger, Martin, & Damian, 2002). For example, in our previous Children's Cross-Modal Picture-Word study with auditory nonsense syllable distractors (picture = "bus"; distractor = "/bʌ/"), results showed facilitation in a sample of children with HL who were successful hearing aid users and who had similar auditory skills, including good phoneme discrimination (Jerger, Lai, & Marchman, 2002). Thus, findings in children with HL across our two studies for congruent distractors with a /ʌ/ vowel nucleus suggest that word distractors interfere with naming and nonword distractors facilitate naming ("buckle" interferes and "bʌ" facilitates). Further research is planned to assess (a) whether the unusual congruency interference effect is associated with the lexical status of the distractors or with their greater phonotactic complexity and (b) whether other CV pairs also produce congruency interference effects. A possibility is that the congruency interference effect in children with HL and poorer articulatory proficiency

might be related to a tendency to neutralize vowels so they sound more schwa-like or more like the vowel /ʌ/. Previous research (Angelocci, Kopp, & Holbrook, 1964; Monsen, 1976; Smith, 1975) has reported this tendency in children with severe to profound hearing loss and poor speech recognition, but studies are lacking in the literature for children with demographic characteristics similar to those of our participants. Further research is needed in this area.

With regard to performance in children with NH versus HL, results for both the conflicting in place or in voicing distractors indicated that the degree of interference produced by audiovisual distractors did not differ significantly between groups. In contrast, the degree of interference produced by auditory distractors differed strikingly between groups for the conflicting conditions. Results in the HL group did not show any significant auditory interference for either of the conflicting distracters, whereas results in the NH group showed consistently robust interference. With regard to the influence of visual speech on phonological processing, the outcome of significant interference from audiovisual distractors coupled with no significant interference from auditory distractors in the HL group represents a sharp difference from the pattern of results in the NH group.

Again, the current group of NH children represents the age range within which there is a temporary loss of the influence of visual speech on performance. This raised a research question about whether the children with HL might also show developmental shifts in the pattern of results as a function of age-related competencies. To explore this possibility, we formed subgroups of children with HL representing more or less mature age-related skills but no significant differences in hearing status. Results showed an influence of visual speech on performance in the less mature subgroup but not in the more mature subgroup. Data suggested that the children with HL show the same developmental shifts that we observed in children with NH, but at a delayed age.

Conclusions

1. Phonological processing is more influenced by visual speech in children with HL than in children with NH.

2. In children with HL, congruent distractors with a /ʌ/ vowel nucleus produced an aberrant interference effect, slowing naming particularly in children with poorer output phonology as indexed by articulatory proficiency. Further research is needed to develop a theory of and understanding of this finding.

3. In children with HL, conflicting distractors produced the expected interference effect on naming

characterized by age-related change. Participants initially showed a disproportionately large influence of visual speech on phonological processing coupled with a negligible influence of auditory speech on performance. With increasing age, this pattern evolved into a robust influence of auditory speech on phonological processing with no benefit on performance from adding visual speech. To the extent that the apparent loss of visual speech on performance in older children is a positive indication that relevant knowledge systems, particularly phonology, are transitioning into more highly developed forms, these results are supportive of normal but delayed development of some aspects of phonology in children with HL.

4. Results with conflicting distractors were consistent with the idea that speech representations in children with HL appear to be initially structured disproportionally in terms of visual speech and/or motoric/articulatory codes, eventually becoming sufficiently specified in terms of auditorily encoded information to support auditory phonological processing.

5. Overall, conflicting distractors produced more nearly normal patterns of phonological processing than congruent distractors.

Acknowledgments

References

Abraham, S. (1989). Using a phonological framework to describe speech errors of orally trained, hearing-impaired schoolagers. *Journal of Speech and Hearing Disorders, 54,* 600–609.

Allen, M. (2004). *Understanding regression analysis.* New York: Plenum.

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

Angelocci, A., Kopp, G., & Holbrook, A. (1964). The vowel formants of deaf and normal-hearing eleven- to fourteen-year-old boys. *Journal of Speech and Hearing Disorders, 29,* 156–170.

Auer, E., & Bernstein, L. (2007). Enhanced visual speech perception in individuals with early-onset hearing impairment. *Journal of Speech, Language, and Hearing Research, 50,* 1157–1165.

Ayres, J. (1978). *Southern California Figure Ground Visual Perception Test* (6th ed.). Los Angeles: Western Psychological Services.

Beery, K., Buktenica, N., & Beery, N. (2004). *Beery VMI: Development test of visual-motor integration* (5th ed.). Minneapolis, MN: NCS Pearson.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological), 57,* 289–300.

Benjamini, Y., Krieger, A., & Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika, 93,* 491–507.

Bertelson, P., & de Gelder, B. (2004). The psychology of multimodal perception. In C. Spence & J. Driver (Eds.), *Crossmodal space and crossmodal attention* (pp. 141–177). Oxford, England: Oxford University Press.

Bird, J., Bishop, D., & Freeman, N. (1995). Phonological awareness and literacy development in children with expressive phonological impairments. *Journal of Speech and Hearing Research, 38,* 446–462.

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. *Journal of Child Psychology and Psychiatry, 42,* 329–340.

Brooks, P., & MacWhinney, B. (2000). Phonological priming in children's picture naming. *Journal of Child Language, 27,* 335–366.

Brownell, R. (2000). *Expressive One-Word Picture Vocabulary Test.* Novato, CA: Academic Therapy Publications.

Campbell, R. (1988). Tracing lip movements: Making speech visible. *Visible Language, 22,* 32–57.

Carroll, J. B., & White, M. N. (1973). Age-of-acquisition norms for 220 picturable nouns. *Journal of Verbal Learning and Verbal Behavior, 12,* 563–576.

Coltheart, M. (1981). *The MRC Psycholinguistic Database.* Retrieved August 9, 2006, from http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm.

Conrad, R. (1977). Lipreading by deaf and hearing children. *British Journal of Educational Psychology, 47,* 60–65.

Cortese, M., & Fugett, A. (2004). Imageability ratings for 3,000 monosyllabic words. *Behavioral Research Methods, Instruments, & Computers, 36,* 384–387.

Dale, P., & Fenson, L. (1996). Lexical development norms for young children. *Behavioral Research Methods, Instruments, & Computers, 28,* 125–127.

Damian, M., & Martin, R. (1999). Semantic and phonological codes interact in single word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 345–361.

Denes, P., & Pinson, E. (1993). *The speech chain. Physics and biology of spoken language.* London: W. H. Freeman.

Desjardins, R., Rogers, J., & Werker, J. (1997). An exploration of why preschoolers perform differently than do adults in audiovisual speech perception tasks. *Journal of Experimental Child Psychology, 66,* 85–110.

Dodd, B. (1977). The role of vision in the perception of speech. *Perception, 6,* 31–40.

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

Dodd, B. (1980). Interaction of auditory and visual information in speech perception. *British Journal of Psychology, 71,* 541–549.

Dodd, B. (1987). The acquisition of lipreading skills by normally hearing children. In B. Dodd & R. Campbell (Eds.), *Hearing by eye: The psychology of lipreading* (pp. 163–175). Hillsdale, NJ: Erlbaum.

Dodd, B., & Burnham, D. (1988). Processing speechread information. *The Volta Review: New Reflections on Speech-reading, 90,* 45–60.

Dodd, B., Holm, A., Hua, Z., & Crosbie, S. (2003). Phonological development: A normative study of British English-speaking children. *Clinical Linguistics & Phonetics, 17,* 617–643.

Dodd, B., & So, L. (1994). The phonological abilities of Cantonese-speaking children with hearing loss. *Journal of Speech and Hearing Research, 37,* 671–679.

Dunn, L., & Dunn, L. (1997). *The Peabody Picture Vocabulary Test–Third Edition.* Circle Pines, MN: AGS.

Dupont, S., Aubin, J., & Menard, L. (2005). A study of the McGurk effect in 4- and 5-year-old French Canadian children. *ZAS Papers in Linguistics, 40,* 1–17.

Fernald, A., Swingley, D., & Pinto, J. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development, 72,* 1003–1015.

Gershkoff-Stowe, L., & Thelen, E. (2004). U-shaped changes in behavior: A dynamic systems perspective. *Journal of Cognition and Development, 5,* 11–36.

Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation, 12,* 395–427.

Goldman, R., & Fristoe, M. (2000). *Goldman–Fristoe Test of Articulation* (2nd ed.). Circle Pines, MN: AGS.

Gow, D., Melvold, J., & Manuel, S. (1996, October). How word onsets drive lexical access and segmentation: Evidence from acoustics, phonology, and processing. In *Proceedings of the Fourth International Conference on Spoken Language Processing* (pp. 66–69), Philadelphia, PA.

Green, K. (1996). The use of auditory and visual information in phonetic perception. In D. Stork & D. Hennecke (Eds.), *Speech reading by humans and machines. Models, systems, and applications* (pp. 55–78). Berlin: Springer-Verlag.

Green, K., Kuhl, P., & Meltzoff, A. (1988). Factors affecting the integration of auditory and visual information in speech: The effect of vowel environment. *The Journal of the Acoustical Society of America, 84,* S155.

Greenberg, H., & Bode, D. (1968). Visual discrimination of consonants. *Journal of Speech and Hearing Research, 11,* 869–874.

Hockley, N., & Polka, L. (1994). A developmental study of audiovisual speech perception using the McGurk paradigm. *The Journal of the Acoustical Society of America, 96,* 3309.

Hollingshead, A. (1975). *Four-Factor Index of Social Status.* New Haven, CT: Yale University.

Jerger, S. (2007). Current state of knowledge: Perceptual processing by children with hearing impairment. *Ear & Hearing, 28,* 754–765.

Jerger, S., Bessonette, S. J., Davies, K. L., & Battenfield, K. (2007). Unpublished data. Dallas, TX: University of Texas at Dallas.

Jerger, S., Damian, M., Spence, M., Tye-Murray, N., & Abdi, H. (2009). Developmental shifts in children's sensitivity to visual speech: A new multimodal picture word task. *Journal of Experimental Child Psychology, 102,* 40–59.

Jerger, S., Grimes, A., Tran, T., Chen, C., & Martin, R. (1997). Childhood hearing impairment: Processing dependencies in multidimensional speech perception for an auditory level of analysis. *Ear and Hearing, 18,* 513–535.

Jerger, S., Lai, L., & Marchman, V. (2002). Picture naming by children with hearing loss: II. Effect of phonologically-related auditory distractors. *Journal of the American Academy of Audiology, 13,* 478–492.

Jerger, S., Martin, R., & Damian, M. (2002). Semantic and phonological influences on picture naming by children and teenagers. *Journal of Memory and Language, 47,* 229–249.

Jerger, S., Stout, G., Kent, M., Albritton, E., Loiselle, L., Blondeau, R., & Jorgensen, S. (1993). Auditory Stroop effects in hearing-impaired children. *Journal of Speech and Hearing Research, 36,* 1083–1096.

Jescheniak, J., Hahne, A., Hoffmann, S., & Wagner, V. (2006). Phonological activation of category coordinates during speech planning is observable in children but not in adults: Evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 373–386.

Jordan, T., & Bevan, K. (1997). Seeing and hearing rotated faces: Influences of facial orientation on visual and audiovisual speech recognition. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 388–403.

Jordan, T., & Sergeant, P. (1998). Effects of facial image size on visual and audio-visual speech recognition. In R. Campbell, B. Dodd, & D. Burnham (Eds.), *Hearing by eye. II: Advances in the psychology of lipreading and auditory-visual speech* (pp. 155–176). Hove, United Kingdom: Psychology Press.

Levelt, W. (2001). Spoken word production: A theory of lexical access. *Proceedings of the National Academy of Sciences, USA, 98,* 13464–13471.

Levelt, W., Roelofs, A., & Meyer, A. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,* 1–75.

Levelt, W., Schriefers, H., Vorberg, D., Meyer, A., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review, 98,* 122–142.

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

Marslen-Wilson, W., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 576–585.

Massaro, D. (1984). Children's perception of visual and auditory speech. *Child Development, 55,* 1777–1788.

Massaro, D., Thompson, L., Barron, B., & Laren, E. (1986). Developmental changes in visual and auditory contributions to speech perception. *Journal of Experimental Child Psychology, 41,* 93–113.

McGregor, K., Friedman, R., Reilly, R., & Newman, R. (2002). Semantic representation and naming in young children. *Journal of Speech, Language, and Hearing Research, 45,* 332–346.

McGurk, H., & McDonald, M. (1976, December 23). Hearing lips and seeing voices. *Nature, 264,* 746–748.

Merikle, P., & Reingold, E. (1991). Comparing direct (explicit) and indirect (implicit) measures to study unconscious memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 224–233.

Miller, G., & Nicely, P. (1955). An analysis of perceptual confusions among some English consonants. *The Journal of the Acoustical Society of America, 27,* 338–352.

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). Hillsdale, NJ: Erlbaum.

Moeller, M., Tomblin, J., Yoshinaga-Itano, C., Connor, C., & Jerger, S. (2007). Current state of knowledge: Language and literacy of children with hearing impairment. *Ear & Hearing, 28,* 740–753.

Monsen, R. B. (1976). Normal and reduced phonological space: The production of English vowels by deaf adolescents. *Journal of Phonetics, 4,* 189–198.

Morrison, C., Chappell, T., & Ellis, A. (1997). Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables. *Quarterly Journal of Experimental Psychology, 50A*(3), 528–559.

Munhall, K., & Vatikiotis-Bateson, E. (1998). The moving face during speech communication. In R. Campbell, B. Dodd, & D. Burnham (Eds.), *Hearing by eye. II: Advances in the psychology of lipreading and auditory-visual speech* (pp. 123–139). Hove, United Kingdom: Psychology Press.

Nusbaum, H. C., Pisoni, D. B., & Davis, C. K. (1984). Sizing up the Hoosier Mental Lexicon: Measuring the familiarity of 20,000 words. *Research on Speech Perception Progress Report No. 10,* 357–376.

Oller, D., & Eilers, R. (1981). A pragmatic approach to phonological systems of deaf speakers. In N. Lass (Ed.), *Speech and language: Advances in basic research and practice* (Vol. 6, pp. 103–141). New York: Academic Press.

Paivio, A., Yuille, J., & Madigan, S. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology: Monograph Supplement, 76,* 1–25.

Peltzer, L. (1997). *Oral Motor Function Questionnaire.* Unpublished manuscript, Baylor College of Medicine.

Pett, M., Lackey, N., & Sullivan, J. (2003). *Making sense of factor analysis.* Thousand Oaks, CA: Sage.

Rader, K. (1977). *Rader Near Point Vision Test.* Tulsa, OK: Modern Education.

Rayner, K., & Springer, C. (1986). Graphemic and semantic similarity effects in the picture-word interference task. *British Journal of Psychology, 77,* 207–222.

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

Ruffin-Simon, C. (1983). Relationship between production and perception of speech: Some developmental issues. In I. Hochberg, H. Levitt, & M. Osberger (Eds.), *Speech of the hearing impaired. Research, training, and personnel preparation* (pp. 233–249). Baltimore: University Park Press.

Schriefers, H., Meyer, A., & Levelt, W. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language, 29,* 86–102.

Sekiyama, K., & Burnham, D. (2004, October). Issues in the development of auditory-visual speech perception: Adults, infants, and children. In *Interspeech 2004: Eighth International Conference on Spoken Language Processing,* (pp. 1137–1140), Jeju Island, Korea.

Smit, A., Hand, L., Freilinger, J., Bernthal, J., & Bird, A. (1990). The Iowa Articulation Norms Project and its Nebraska replication. *Journal of Speech and Hearing Disorders, 55,* 779–798.

Smith, C. (1975). Residual hearing and speech production in the deaf. *Journal of Speech and Hearing Research, 19,* 795–811.

Smith, L., & Thelen, E. (2003). Development as a dynamic system. *Trends in Cognitive Sciences, 7,* 343–348.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 174–215.

Stanovich, K., Cunningham, A., & Cramer, B. (1984). Assessing phonological awareness in kindergarten children: Issues of task comparability. *Journal of Experimental Child Psychology, 38,* 175–190.

Starreveld, P. (2000). On the interpretation of onsets of auditory context effects in word production. *Journal of Memory and Language, 42,* 497–525.

Stoel-Gammon, C., & Dunn, C. (1985). *Normal and disordered phonology in children.* Baltimore: University Park Press.

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

Tye-Murray, N. (1998). *Foundations of aural rehabilitation.* San Diego, CA: Singular.

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. *The Journal of the Acoustical Society of America, 98,* 2454–2460.

Tye-Murray, N., Spencer, L., & Woodworth, G. (1995). Acquisition of speech by children who have prolonged cochlear implant experience. *Journal of Speech and Hearing Research, 38,* 327–337.

Vitevitch, M. S., & Luce, P. A. (2004). A web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, Instruments, & Computers, 36,* 481–487.

Wechsler, D. (1974). *Wechsler Intelligence Scale for Children–Revised.* New York: The Psychological Corporation.

Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale–Revised.* New York: The Psychological Corporation.

Weikum, W., Vouloumanos, A., Navarra, J., Soto-Faraco, S., Sebastian-Galles, N., & Werker, J. (2007, May 25). Visual language discrimination in infancy. *Science, 316,* 1159.

Wightman, F., Kistler, D., & Brungart, D. (2006). Informational masking of speech in children: Auditory-visual integration. *The Journal of the Acoustical Society of America, 119,* 3940–3949.

Appendix. Table A-1. Pictures and distractors of Children's Multimodal Picture–Word Test (Jerger et al., 2009).

Phonology items					
Pictured objects			**Distractors**		
Bees	Deer	Beach	Peacock	Demon	Teacher
Bus	Duck	Beast	Peach	Detective	Tee-shirt
Pizza	Teeth	Buckle	Potato	Dumbo	Tomato
Pumpkin	Tongue	Butterfly	Puddle	Dumptruck	Tugboat
			Eagle	Onion	

Semantic filler items				
Pictured objects		**Distractors**		
Boot	Dog	Bear	Bed	Cat
Doll	Pants	Cheese	Dress	Flag
Pickle	Pizza	Glove	Horse	Hot dog
Tiger		Lemon	Puppet	Shirt
		Slipper	Worm	

Note. Phonology pictures were administered in the presence of three types of phonologically related distractors (congruent, one feature conflicting in place, and one feature conflicting in voicing onsets; e.g., the picture "Bus" coupled with "Buckle," "Dumptruck," and "Potato," respectively) plus the baseline distractor "Onion" for /ʌ/ vowel-nucleus pictures and "Eagle" for /i/ vowel-nucleus pictures. Filler-item pictures were presented in the presence of two types of distractors (semantically related and semantically unrelated; e.g., the picture "Boot" coupled with "Slipper" and "Flag," respectively).