Jerger, S., Damian, M. F., McAlpine, R. P., & Abdi, H. (2017). Visual speech alters the discrimination and identification of non-intact auditory speech in children with hearing loss. International Journal of Pediatric Otorhinolaryngology, 94, 127-137. DOI: 10.1016/j.ijporl.2017.01.009

International Journal of Pediatric Otorhinolaryngology 94 (2017) 127–137

Contents lists available at ScienceDirect

International Journal of Pediatric Otorhinolaryngology

journal homepage: http://www.ijporlonline.com/

Visual speech alters the discrimination and identification of non-intact auditory speech in children with hearing loss☆

Susan Jerger [a, b, *], Markus F. Damian [c], Rachel P. McAlpine [a, b], Hervé Abdi [a]

[a] School of Behavioral and Brain Sciences, GR4.1, University of Texas at Dallas, 800 W. Campbell Rd, Richardson, TX, 75080, USA
[b] Callier Center for Communication Disorders, 811 Synergy Park Blvd., Richardson, TX, 75080, USA
[c] University of Bristol, School of Experimental Psychology, 12a Priory Road, Room 1D20, Bristol, BS8 1TU, United Kingdom

ARTICLE INFO

Article history:
Received 19 July 2016
Received in revised form
5 January 2017
Accepted 6 January 2017
Available online 9 January 2017

Keywords:
Visual speech
Audiovisual speech
Phoneme discrimination
Childhood hearing loss
Lipreading

ABSTRACT

Objectives: Understanding spoken language is an audiovisual event that depends critically on the ability to discriminate and identify phonemes yet we have little evidence about the role of early auditory experience and visual speech on the development of these fundamental perceptual skills. Objectives of this research were to determine 1) how visual speech influences phoneme discrimination and identification; 2) whether visual speech influences these two processes in a like manner, such that discrimination predicts identification; and 3) how the degree of hearing loss affects this relationship. Such evidence is crucial for developing effective intervention strategies to mitigate the effects of hearing loss on language development.
Methods: Participants were 58 children with early-onset sensorineural hearing loss (CHL, 53% girls, *M* = 9;4 yrs) and 58 children with normal hearing (CNH, 53% girls, *M* = 9;4 yrs). Test items were consonant-vowel (CV) syllables and nonwords with intact visual speech coupled to non-intact auditory speech (excised onsets) as, for example, an intact consonant/rhyme in the visual track (*Baa or Baz*) coupled to non-intact onset/rhyme in the auditory track (*/–B/aa or/–B/az*). The items started with an easy-to-speechread/B/or difficult-to-speechread/G/onset and were presented in the auditory (static face) vs. audiovisual (dynamic face) modes. We assessed discrimination for intact vs. non-intact different pairs (e.g., Baa:/–B/aa). We predicted that visual speech would cause the non-intact onset to be perceived as intact and would therefore generate more same—as opposed to different—responses in the audiovisual than auditory mode. We assessed identification by repetition of nonwords with non-intact onsets (e.g.,/–B/az). We predicted that visual speech would cause the non-intact onset to be perceived as intact and would therefore generate more Baz—as opposed to az— responses in the audiovisual than auditory mode.
Results: Performance in the audiovisual mode showed more same responses for the intact vs. non-intact different pairs (e.g., Baa:/–B/aa) and more intact onset responses for nonword repetition (Baz for/–B/az). Thus visual speech altered both discrimination and identification in the CHL—to a large extent for the/B/ onsets but only minimally for the/G/onsets. The CHL identified the stimuli similarly to the CNH but did not discriminate the stimuli similarly. A bias-free measure of the children's discrimination skills (i.e., d' analysis) revealed that the CHL had greater difficulty discriminating intact from non-intact speech in both modes. As the degree of HL worsened, the ability to discriminate the intact vs. non-intact onsets in the auditory mode worsened. Discrimination ability in CHL significantly predicted their identification of the onsets—even after variation due to the other variables was controlled.
Conclusions: These results clearly established that visual speech can fill in non-intact auditory speech, and this effect, in turn, made the non-intact onsets more difficult to discriminate from intact speech and more likely to be perceived as intact. Such results 1) demonstrate the value of visual speech at multiple

☆ This work was supported by the National Institute on Deafness and Other Communication Disorders, grant DC-00421 to the University of Texas at Dallas.

* Corresponding author. School of Behavioral and Brain Sciences, GR4.1, University of Texas at Dallas, 800 W. Campbell Rd, Richardson, TX, 75080, USA.
E-mail addresses: sjerger@utdallas.edu (S. Jerger), m.damian@bristol.ac.uk (M.F. Damian), rachelmparra@gmail.com (R.P. McAlpine), herve@utdallas.edu (H. Abdi).

http://dx.doi.org/10.1016/j.ijporl.2017.01.009

levels of linguistic processing and 2) support intervention programs that view visual speech as a powerful asset for developing spoken language in CHL.

1. Introduction

Spoken conversations—a daily challenge for children with prelingual sensorineural hearing loss (CHL)—are audiovisual events that critically depend on the ability to discriminate and identify phonemes yet we have little evidence about the role of early auditory experience and visual speech on the development of these two fundamental skills underlying accurate speech perception. To address this gap in the literature, the objectives of this research were to determine *1*) how visual speech influences phoneme discrimination and identification, *2*) whether visual speech influences these two processes in a like manner, such that discrimination predicts identification, and *3*) how the degree of hearing loss affects this relationship. A proposed link between audiovisual speech and the development of phoneme discrimination and identification is widely accepted, especially in CHL [1], but it is patently understudied in CHL with mild to severe losses who are hearing aid users with access to many of the cues of speech. Nonetheless, the existence of such a link is supported by the studies below. In these studies, the tasks typically required CHL and children with normal hearing (CNH) to: *1*) detect the difference, if any, between the speech sounds of successive utterances and respond same or different (discrimination) or *2*) abstract/identify the speech sounds of an utterance and respond by repeating the utterance or pointing to a picture (identification). These two paradigms tap different levels of linguistic processing [see [2]].

1.1. Effect of visual speech on phoneme discrimination

1.1.1. CNH

Visual speech improves phoneme discrimination by infants with NH [3]. CNH also perform better for audiovisual than auditory input for *1*) feature contrast discrimination [4], *2*) vowel phoneme monitoring [5], and *3*) phoneme discrimination for visually distinct ('ba' vs. 'ga') contrasts [[6]; but see [7, 8] for exceptions]. Some data also suggest that CNH who show greater sensitivity to visual speech show better perceptual tuning to the phoneme contrasts of the native language relative to non-native contrasts [9,10].

1.1.2. CHL

With regard to CHL with mild to severe losses who are hearing aid users, the research on phoneme discrimination is meager, and the results are mixed. One study [11]—on a task assessing repetition of phonetic feature contrasts (daa vs. baa)—reported that performance in CHL did not differ in the audiovisual vs. auditory modes so CHL showed no benefit of visual speech. Another study [12], however, revealed the importance of visual speech for phonological knowledge. This research initially assessed phoneme discrimination in CHL (as supplementary data for a picture-word study) in order to divide the children into two groups depending on whether they Could or Could-Not discriminate the phonemes above chance. The groups were tested on phonological tasks in the audiovisual vs. auditory modes. The Could-Not-Discriminate CHL performed significantly better in the audiovisual than auditory mode (onset: 97% vs. 80%; rhyme: 93% vs. 77%). The Could-Discriminate CHL performed near ceiling for both modes, and so the effect of visual speech in these children could not be evaluated.

1.2. Effect of visual speech on phoneme or nonword identification

1.2.1. CNH

In general, the influence of visual speech on identification seems to increase with age. Identification by children from roughly 5–11-yrs-old does not show adult-like benefit from visual speech whereas identification by preteens–teenagers does [9,10,13–17]. A notable complication to this story is created, however, by other studies reporting that visual speech significantly benefits identification—for at least some conditions—in 3–5-yr-olds [6], 8-yr-olds [18,19], and all ages between 4 and 14-yr-olds [20].

1.2.2. CHL

Due to the limited number of studies of phoneme/nonword identification in CHL with mild to severe losses who are hearing aid users, this section also includes studies of word identification. It should be acknowledged, however, that identifying, discriminating, and detecting phonemes in words can reflect lexical-semantic influences [5,21,22]. In general, audiovisual speech perception in CHL exceeds perception for either modality alone—although large individual variability characterizes these results (e.g., [23–26]. In our series of studies—although many CHL performed at ceiling because their losses were mild to moderate—identification of speech stimuli was always numerically better in the audiovisual than in the auditory mode (i.e., about 95% vs. 85%) [see, e.g., [27, 28]]. Finally, in early development, experience with auditory speech (i.e., degree of HL) also appears to alter recognition of the correspondences between auditory and visual speech. For example, infants/toddlers with NH or mild-moderate HL—when hearing a word while watching side-by-side images of two talkers, one mouthing the heard word and one mouthing a different word—looked longer at the matching visual speech input whereas infants/toddlers with severe-to-profound HL did not [29].

1.3. Effect of visual speech on relation between discrimination and identification

1.3.1. CNH

Current, but scant, evidence indicates that: *1*) Visual speech (salient contrasts) improved both phoneme discrimination and identification [6], and 2) Visual speech *2a*) produced adult-like benefit at an earlier age for word discrimination than word recognition as measured in an AX paradigm, and *2b*) benefited word recognition as measured in an AX paradigm more than word discrimination in adults but not in young children [30]. If we supplement this limited information with the auditory only evidence, the findings are mixed [31–35]. Finally, research with infants with NH indicates that auditory phoneme discrimination at 6-mos predicted word identification at later ages [36].

1.3.2. CHL

To the best of our knowledge, previous research has not addressed the relationship between discrimination and identification—and the influence of visual speech on such—in CHL with mild to severe losses who are hearing aid users. This is an important gap because previous studies mostly studied CHL with severe to profound hearing loss and/or focused on either discrimination or identification. An example of a particularly relevant auditory only study—that assessed the discrimination of phonetic feature contrasts for pairs of stimuli (the French words "bouche–mouche") and the identification of each of these same stimuli presented individually ("bouche" or "mouche")—reports that CHL who use cochlear implants had poorer discrimination and identification than CNH, and that both groups showed slightly better discrimination than identification [37].

Theoretical support for the importance of studying the relationship between discrimination and identification is provided by our early speech perception/production models, which posit that discrimination (perceptual representational level) develops before identification (cognitive/linguistic level) [38,39]—a proposal consistent with the idea that discrimination and identification reflect different levels of speech representation [2]. Support for the importance of visual speech to this relationship is the compelling evidence that visual speech plays a major role in helping children learn to discriminate and identify phonemes [e.g., [40–43]]. Further cogent evidence in this regard is the finding that—in individuals with early-onset blindness—phonology and early expressive language skills are delayed and/or different (e.g., [44–46]).

1.4. Predictions and overview

1.4.1. CNH and CHL

Overall, the above results in the literature generally predict: *1*) that CNH will discriminate phonemes better audiovisually than auditory only—whereas mixed evidence indicates that CHL may or may not, and *2*) that CHL will identify words better audiovisually than auditory only—whereas mixed evidence indicates that CNH may or may not. With regard to our new stimuli with non-intact auditory onsets—a primary focus of this research—these experimental stimuli are detailed in the "Methods" section along with predicted findings. As will be seen, we predict that visual speech will restore or fill in the non-intact auditory onsets and that this effect, in turn, will make the non-intact onsets more difficult to discriminate from intact speech and more likely to be perceived as intact in the audiovisual mode (see 20). With regard to the relation between discrimination and identification, limited results in the literature—to the extent that word discrimination and recognition in an AX paradigm transfer to our research—suggest that CNH will discriminate and identify phonemes better audiovisually than auditory only. Results in the auditory study with CHL who used cochlear implants and CNH suggest that the current participants will discriminate speech better than they identify it [37]. Finally, this research will, for the first time, indicate if children's audiovisual phoneme discrimination predicts their audiovisual nonword identification. The aforementioned models of the development of speech perception/production—which propose that discrimination develops before identification—suggest that children's proficiency in discriminating phonemes may be related to their capability in identifying phonemes [38,39]. And, again, we are studying nonwords in order to assess both discrimination (Baa-Baa) and identification (Baz) at the phonological level of linguistic representation, thus eliminating or at least minimizing lexical-semantic influences on performance [5,21,22].

2. Method

2.1. Participants

Participants were 58 CHL with early-onset sensorineural loss (53% girls) and 58 CNH (53% girls). The CNH group—with a mean and distribution of ages akin to that in the CHL group—was formed from a pool of 132 typically developing children from associated projects [see 20, 47]. Sixty-four percent of the current CNH participated in our previous study assessing identification results for words vs. nonwords (20). Ages (yr;mo) ranged from 4;3 to 14;9 ($M = 9;4$, $SD = 2;11$) in the CHL and 4;3 to 14;5 ($M = 9;4$, $SD = 2;10$) in the CNH. The racial distributions were 73% Whites, 22% Blacks, and 5% Asian in CHL and 82% Whites, 5% Blacks, 10% Asian, and 3% Multiracial in CNH, with 12% of CNH reporting Hispanic ethnicity. All participants met the following criteria: 1) English as a native language, 2) ability to communicate successfully aurally/orally, and 3) no diagnosed or suspected disabilities other than HI and its accompanying speech and language problems.

2.1.1. Audiological characteristics

Hearing sensitivity in the CNH at hearing levels (HLs) of 500, 1000, and 2000 Hz (pure-tone average, PTA) [48] averaged 2.50 dB HL ($SD = 4.61$, right ear) and 3.39 dB HL ($SD = 5.34$, left ear). The PTAs in the CHL averaged 45.20 dB HL (better ear) and 55.95 dB HL (poorer ear). The PTAs on the better/poorer ears respectively were distributed as follows: ≤ 20 dB (10%/3%), 21 to 40 dB (29%/24%), 41 to 60 dB (37%/36%), 61 to 80 dB (21%/22%), 81 to 100 dB (3%/10%), and greater than 100 dB (0%/5%). The children with PTAs of ≤20 dB had losses in restricted frequency regions (e.g., falling audiometric contours). Hearing aids were used by 90% of these children. Participants who wore amplification were tested while wearing their devices. The estimated age at which the children who wore amplification received their first aid averaged 2.55 yrs ($SD = 1.84$); the estimated duration of device use averaged 8.84 yrs ($SD = 2.77$). Forty-four children were mainstreamed in a public school setting and 14 children were enrolled in an aural/oral school for CHL.

2.2. Comparison of groups

2.2.1. Materials

Vocabulary skills were estimated with the Peabody Picture Vocabulary Test-III [49] and the Expressive One-Word Picture Vocabulary Test [50]. Phonological awareness was estimated with subtests of the Pre-Reading Inventory of Phonological Awareness [51] measuring children's ability to isolate onset phonemes, recognize alliterative onset phonemes, and segment the phonemes within a word. Articulation proficiency was estimated with the Goldman Fristoe Test of Articulation [52]. Visual perception was estimated by the Beery-Buktenica Developmental Test of Visual Perception [53]. Spoken word recognition at 70 dB SPL was estimated with the Word Intelligibility by Picture Identification (WIPI) Test (auditory mode) [54] and the Children's Audiovisual Enhancement Test (CAVET; auditory, audiovisual, and visual only (lipreading) modes) [55].

2.2.2. Findings

Table 1 compares performance on the set of verbal and nonverbal measures in the groups. We carried out multiple *t*-tests on a subset of these measures (vocabulary, phonological awareness, visual perception, and lipreading) with the problem of multiple comparisons controlled with the False Discovery Rate (FDR) procedure [56]. We did not include articulatory proficiency and auditory word recognition because more than half of the children in each group made few errors: respectively ≤1 error in 53% (CHL) and 90% (CNH) and >90% correct in 65% (CHL) and 100% (CNH). Not surprisingly, average results for articulatory proficiency and auditory word recognition were numerically poorer in CHL than CNH—a result consistent with long term previous findings [e.g., 12]. Results of the *t*-tests indicated that the CNH had significantly better vocabulary skills, phonological awareness, and visual perception. The difference between groups in verbal skills is expected but reasons for the difference in visual perception are

Table 1
Average (standard deviation in parentheses) performance on a set of verbal and nonverbal measures in the CHL vs CNH.

Groups		
Measures	CHL N = 58	CNH N = 58
Verbal Skills		
Vocabulary (standard score)		
Receptive*	94.85 (16.65)	121.22 (11.47)
Expressive*	87.02 (13.78)	120.40 (11.43)
Phonological Awareness (%)*	69.38 (23.05)	77.73 (10.43)
Articulation Proficiency (# errors)	4.05 (6.79)	0.62 (1.58)
Nonverbal Skills		
Visual Perception (standard score)*	100.57 (16.58)	113.71 (14.07)
Word Recognition (%)		
Auditory	88.03 (10.82)	99.59 (1.53)
Audiovisual	94.91 (10.78)	–
Lipreading Onsets#	69.22 (21.82)	66.02 (17.24)

Note: * Indicates performance in CNH vs CHL differed significantly (adjusted $p < 0.05$).
Tests included in the statistical analyses were vocabulary, phonological awareness, visual perception, and lipreading (see text).
—Audiovisual mode for word recognition was not administered in CNH due to ceiling performance in auditory mode.
Lipreading was quantified by the percentage of word-onsets repeated correctly in visual only mode, with visemes (e.g., pat for bat) scored as correct. We focused on lipreading onsets because we are assessing the discrimination of onsets. The percentage of words repeated correctly in the visual only mode (lipread) averaged 15.71% in CNH and 25.60% in CHL.

unclear—although poorer visual perception in CHL has been observed previously [57]. Please note, however, that visual perception in both of the current groups (CHL and CNH) was within the average normal range, and lipreading the onsets did not differ between groups.

2.3. Materials and instrumentation: stimuli

2.3.1. Recording

The stimuli were recorded (Quicktime movie files) by an 11-yr-old boy actor with clearly intelligible speech without pubertal characteristics. His full facial image and upper chest were recorded. The color video signal was digitized at 30 frames/s with 24-bit resolution at a 720 × 480 pixel size. The auditory signal was digitized at a 48 kHz sampling rate with 16-bit amplitude resolution. The utterances were adjusted to equivalent A-weighted root mean square sound levels (see 20, 47 for details). The set of items consisted of experimental items—4 vowels (/i/,/æ/,/ʌ/,/o/or/ee/,/aa/,/uh/,/oh/); 8 consonant-vowel (CV) syllables (the onsets/B/or/G/coupled with each vowel, e.g., Baa, Gaa); and 8 nonwords (the onsets/B/or/G/coupled with each vowel and a final consonant, e.g., Baz, Gak)—and 14 filler items (vowel-onsets or not/B/or/G/onsets coupled with varying offsets, e.g., Cheeg, Doss, Tavel, Eeble, Osh-uck). Note that the experimental items reflected an easy-to-speechread onset/B/and a difficult-to-speechread onset/G/, a design that allowed us to assess the effects of speechreadability [58].

2.3.2. Low fidelity (non-intact) auditory onsets

We edited the auditory track of the/B/and/G/experimental items by locating the onsets visually and auditorily with Adobe Premiere Pro and Soundbooth (Adobe Systems Inc., San Jose, CA) and loudspeakers. We excised the waveforms in 1 ms steps from the identified auditory onsets to the point in the later waveforms for which at least 4 of 5 trained adult listeners heard the vowel—not the consonant—as the onset in the auditory mode. Splice points were always at zero axis crossings. Using this perceptual criterion, we excised on average from the/B/and/G/onsets respectively 51 ms and 63 ms for the CV syllables and 63 ms and 72 ms for the nonwords. The visual track of the utterances was also edited to form audiovisual (dynamic face) vs auditory (static face) modes of presentation (see also 20, 47). The video track was routed to a high-resolution computer monitor and the auditory track was routed through a speech audiometer to a loudspeaker.

2.3.3. Audiovisual vs auditory modes

The audiovisual stimuli consisted of a brief period of the talker's still neutral face and upper chest followed by an audiovisual presentation of either the to-be-discriminated pair of CV syllables or the to-be-repeated nonword followed by the talker's still neutral face and upper chest. The auditory mode consisted of exactly the same auditory track but the visual track contained the talker's still neutral face and upper chest for the entire trial. As an example of a nonword stimulus, the audiovisual mode consisted of 1) an intact consonant/rhyme (e.g., *Baz*) in the visual track coupled to 2) a non-intact onset/rhyme (e.g.,*/–B/az*) in the auditory track. The auditory mode consisted of 1) a static face in the visual track coupled to 2) the same non-intact auditory track (*/–B/az).*

Our question was whether the intact visual speech would restore or fill in the non-intact auditory onset. If yes, then performance for the *same* auditory stimulus would differ depending upon the mode (for example, perceiving/*Baz*/in the audiovisual mode but/*az*/in the auditory mode). We should also note that the auditory mode controls for any influence on performance due to any remaining coarticulatory cues in the stimulus and for any strategic effects that might be characterizing the children's performance. With these influences on performance controlled, we can identify whether the addition of visual speech affected performance on our tasks.

To preview the experimental paradigms presented below, for the discrimination task, we asked children to judge whether two CV syllables were the same (e.g., Baa:Baa) or different (e.g., Baa:Gaa). The items of particular interest, however, were different pairs that consisted of one intact vs. one non-intact onset (e.g., Baa:/–B/aa). For the identification task, we asked children to repeat what they perceived for nonwords with intact (Baz) or non-intact (–B/az) onsets. We predicted that visual speech would restore or fill in the non-intact auditory onsets, and thus the non-intact onsets would be perceived as intact in the presence of visual speech. This would generate more same—as opposed to different—responses in the audiovisual than auditory mode for the discrimination task and more Baz—as opposed to az—responses in the audiovisual than auditory mode for the identification task.

2.3.4. Set of items: discrimination

The pairs of items—presented in the audiovisual and auditory modes—consisted of CV syllables with intact/B/and/G/onsets (e.g., Boh), CV syllables with non-intact/B/and/G/onsets (e.g.,/–B/oh), and intact vowel syllables (e.g., oh). Each trial presented two CV or two vowel syllables, which were the same (e.g., aa:aa, Baa:Baa, or/–B/aa:/–B/aa) or different (e.g., aa:ee, Baa:Gaa, or Baa:/–B/aa). The two syllables were separated by a silent interval of 1400 ms. Pilot studies indicated that the administration of all possible pairs from this set of items was ill-advised because the children expressed an unshakable dislike of this task, an effect consistent with previous experiences, see [e.g., [59]]. Thus we shortened the task and administered only a subset of items to each child.

We formed four lists representing four subsets of the items (the lists were presented forwards and backwards to yield eight variations) [see [60] for our principled approach for randomly omitting items and one illustrative subset]. Each list ($N = 70$, 35 items in each mode) contained the following items in each mode:

4 INTACT-SAME CV pairs (Boh:Boh, Gee:Gee),
4 INTACT-DIFFERENT CV pairs (Baa:Gaa, Guh:Buh),
3 INTACT-SAME VOWEL pairs (oh:oh, uh:uh),
6 INTACT-DIFFERENT VOWEL pairs (aa:ee, ee:oh),
6 NON-INTACT-SAME CV pairs (/-B/oh:/-B/oh,/-G/uh:/-G/uh), and —the pairs of particular interest—
12 CV pairs DIFFERING IN INTACTNESS (/-B/oh:Boh, Guh:/-G/uh).

For the last condition, some CV pairs, selected randomly, were presented twice. The *a priori* probabilities of a list were 41% same − 59% different for the intact items and 33% same − 67% different for the non-intact items. Note, however, that the probabilities for the non-intact items (based on physical characteristics) are not precise because the perceptions of listeners vary—but, nevertheless, with a general tendency to perceive: 1) a vowel-onset in the auditory mode and a consonant onset in the audiovisual mode for the/B/onsets and 2) a vowel-onset in both modes for the/G/onsets [20].

The items of a list were randomly intermixed under the constraints that 1) no item could repeat; 2) the intact and non-intact versions of the same item must be separated by at least two intervening items; 3) the mode must alternate after three repetitions; and 4) the modes (audiovisual, auditory), the judgments (same, different), the types of pairs (intact, non-intact, intact:non-intact), and the types of items (intact vowel, intact/B/and/G/, non-intact/B/and/G/) must be dispersed uniformly throughout the lists. The presentation of individual items was counterbalanced such that 50% of the items occurred first in the auditory mode and 50% occurred first in the audiovisual mode. The response board contained two keys designated same/alike and different/not alike by 1) two copies of the same shape in the same color and 2) two different shapes in different colors. The side corresponding to each response was counterbalanced across participants.

2.3.5. Set of items: identification

The items consisted of experimental items—8 intact and 8 non-intact nonwords with/B/and/G/onsets (e.g., Beece or/−B/eece; Geen or/−G/een)—and 14 filler items (nonwords with vowel onsets, e.g., Apper, Onyit, or not/B/or/G/consonant onsets, e.g., Hork, Tyfer). All of these items were presented in the audiovisual and auditory modes with each experimental item presented twice in each mode. Thus, listeners heard trials randomly alternating between intact and non-intact onsets, audiovisual and auditory modes, and test and filler items. The entire set of experimental and filler items were randomly intermixed and formed into four lists (presented forwards and backwards for eight variations). Each list consisted of 64 experimental items and 48 filler items, yielding 57% test trials. Each list varied randomly with the constraints that (a) no onset could repeat, (b) the intact and non-intact versions of the same item could not occur without at least two intervening items, (c) a non-intact onset must be followed by an intact onset, (d) the mode must alternate after three repetitions, and (e) all types of onsets (intact/B/and/G/, non-intact/B/and/G/, vowels, and not/B/or/G/) needed to be dispersed uniformly throughout the lists.

2.4. Procedure

2.4.1. General

Testing was carried out within a sound-treated booth. The tester sat at a computer workstation, and the children, with a co-tester alongside, sat at a distance of 71 cm directly in front of an adjustable height table containing the computer monitor and loudspeaker. Their view of the talker's face subtended a visual angle of 7.17° vertically (eyebrow to chin) and 10.71° horizontally (eye level). The stimuli were presented at an intensity level of approximately 70 dB SPL. These data were gathered as part of a larger experimental protocol administered over three testing sessions of about 1 h each [20,28,47]. The interval between sessions averaged 12 days in each group.

2.4.2. Discrimination

The children were instructed as follows:

A boy is going to say two sounds and sometimes they will be the same/alike (demonstrate: aa-aa or Bee-Bee) and sometimes they will be different/not alike (demonstrate: aa:ee or Guh:Buh). And sometimes the boy's mouth will move and sometimes it will not move. Your job is to push this button if the sounds are the same/alike (demonstrate: oh-oh; Baa-Baa) and push this button if the sounds are different/not alike (ee:aa; Buh:Guh).

Each child completed one list, with one-half of items presented in the second testing session and one-half in the third session. Each half list began with practice items. The children voted same/alike or different/not alike by pushing the correct button in a two-alternative forced-choice paradigm.

2.4.3. Identification

For the identification task, we asked children to *repeat exactly* what the talker said. The children's utterances were transcribed independently by the tester and co-tester and also digitally recorded. For the utterances with non-intact onsets, the transcribers disagreed on 2.28% of responses. For these responses, another trained listener independently transcribed the recorded utterances. Her transcription, which always agreed with one of the other transcribers, was recorded as the response. The criteria for scoring responses to the non-intact onsets (illustrated for/−B/az) were as follows:

1. Correct vowel onsets (“az”) scored as an auditory-based response for both modes.
2. Correct consonant onsets (“Baz”) scored as a visual-based response for the audiovisual mode and as a coarticulatory response for the auditory mode.
3. Incorrect vowel or consonant onsets (“Faz”) scored as errors.

Each child completed one list, subdivided into 4 sections. For one-half of children, 3 sections of a list were completed (in separated listening conditions) in the first testing session with the 4th section completed in the second testing session. For the other one-half of children, 1 section of a list was completed in the second testing session with the remaining 3 sections completed (in separated listening conditions) in the third testing session [see 20 for details). A variable number of practice items preceded each section.

Finally, with regard to comparing discrimination of the CV syllable onsets vs. identification of the nonword onsets, we should note that a pilot study with young adults indicated that identification of the onsets of CV syllables vs. nonwords did not differ. This research protocol was approved by the Institutional Review Boards of University of Texas at Dallas and Washington University St. Louis.

3. Results

3.1. Discrimination and identification of the intact onsets

3.1.1. Discrimination

The CNH discriminated the intact different pairs (Baa:Gaa, ee:aa) at 100% accuracy for both modes. The CHL discriminated the intact different pairs in the auditory and audiovisual modes respectively, on average, at 91%–92% accuracy for the vowels and 84%–88% accuracy for the consonants.

3.1.2. Identification

The children repeated the intact nonword onsets (Baz, Gak) in the auditory and audiovisual modes, on average, at 100%–100% (CNH) and 85%–92% (CHL). All of the CHL could articulate the/B/ and/G/onsets correctly, and errors were primarily due to lexicalization of the nonwords, e.g., "back" for Gak. The children repeated the intact offsets (i.e., the remainder of the utterance) in the auditory and audiovisual modes respectively, on average, at 100%–100% (CNH) and 68%–71% (CHL). Below we analyze performance for the non-intact onsets.

3.2. Discrimination vs. identification of the non-intact onsets

In the discrimination task, we wished to determine whether visual speech made it harder to discriminate non-intact from intact speech (e.g., Baa:/–B/aa perceived as Baa:Baa). Thus we focused on the intact vs non-intact different pairs (e.g., Baa:/–B/aa) and determined the percentage of same responses to the pairs differing in intactness. In the identification task, we wished to determine whether visual speech made it more likely to perceive the non-intact onsets as intact (e.g.,/–B/az perceived as Baz). Thus we focused on the nonwords with non-intact/B/and/G/onsets and determined the percentage of correct onset responses for the non-intact nonwords. The goals of the data analyses were to determine 1) how hearing loss affects children's discrimination and identification, 2) how visual speech may offset any such effects, and 3) whether discrimination predicts identification. The Bonferroni correction controlled the familywise alpha in all analyses [61].

3.2.1. Results in total group

Table 2 compares the groups'—CHL (2A) and CNH (2B)—average ability to discriminate vs. identify the stimuli for the/B/and/G/onsets in the auditory and audiovisual modes. The arrows and boxes will be explained momentarily. The data were analyzed with a repeated-measures mixed design analysis of variance (ANOVA) with one between-participant factor (Group: CHL vs. CNH) and three within-participant factors (Task: discrimination vs. identification; Mode: auditory vs. audiovisual; Onset:/B/vs/G/). Results (see Table 3A) revealed a significant difference between *1*)

Table 3
Results of repeated-measures mixed design analysis of variance (ANOVA) for data in the two groups (Table 2). The between-participant factor was Group (CHL vs. CNH) and the within-participant factors were Task: identification vs. discrimination, Mode: auditory vs. audiovisual, and Onset: /B/vs/G/). The dependent variable was the discrimination and identification responses (% same or % correct).

A. Significant Statistical Outcomes (ANOVA)

Factors	Mean Square Error	*F* value	*p* value	Partial η^2
Mode	483.18	296.86	<0.0001	0.723
Onset	515.89	198.82	<0.0001	0.636
Task	804.96	45.33	<0.0001	0.285
Task × Group	804.96	25.80	<0.0001	0.185
Onset × Mode	404.71	156.55	<0.0001	0.579
Onset × Mode × Task	232.90	20.77	<0.0001	0.154

Note: df's = 1, 114 for all effects.

B. F contrast analysis

	/B/onset				/G/onset			
	Δ	*F*	*p*	partial η2	Δ	*F*	*p*	partial η2
Task: Discrimination – Identification Contrast								
CHL								
Auditory	**26.6**	**88.0**	**<0.001**	**0.44**	**26.6**	**87.9**	**<0.001**	**0.44**
Audiovisual	7.6	7.2	ns	0.06	**27.3**	**92.7**	**<0.001**	**0.45**
CNH								
Auditory	5.0	3.2	ns	0.03	3.1	2.4	ns	0.02
Audiovisual	−5.3	3.5	ns	0.03	**9.5**	**11.3**	**0.001**	**0.09**
Mode: Audiovisual – Auditory Contrast								
CHL								
Discrimination	**27.6**	**94.8**	**<0.001**	**0.45**	**9.9**	**12.2**	**<0.001**	**0.10**
Identification	**46.6**	**270.4**	**<0.001**	**0.70**	**9.2**	**10.5**	**0.002**	**0.09**
CNH								
Discrimination	**40.5**	**204.3**	**<0.001**	**0.64**	**10.4**	**13.3**	**<0.001**	**0.11**
Identification	**50.9**	**322.1**	**<0.001**	**0.74**	3.9	1.9	ns	0.02

Note: ns = not significant (i.e., adjusted p value > 0.05). Significant results are bolded.
The mean square error for the contrasts was 232.90.

Table 2
Mean results (standard errors in parentheses) for the discrimination vs. identification of the stimuli in the auditory and audiovisual modes for the CHL (2A) and CNH (2B). Discrimination was quantified by percentage of same responses for the pairs differing in intactness (i.e., Baa:/–B/aa perceived as Baa:Baa). Identification was quantified by the percentage of correct onset responses for the non-intact onsets (i.e.,/–B/az perceived as Baz).

A. **CHL** (N = 58)

	Onset					
	/B/			/G/		
	Mode			Mode		
Task	Audiovisual		Auditory	Audiovisual		Auditory
Discrimination (% Same Responses)	77.16 (2.61)	>	49.57 (3.10)	48.06 (3.21)	>	38.15 (3.05)
			v	v		v
Identification (% Correct Onsets)	69.58 (2.80)	>	22.98 (3.16)	20.78 (2.90)	>	11.58 (2.75)

B. **CNH** (N = 58)

	Onset					
	/B/			/G/		
	Mode			Mode		
Task	Audiovisual		Auditory	Audiovisual		Auditory
Discrimination (% Same Responses)	66.67 (3.75)	>	26.16 (3.07)	37.93 (3.59)	>	27.58 (3.24)
				v		
Identification (% Correct Onsets)	71.98 (3.82)	>	21.12 (3.56)	28.41 (3.09)		24.51 (3.16)

Note: The boxes integrate these results with findings of the *F* contrast analyses. Please see text.

the auditory vs. audiovisual Modes, *2*) the easy-vs. hard-to-speechread Onsets, and *3*) the discrimination vs. identification Tasks. The finding of a significant difference between the Modes is particularly noteworthy because it indicates that visual speech significantly influenced both discrimination and identification in these children. A simple interpretation of these overall effects, however, is complicated by significant interactions as seen in Table 3A. To probe these interactions, we carried out an *F* contrast analysis for each onset in each group [62]. The contrasts analyzed the difference between the: *1*) Discrimination − Identification Tasks and *2*) Audiovisual − Auditory Modes. Table 3B presents the *F* contrast results, along with the absolute differences between the two means (shown in Table 2) that formed each contrast. The boxes and arrows in Table 2 graphically integrate these findings and the data. When the two means forming a Task or Mode Contrast did not differ, they are boxed together. When the two means forming a Contrast did differ, a greater than sign (>) indicates which mean was significantly better.

3.2.2. F contrast: discrimination − identification task

3.2.2.1. /B/onset. The *F* contrasts (***column effects***, Table 2) indicate that the Tasks differed only for the auditory mode in the CHL; otherwise discrimination vs. identification did not differ (boxed together). The significant *F* contrast occurred because the CHL had difficulty discriminating the intact vs. non-intact onsets in the auditory mode. Thus, these CHL voted same disproportionately on the Discrimination Task, and this behavior yielded a significant difference between Tasks.

3.2.2.2. /G/onset. The *F* contrasts indicate that the Tasks differed for both modes in the CHL and for the audiovisual mode in the CNH. The significant *F* contrast occurred because both the CHL and CNH had disproportionate difficulty discriminating the intact vs. non-intact onsets—this time in both modes for the CHL and in the audiovisual mode for the CNH. Results in the CNH indicate that visual speech—minimally, but significantly—affected discrimination more than identification for this hard-to-speechread onset.

3.2.3. F contrast: audiovisual − auditory mode

3.2.3.1. /B/onset. The *F* contrasts (***row effects***, Table 2) indicate that visual speech significantly altered both discrimination and identification in the CHL and CNH, by 28%–51% (Table 3B).

3.2.3.2. /G/onset. The *F* contrasts indicate that visual speech significantly but minimally (≤10%, Table 3B) affected both discrimination and identification in the CHL, but only discrimination in the CNH.

The above bundle of effects produced the interactions seen in Table 3A. One particular interaction worthy of further note is the Task × Group effect, which was the only significant difference between the Groups. We can examine this difference by imaging the results of Table 2 collapsed across the modes and the onsets. Such results reveal that whereas the Groups identified the stimuli similarly (CHL: 31%; CNH: 36%), they did not discriminate the stimuli similarly (CHL: 53%; CNH: 39%). Stated differently, the CHL performed differently on these two Tasks (Discrimination: 53%; Identification: 31%) whereas the CNH performed similarly on these two Tasks (Discrimination: 39%; Identification: 36%). Again, the CHL had difficulty discriminating the intact vs. non-intact onsets, and thus these children voted same disproportionately whereas they did not disproportionately identify the non-intact onsets as intact.

Below we address the question of whether the results in Table 2 reflect *1*) how well the children discriminated the pairs vs. *2*) a response bias because the children responded same most of the time to both the different (e.g., Baa:/−B/aa) and same (e.g.,/−B/aa:/−B/aa) pairs. Prior to proceeding, we should note that—when results were collapsed across all items—the children did not exhibit a response bias: They pushed the same and different buttons respectively 49% and 51% (CNH) and 55% and 45% (CHL) of the time. Due to the minimal effects of visual speech on the/G/onset, all further analyses were limited to the/B/onset.

3.3. D′ analysis: discrimination task

The *d*′ analysis determined the percentage of different responses a) to the different pairs (called hits) vs. b) to the same pairs (called false alarms) [63]. Table 4 details these percentages, along with the complementary percentages for the same responses (B onsets), in the CHL (4A) and the CNH (4B). We quantified the children's excellence in discrimination or sensitivity by the difference between the proportion of hits (H) and false alarms (FA)—transformed into *Z*-scores. These results showed two relevant findings. First—and critically important for our purposes—the presence of visual speech notably reduced the children's (both CHL and CNH) ability to discriminate the intact from non-intact speech. Second, for both modes, the CHL showed poorer discriminability than the CNH. This latter result confirms, with a bias-free measure, that the CHL had greater difficulty discriminating the intact from non-intact speech. As we turn to an examination of the effects of degree of HL on performance, the previous data predict that the discrimination task will be more sensitive to the effects of HL than the identification task.

3.4. Effects of degree of HL

To address whether the degree of HL—defined by the average hearing levels at 500, 1000, 2000, and 4000 Hz (PTA) on both ears—could account for the variability of performance in the auditory and audiovisual modes, we conducted a multiple regression analysis in the CHL with degree of HL as the predictor variable and discrimination/identification performance in the two modes as the criterion variables. Table 5 summarizes these findings. Results indicated that the degree of HL had a significant effect on overall performance (ALL variables considered simultaneously). However, the slope coefficients indicated that *only* the ability to discriminate the intact vs. non-intact onsets in the auditory mode worsened as the degree of HL worsened.

Table 4
The percentage of different and same responses to the different and same pairs of the Discrimination task (B onsets) formatted for *d* analysis: CHL (4A) and CNH (4B).

Targets	Mode			
	Response (%)		Response (%)	
	Audiovisual		Auditory	
	Different (%)	Same (%)	Different (%)	Same (%)
A. **CHL** ($N = 58$)				
Different (Baa:/−B/aa)	22.84 (H)	77.16	50.43 (H)	49.57
Same (/−B/aa:/−B/aa)	10.31 (FA)	89.69	9.90 (FA)	90.10
Difference	*12.53*		*40.53*	
d′	*0.51*		*1.09*	
B. **CNH** ($N = 58$)				
Different (Baa:/−B/aa)	33.33 (H)	66.67	73.84 (H)	26.16
Same (/−B/aa:/−B/aa)	0.00 (FA)	100.00	0.00 (FA)	100.00
Difference	33.33		73.84	
d′	*2.16*		*3.23*	

Note. (H) = Hits; FA = False Alarms. To address the value of 0 (zero) FA in the CNH, we applied the log-linear rule to all response percentages before conversion into z scores in both the CNH and CHL [68].

Table 5
Multiple regression analysis in CHL with degree of HL as the predictor variable and discrimination/identification performance in the two modes as the criterion variables. Results show multiple correlation coefficient and omnibus *F* for all variables considered simultaneously followed by the slope coefficients and the partial *F* statistics evaluating the variation in Discrimination/Identification in each mode uniquely accounted for by degree of HL (after removing the influence of the other variables).

Variables	/B/onset		
	Multiple *R*	Omnibus *F*	*p*
ALL	0.405	2.60	0.046
	Slope Coefficient	Partial *F*	*p*
Discrimination: Auditory	**0.330**	**8.39**	**0.005**
Audiovisual	0.041	0.07	0.790
Identification: Auditory	0.091	0.69	0.408
Audiovisual	-0.026	0.032	0.844

Note. Significant results are bolded.
MSE = 356.13, *df*'s = 4,53 for omnibus *F* and 1,53 for partial *F*.

To illustrate these findings, the CHL were separated into two groups on the basis of a median split. The PTA score averaged 35.45 dB HL (*SD* = 10.15; range: 5–49 dB HL) in the Better CHL Group and 68.30 dB HL (*SD* = 12.03; range: 51–98 dB HL) in the Poorer CHL Group. The average age (yr;mo) was 8;8 in the Better Group and 9;11 in the Poorer Group. Fig. 1 portrays average discrimination and identification in the Better vs. Poorer Groups in the audiovisual and auditory modes. The sticks alongside the bars detail average performance in the CNH (Table 2). Fig. 1 clearly illuminates the findings of the regression analysis, namely that the degree of HL only affected the discrimination task in the auditory mode.

In the section below, we queried whether the children's discrimination of the intact vs. non-intact onsets predicted their identification of the onsets. Supporting evidence in infants is that their ability to discriminate auditory stimuli predicts their later ability to identify words [36]. Clearly a simple developmental model of speech perception would suggest that children must detect and discriminate stimuli before they can identify and label each one individually [e.g., [64, 65]]. Again, due to the minimal effects of visual speech on the/G/onset, results below are limited to the/B/onset.

3.5. Does discrimination predict identification?

To evaluate whether the children's ability to discriminate speech predicts their ability to identify speech, we carried out regression analyses separately in the CHL and CNH groups. The predictor variable was each child's audiovisual discrimination; the criterion variables were audiovisual identification of the onsets and also auditory word identification as quantified by standardized tests (average performance on the Word Intelligibility by Picture Identification Test [54] and the Children's Audio-Visual Enhancement Test in the auditory mode [55]. We included standardized measures of auditory word identification due to the aforementioned findings in infants and also in an attempt to broaden the clinical implications of our findings. Degree of HL was included as a control variable. We included auditory word identification and degree of hearing sensitivity in the CNH to maintain a consistent regression model. However, there was no variability to analyze for these two measures in the CNH.

We computed the part (a.k.a. semi-partial) *r* and partial *F* statistics to determine whether discrimination predicted a significant proportion of the variation in identification (onsets or words) after the variability due to the degree of HL and the remaining variables was partialled out [66]. These statistics are presented in Table 6 along with the omnibus results for all variables considered together. The omnibus results (ALL) showed that discrimination ability significantly predicted overall performance in both the CHL and CNH, accounting for 26% to 43% of the variability. The part *r*'s indicated that discrimination skills significantly predicted both the CHL's and CNH's identification of the onsets when the variation due to the other variables was controlled (accounting for 21% to 41% of the variability). Finally, results suggested that the discrimination ability of CHL also predicted their word recognition performance (however, this association did not achieve significance, $p = 0.06$). Albeit non significant, this result seems worthy of mention because our statistical approach was stringent as it constrained prediction to only the variance that was uniquely shared between discrimination and word recognition.

4. Discussion

In adults, visual speech provides cues that can enhance multiple levels of linguistic processing such as discrimination and identification [67]. Yet there is scant evidence in CNH and CHL to clarify how visual speech affects these fundamental processes underlying accurate perception. The purpose of this research was to assess the relation between discrimination and identification in CNH vs. CHL and to determine whether visual speech makes it 1) harder to

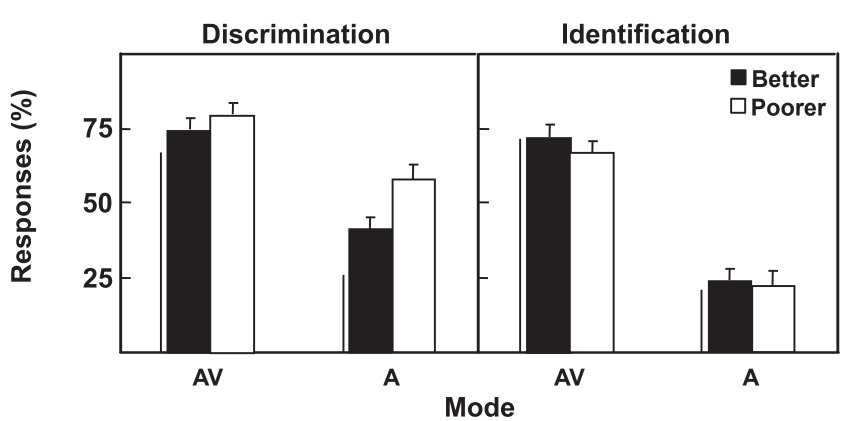

Fig. 1. Average discrimination (% same) and identification (% correct) in the Better vs. Poorer CHL Groups in the audiovisual and auditory modes. The PTA averaged 35.45 dB HL (*SD* = 10.15; range: 5–49 dB HL) in the Better CHL Group and 68.30 dB HL (*SD* = 12.03; range: 51–98 dB HL) in the Poorer CHL Group. The sticks alongside the bars detail average performance in the CNH (Table 2). If visual speech filled in the non-intact auditory onsets, this effect, in turn, should make the non-intact onsets more difficult to discriminate from intact speech and more likely to be identified as intact in the audiovisual mode.

Table 6
Summary of multiple regression results for the prediction of Identification (Audiovisual Onsets and Auditory Words) by Audiovisual Discrimination. Data were the stimuli with/B/onsets. Results show multiple correlation coefficient and omnibus F for all variables considered simultaneously followed by the slope coefficient, part r, and partial F statistics evaluating the variation in Identification uniquely accounted for by Audiovisual Discrimination (after removing the influence of the other variables).

A. CHL (with Degree of HL as a control variable)				
Variables		Multiple *R*	Omnibus *F*	*p*
ALL (Onset ID, Word ID, Degree of HL)		0.514	6.35	0.0009
	Slope Coefficient	Part *r*	Partial *F*	*p*
Onset ID: Audiovisual	**0.425**	**0.456**	**15.05**	**0.0003**
Word ID: Auditory	*0.487*	*0.224*	*3.59*	*0.063*
Note: $MSE = 305.52$.				
B. CNH (with Hearing Sensitivity as a control variable)				
Variables		Multiple *R*	Omnibus *F*	*p*
ALL (Onset ID, Word ID, Hearing PTA)		0.660	13.88	<0.0001
	Slope Coefficient	Part *r*	Partial *F*	*p*
Onset ID: Audiovisual	0.655	0.643	39.61	<0.0001
Word ID: Auditory	−0.617	0.032	0.10	0.751

Note: $MSE = 487.96$.
Grand *Note*. Significant results are bolded. *df*'s = 3,54 for omnibus *F* and 1,54 for partial *F*.

discriminate non-intact from intact speech (e.g., Baa:/−B/aa perceived as Baa:Baa), and 2) more likely to perceive non-intact onsets as intact (e.g.,/−G/ak perceived as Gak).

4.1. Did visual speech affect discrimination and identification?

A primary finding was that visual speech can fill in non-intact auditory speech. Visual speech consistently made it 1) harder to discriminate the non-intact from intact speech and 2) more likely to perceive the non-intact onsets as intact in both CHL and CNH for the easy to speechread onsets (/B/). These results agree with the finding that visual speech benefits both word discrimination and recognition (obtained in AX paradigm) in CNH [30]. For the hard-to-speechread onsets/G/, visual speech minimally (≤10%) affected both discrimination and identification in the CHL but only discrimination in the CNH. With regard to the relation between discrimination and identification, the tasks did not differ in either group for the easy to speechread onsets (/B/), with one exception: The CHL exhibited unusual difficulty discriminating the intact vs. non-intact onsets in the auditory mode and thus they voted same disproportionately whereas they did not identify the non-intact onsets as intact disproportionately. For the hard-to-speechread onsets/G/, both groups exhibited slightly (≤10%) greater difficulty discriminating the intact from non-intact onsets—a pattern of results that produced a significant difference between tasks in both modes for the CHL but only in the audiovisual mode for the CNH. This latter result indicated that visual speech affected discrimination more than identification in CNH for the hard-to-speechread onsets—although again this effect was noticeably slight.

In short, overall findings in the groups were more similar than different. The only significant difference between groups was that—when results were collapsed across modes and onsets—the groups identified the stimuli similarly (CHL: 31%, CNH: 36%) whereas they discriminated the stimuli differently (CHL: 53%, CNH: 39%). Again, the CHL had disproportionate difficulty discriminating the non-intact from intact onsets (CHL: voted same 53% of time, CNH: voted same 39% of time) whereas they did not disproportionately identify the non-intact onset as intact (CHL: 31%, CNH: 36%).

A bias-free measure of the children's discrimination (i.e., a *d'* analysis) indicated that visual speech impressively lessened both groups' ability to discriminate the intact vs. non-intact speech. Further the analysis confirmed that the CHL exhibited poorer discriminability of the intact vs. non-intact speech in both modes than the CNH. This latter finding agrees with results for the discrimination of auditory speech features in CNH vs. CHL who use cochlear implants [37]. Relative to the *d'* values (auditory mode) obtained by Bouton et al. [37], the *d'* values (auditory mode) in our CNH were reasonably similar (3.23 vs. 3.5 estimated from Bouton et al.), but the values in our CHL were lower than those of Bouton in cochlear implant users (1.09 vs. 2.0 estimated).

4.2. Did degree of HL affect discrimination and identification?

Results indicated that only the ability to discriminate the intact vs. non-intact onsets in the auditory mode worsened as the degree of HL worsened. We illustrated this finding by examining results in three groups: CNH, CHL with Better Sensitivity (35 dB HL), and CHL with Poorer Sensitivity (68 dB HL). These results clearly illustrated that the degree of HL affected only auditory discrimination. Also the effect of HL was systematic; when children discriminated the intact vs. non-intact onsets in the auditory mode, the CNH voted same less often than the CHL: Better group who voted same less often than the CHL: Poorer group (CNH ≠ Better ≠ Poorer). With regard to the effects of visual speech, discrimination in the audiovisual mode differed only between the CNH ≠ Poorer CHL. Thus, visual speech appeared to eliminate the difference between the CNH − Better CHL in the auditory mode, a pattern suggesting that visual speech minimized or eliminated the effects of mild HL on auditory discrimination.

4.3. Did discrimination predict identification?

Our specific question was whether a child's discrimination of the intact vs. non-intact onsets predicted her identification of the onsets in the audiovisual mode. Results indicated that discrimination ability did significantly predict—even after the variation due to the other variables was controlled— both the CHL's and CNH's identification of the onsets (accounting for 21% to 41% of the variability).

4.4. Conclusion

This research documented that visual speech significantly alters the discrimination and identification of non-intact auditory speech in children. Overall findings in the CHL and CNH were more similar than different. Visual speech filled in the non-intact auditory speech, and this effect, in turn, made the non-intact onsets 1) more difficult to discriminate from intact speech and 2) more likely to be identified as intact. Such results clearly demonstrate the value of visual speech cues at multiple levels of linguistic processing in children. Clinically these data support intervention programs that view visual speech as a powerful asset for developing spoken language in CHL. The incorporation of visual speech would focus attention: 1) on the talker's face, and this should promote sociolinguistic competencies, and 2) on the correspondences between the auditory and visual speech inputs, and this should promote both speech perception and production skills.

Acknowledgements

This research was supported by the National Institute on Deafness and Other Communication Disorders, grant DC-00421 to the University of Texas at Dallas. Dr. Abdi would like to acknowledge the support of an EURIAS fellowship at the Paris Institute for Advanced Studies (France), with the support of the European Union's 7th Framework Program for research, and from a funding from the French State managed by the "Agence Nationale de la Recherche (program: *Investissements d'avenir*, ANR-11-LABX-0027-01 Labex RFIEA+)." Sincere appreciation to 1) speech science colleagues for their guidance and advice to adopt a perceptual criterion for editing the non-intact stimuli. We thank Dr. Nancy Tye-Murray of Washington University School of Medicine for supervision of data collection in the CHL and for comments on a previous version of this paper. We thank the children and parents who participated and the research staff who assisted, namely Aisha Aguilera, Carissa Dees, Nina Dinh, Nadia Dunkerton, Alycia Elkins, Brittany Hernandez, Cassandra Karl, Demi Krieger, Michelle McNeal, Jeffrey Okonye, and Kimberly Periman of University of Texas at Dallas (data collection, analysis, presentation), and Derek Hammons and Scott Hawkins of The University of Texas at Dallas and Dr. Brent Spehar of Washington University School of Medicine (audiovisual recording and editing, computer programming).

References

[1] L. Woodhouse, L. Hickson, B. Dodd, Review of visual speech perception by hearing and hearing-impaired people: clinical implications, Int. J. Lang. Commun. Disord. 44 (2009) 253–270.
[2] C. Wood, Discriminability, response bias, and phoneme categories in discrimination of voice onset time, J. Acoust. Soc. Am. 60 (1976) 1381–1389.
[3] T. Teinonen, R. Aslin, P. Alku, G. Csibra, Visual speech contributes to phonetic learning in 6-month-old infants, Cogn 108 (2008) 850–855.
[4] T. Hnath-Chisolm, E. Laipply, A. Boothroyd, Age-related changes on a children's test of sensory-level speech perception capacity, J. Speech Lang. Hear Res. 41 (1998) 94–106.
[5] M. Fort, E. Spinelli, C. Savariaux, S. Kandel, Audiovisual vowel monitoring and the word superiority effect in children, Int. J. Behav. Dev. 36 (2012) 457–467.
[6] K. Lalonde, R. Holt, Preschoolers benefit from visually salient speech cues, J. Speech Lang. Hear Res. 58 (2015) 135–150.
[7] A. Boothroyd, L. Eisenberg, A. Martinez, An on-line imitative test of speech-pattern contrast perception (OlimSpac): developmental effects in normally hearing children, J. Speech Lang. Hear Res. 53 (2010) 531–542.
[8] B. Dodd, B. Mcintosh, D. Erdener, D. Burnham, Perception of the auditory-visual illusion in speech perception by children with phonological disorders, Clin. Linguist. Phon. 22 (2008) 69–82.
[9] D. Erdener, D. Burnham, Development of auditory-visual speech perception in English speaking children: the role of language-specific factors, in: Proceedings International Speech Communication Association (ISCA) Conference on Auditory-visual Speech Processing, 2005, pp. 57–62.
[10] D. Erdener, D. Burnham, The relationship between auditory-visual speech perception and language-specific speech perception at the onset of reading instruction in English-speaking children, J. Exp. Child. Psychol. 114 (2013) 120–138.
[11] L. Eisenberg, A. Martinez, A. Boothroyd, Auditory-visual and auditory-only perception of phonetic contrasts in children, Volta. Rev. 103 (2003) 327–346.
[12] S. Jerger, L. Lai, V. Marchman, Picture naming by children with hearing loss: II. Effect of phonologically-related auditory distractors, J. Am. Acad. Audiol. 13 (2002b) 478–492.
[13] Y. Chen, V. Hazan, Developmental factors and the non-native speaker effect in auditory-visual speech perception, J. Acoust. Soc. Am. 126 (2009) 858–865.
[14] R. Desjardins, J. Rogers, J. Werker, An exploration of why preschoolers perform differently than do adults in audiovisual speech perception tasks, J. Exp. Child. Psychol. 66 (1997) 85–110.
[15] D. Massaro, Children's perception of visual and auditory speech, Child. Dev. 55 (1984) 1777–1788.
[16] H. McGurk, M. McDonald, Hearing lips and seeing voices, Nature 264 (1976) 746–748.
[17] C. Tremblay, R. Champoux, P. Voss, B. Bacon, F. Lepore, H. Theoret, Speech and non-speech audio-visual illusions: a developmental study, PLoS One 2 (2007) e742, http://dx.doi.org/10.1371/journal.pone.0000742.
[18] K. Sekiyama, D. Burnham, Issues in the development of auditory-visual speech perception: adults, infants, and children, Interspeech 2004 (2004) 1137–1140.
[19] K. Sekiyama, D. Burnham, Impact of language on development of auditory-visual speech perception, Dev. Sci. 11 (2008) 306–320.
[20] S. Jerger, M.F. Damian, N. Tye-Murray, H. Abdi, Children use visual speech to compensate for non-intact auditory speech, J. Exp. Child. Psychol. 126 (2014) 295–312.
[21] A. Boothroyd, Linguistic factors in speechreading, Volta. Rev. 90 (1988) 77–87.
[22] S. Bouton, P. Cole, W. Serniclaes, The influence of lexical knowledge on phoneme discrimination in deaf children with cochlear implants, Speech Commun. 54 (2012) 189–198.
[23] N. Erber, Auditory and audiovisual reception of wordsd in low-frequency noise by children with normal hearing and by children with impaired hearing, J. Speech Hear Res. 14 (1971) 496–512.
[24] N. Erber, Auditory, visual, and auditory-visual recognition of consonants by children with normal and impaired hearing, J. Speech Hear Res. 15 (1972) 413–422.
[25] K. Kirk, L. Prusick, B. French, C. Gotch, L. Eisenberg, N. Young, Assessing spoken word recognition in children who are deaf or hard of hearing: a translational approach, J. Am. Acad. Audiol. 23 (2012) 464–475.
[26] P. Lamore, T. Huiskamp, N. van Son, A. Bosman, G. Smoorenburg, Auditory, visual, and audiovisual perception of segmental speech features by severely hearing-impaired children, Audiol 37 (1998) 396–419.
[27] S. Jerger, N. Tye-Murray, H. Abdi, Role of visual speech in phonological processing by children with hearing loss, J. Speech Lang. Hear Res. 52 (2009) 412–434.
[28] S. Jerger, N. Tye-Murray, M.F. Damian, H. Abdi, Phonological priming in children with hearing loss: effect of speech mode, fidelity, and lexical status, Ear Hear 7 (2016) 623–633.
[29] T. Bergeson, D. Houston, R. Miyamoto, Effects of congenital hearing loss and cochlear implantation on audiovisual speech perception in infants and children, Restor. Neurol. Neurosci. 28 (2010) 157–165.
[30] K. Lalonde, R. Holt, Audiovisual speech perception development at varying levels of perceptual processing, J. Acoust. Soc. Am. 139 (2016) 1713–1723.
[31] L. Elliott, C. Longinotti, D. Meyer, I. Raz, K. Zucker, Developmental differences in identifying and discriminating CV syllables, J. Acoust. Soc. Am. 70 (1981) 669–677.
[32] J. Hornickel, E. Skoe, T. Nicol, S. Zecker, N. Kraus, Subcortical differentiation of stop consonants relates to reading and speech-in-noise perception, Proc. Natl. Acad. Sci. U. S. A. 106 (2009) 13022–13027.
[33] C. Goetzinger, D. Dirks, C. Baer, Auditory discrimination and visual perception in good and poor readers, Ann. Otol. Rhinol. Laryngol. 69 (1960) 121–136.
[34] I. Hoonhorst, V. Medina, C. Colin, E. Markessis, M. Radeau, P. Deltenre, W. Serniclaes, Categorical perception of voicing, colors and facial expressions: a developmental study, Speech Commun. 53 (2011) 417–430.
[35] J. Sussman, A. Carney, Effects of transition length on the perception of stop consonants by children and adults, J. Speech Hear Res. 32 (1989) 151–160.
[36] F. Tsao, H. Liu, P. Kuhl, Speech perception in infancy predicts language development in the second year of life: a longitudinal study, Child. Dev. 75 (2004) 1067–1084.
[37] S. Bouton, W. Serniclaes, J. Bertoncini, P. Cole, Perception of speech features by French-speaking children with cochlear implants, J. Speech Lang. Hear Res. 55 (2012) 139–153.
[38] R. Aslin, L. Smith, Perceptual development, Ann. Rev. Psychol. 39 (1988) 435–473.
[39] A. Carney, Audition and the development of oral communication competency, in: F. Bess, J. Gravel, A. Tharpe (Eds.), Amplication for Children with Auditory Deficits, Bill Wilkerson Center Press, Nashville, TN, 1996, pp. 29–54.
[40] B. Dodd, R. Campbell, Hearing by Eye: the Psychology of Lip-Reading, Lawrence Erlbaum Assoc., London, 1987.
[41] D. Lewkowicz, A. Hansen-Tift, Infants deploy selective attention to the mouth of a talking face when learning speech, Proc. Natl. Acad. Sci. U. S. A. 109 (2012) 1431–1436.
[42] J. Locke, The Child's Path to Spoken Language, Harvard University Press, Cambridge, MA, 1993.
[43] W. Weikum, A. Vouloumanos, J. Navarra, S. Soto-Faraco, N. Sebastian-Galles, J. Werker, Visual language discrimination in infancy, Science 316 (2007) 1159.
[44] H. McConachie, V. Moore, Early expressive language of severely visually impaired children, Dev. Med. Child. Neurol. 36 (1994) 230–240.
[45] L. Menard, C. Toupin, S. Baum, S. Drouin, J. Aubin, M. Tiede, Acoustic and articulatory analysis of French vowels produced by congenitally blind adults and sighted adults, J. Acoust. Soc. Am. 134 (2013) 2975–2987.
[46] A. Mills, The development of phonology in the blind child, in: B. Dodd,

R. Campbell (Eds.), Hearing by Eye: the Psychology of Lipreading, Erlbaum, London, 1983, pp. 145–161.
[47] S. Jerger, M.F. Damian, N. Tye-Murray, H. Abdi, Children perceive speech onsets by ear and eye, J. Child. Lang. 44 (2017) 185–215.
[48] American National Standards Institute (ANSI), Specifications for Audiometers (ANSI S3.6-2004), New York, 2004.
[49] L. Dunn, D. Dunn, D, The Peabody Picture Vocabulary Test-iv, fourth ed., NCS Pearson, Inc, Minneapolis, MN, 2007.
[50] R. Brownell, Expressive One-word Picture Vocabulary Test, third ed., Academic Therapy Publications, Novato, CA, 2000.
[51] B. Dodd, S. Crosbie, B. McIntosh, T. Teitzel, A. Ozanne, Pre-Reading Inventory of Phonological Awareness, The Psychological Corporation Ltd, San Antonio, TX, 2003.
[52] R. Goldman, M. Fristoe, Goldman Fristoe 2 Test of Articulation, American Guidance Service, Circle Pines, MN, 2000.
[53] K. Beery, N. Beery, The Beery-buktenica Developmental Test of Visual-motor Integration with Supplemental Developmental Tests of Visual Perception and Motor Coordination, fifth ed., NCS Pearson, Inc, Minneapolis, 2004.
[54] M. Ross, J. Lerman, Word Intelligibility by Picture Identification, Stanwix House, Pittsburgh, 1971.
[55] N. Tye-Murray, A. Geers, Children's Audio-visual Enhancement Test, Central Institute for the Deaf, St. Louis, MO, 2001.
[56] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, J. R. Stat. Soc. Ser. B Stat. Methodol. 57 (1995) 289–300.
[57] Z. Erden, S. Otman, V. Tunay, Is visual perception of hearing-impaired children different from healthy children? Int. J. Pediatr. Otorhinolaryngol. 68 (2004) 281–285.
[58] N. Tye-Murray, Foundations of Aural Rehabilitation: Children, Adults, and Their Family Members, fourth ed., Cengage Learning, Stamford, CT, 2015.
[59] S. Jerger, R. Martin, J. Jerger, Specific auditory perceptual dysfunction in a learning disabled child, Ear Hear 8 (1987) 78–86.
[60] S. Jerger, M.F. Damian, H. Abdi, Laboratory Report, Speech Lab, School of Behavioral and Brain Sciences, University of Texas at Dallas, 2014.
[61] H. Abdi, The Bonferonni and Sidak corrections for multiple comparisons, in: N. Salkind (Ed.), Encyclopedia of Measurement and Statistics, Sage, Thousand Oaks, CA, 2007, pp. 103–107.
[62] H. Abdi, H.L. Williams, Contrast analysis, in: N. Salkind (Ed.), Encyclopedia of Research Design, Sage, Thousand Oaks, CA, 2010, pp. 243–251.
[63] H. Abdi, Signal detection theory (SDT), in: third ed., in: B. McGaw, P. Peterson, E. Baker (Eds.), International Encyclopedia of Education, vol. 7, Elsevier, New York, 2010, pp. 407–410.
[64] N. Erber, Auditory Training, Alexander Graham Bell Association for the Deaf, Washington, DC, 1982.
[65] L. Eisenberg, Current state of knowledge: speech recognition and production in children with hearing impairment, Ear Hear 28 (2007) 766–772.
[66] H. Abdi, B. Edelman, D. Valentin, D.,W. Dowling, Experimental Design and Analysis for Psychology, Oxford University Press, New York, 2009.
[67] B. Files, B. Tjan, J. Jiang, J.,L. Bernstein, Visual speech discrimination and identification of natural and synthetic consonant stimuli, Front. Psychol. 6 (2015) 878, http://dx.doi.org/10.3389/fpsyg.2015.00878.
[68] M. Hautus, Corrections for extreme proportions and their biasing effects on estimated values of d , Behav. Res. Methods 27 (1995) 46–51.