Children's acoustic and linguistic adaptations to peers with hearing-impairment☆

Sonia Granlund[a,*], Valerie Hazan[a], Merle Mahon[b]

[a]*Speech, Hearing & Phonetic Sciences, University College London (UCL)*
[b]*Language & Cognition, University College London (UCL)*

Abstract

Purpose: This study aims to examine the clear speaking strategies used by older children when interacting with a peer with hearing loss, focusing on both acoustic and linguistic adaptations in speech.

Method: The Grid task, a problem-solving task developed to elicit spontaneous interactive speech, was used to obtain a range of global acoustic and linguistic measures. Eighteen 9- to 14-year-old children with normal-hearing (NH) performed the task in pairs, once with a friend with NH, and once with a friend with a hearing-impairment (HI).

Results: In HI-directed speech, children increased their fundamental frequency range and mid-frequency intensity, decreased the number of words per phrase, and expanded their vowel space area by increasing F1 and F2 range, relative to NH-directed speech. However, participants did not appear to make changes to their articulation rate, the lexical frequency of content words, or to lexical diversity, when talking to their friend with HI compared to their friend with NH.

Conclusions: Older children show evidence of listener-oriented adaptations to their speech production; although their speech production systems are still developing, they are able to make speech adaptations to benefit the needs of a peer with HI, even without being given specific instruction to do so.

Keywords: clear speech, speech adaptations, children's speech, spontaneous speech, referential communication task

1. Introduction

Speech communication often takes place in less than ideal listening environments. Barriers to effective communication can be acoustic, such as background noise in a school setting, or linguistic, such as situations in which people differ from each other in their native language, or in their ability to comprehend a message. Thus an important aspect of a speaker's communicative competence is being able to adapt their speech and language to different situations and to the needs of different interlocutors.

☆This paper was accepted for publication in the Journal of Speech, Language and Hearing Research (http://jslhr.pubs.asha.org/) on 13/10/2017

*Corresponding author; s.granlund@ucl.ac.uk

Previous research has shown that adults are skilled at using these 'clear speaking styles' in adapting to their listener; different types of acoustic and linguistic barriers may elicit different patterns of acoustic-linguistic adaptations by adult speakers (e.g., Cooke and Lu, 2010; Hazan and Baker, 2011). However, little research has explored how speakers adapt to the needs of interlocutors with hearing-impairments (HI) in a communicative interaction, or the development of this ability in children. Therefore the present study examines whether 9- to 14-year-old children change their speech and language according to the hearing status of their interlocutor.

The development of the ability to adapt to a listener's needs is especially important for children who are in frequent contact with peers with HI. For example, although in the UK approximately 85% of children with HI attend mainstream schools (CRIDE, 2015), many children with normal hearing (NH) report not knowing how to communicate with their HI peers (NDCS, 2012a). A recent campaign by the UK National Deaf Children's Society (NDCS) encouraged children to adopt various strategies for talking to peers with hearing-impairments – but very little research has explored the speech communication strategies used by peers with NH and HI when interacting with each other. The importance of robust interaction strategies being used in peer communication is heightened in older children with HI, as peers become increasingly important for a child's social and emotional development in later childhood (e.g., Batten et al., 2014; Antia et al., 2011). Thus it is vital to explore the speech communication strategies being used by children in these environments to ensure that children with HI obtain maximum benefit from inclusive education.

Children with HI experience both acoustic and linguistic barriers in communication. Despite recent advances in hearing aid technology, modern digital hearing aids (HAs) are currently unable to rectify losses in frequency resolution and selectivity in the cochlea caused by sensorineural hearing loss (Moore, 2007) and cannot adequately compensate for losses at higher frequencies (Stelmachowicz et al., 2004). Similarly, cochlear implants (CI) provide reduced spectral resolution and temporal fine structure information compared to a normally-functioning cochlea (e.g., Dorman et al., 2002), can introduce cochlear shifts in the correspondences between frequency and place (Shannon, 2002), and are able to transmit F0 information only weakly (Kuo et al., 2008). On the other hand, due to early auditory deprivation and continued poorer quality of auditory input, children with HI can exhibit delays in receptive vocabulary (Blamey et al., 2001; Uziel et al., 2007) (although see Geers and Sedey, 2011; Ruffin et al., 2013) and even in syntactic knowledge (Spencer, 2004).

To alleviate these problems, the communication partners of children with HI are often instructed on the strategies to use to enhance communication (e.g., Marschark and Hauser, 2012; Caissie and Tranquilla, 2010; NDCS, 2012b; Action on Hearing Loss, 2012). On the global acoustic level, speakers may be asked to increase the intensity of the speech produced; this may be especially beneficial in the mid-frequency 1-3 kHz range which provides important cues to phonetic distinctions (Cooke et al., 2014a). Speakers may also be instructed to increase the salience of their speech by increasing F0 variability, and to decrease their speech rate (Caissie and Tranquilla, 2010).

Another common instruction is to "enunciate carefully" (Caissie and Tranquilla, 2010, p. 100), thus making more extreme articulatory gestures – in vowels, an increased F1 and F2 range, leading to a larger vowel space area and a greater spectral distance between vowels, has been found to positively correlate with the intelligibility of a speaker (e.g., Hazan and Markham, 2004; Bradlow et al., 1996). On the linguistic level, using shorter and less complex utterances (Marschark and Hauser, 2012), increasing sentence contextual information (NDCS, 2012b), and rephrasing rather than repeating the message in cases of misunderstanding (Marschark and Hauser, 2012; Doyle and Dye, 2002; Action on Hearing Loss, 2012) are encouraged.

Many clear speech studies refer to the hyper-hypo (H&H) model of speech production (Lindblom, 1990) to explain speakers' abilities in adapting the acoustic-phonetic properties of their speech. The theory posits that there is a continuum of 'hyper-' to 'hypo-articulated' speech; speakers are driven to maximise the ease of communication between speaker and listener by increasing articulatory effort when the listener has difficulty understanding them (hyper-articulation), but applying as little effort as possible in conditions where there are no communication difficulties (hypo-articulation). Thus speakers dynamically increase or decrease their production effort according to the communicative success of the situation. Similarly, interactionist and usage-based theories of language acquisition (e.g., Vygotsky, 1978; Gallaway and Woll, 1994; Tomasello, 2003) hypothesise that when competent speakers talk to those with less linguistic competence, the ensuing language learning-teaching process can be considered a mutual accomplishment where the more mature speakers adapt their style of talking to the less competent language user, for example by adults modifying the structure of their language to suit the perceived ability of a young child to understand them.

Studies have demonstrated that adults are indeed adept at adapting to the needs of different interlocutors, including to interlocutors of different ages (e.g., infant-directed speech; Cristia, 2010; Kuhl et al., 1997) and of differing linguistic knowledge (e.g., foreigner-directed speech; pet-directed speech; Costa et al., 2008; Burnham et al., 2002), and to interlocutors in different acoustic environments (e.g., Cooke and Lu, 2010; Hazan and Baker, 2011) (c.f., Cooke et al., 2014a, for a review). In HI-directed speech, previous studies suggest that adult speakers make some of the suggested listener-oriented adaptations when asked to 'speak as if to a hearing-impaired listener', as compared to speaking 'casually, as if to a friend' while reading sentences on a screen (e.g., Smiljanić and Bradlow, 2005; Picheny et al., 1986) (for a review, see Smiljanić and Bradlow, 2009). The types of adaptations made by speakers when reading to imaginary listeners with HI may however not reflect those made spontaneously to a real listener in a communicative situation (Hazan and Baker, 2011; Charles-Luce, 1997; Scarborough et al., 2007; Garnier et al., 2010), and read speech studies prevent speakers from using linguistic simplification, such as using more familiar words or shorter sentences, as an adaptation strategy. Very few communicative studies have been conducted on adult speakers' HI-directed speech. Japanese teachers of the deaf playing a game with children with profound hearing losses compared with children with NH simplified their utterances, and also devoiced certain segments, to a greater extent in HI-directed than in NH-directed speech, probably to aid the speech segmentation and understanding of the children with HI (Imaizumi et al., 1993, 1995). Lind et al. (2010) compared non-repair and repair sequences of speech produced by a speaker in conversation with her hearing-impaired husband and observed that, as well as using lexical cues, she increased the maximum F0, mean intensity and duration of words, and the number of pauses in repair sentences – all of these changes would be compatible with the speaker using both acoustic and linguistic enhancements in speech.

In our study, the peers with NH, as more competent speakers than their partner with HI, may adapt their speech to what they consider to be appropriate for the progress of the conversation, effectively becoming like the 'adult' in child-directed speech. However, the extent to which child speakers may be able to adapt their speech to an interlocutor is unclear; the H&H model does not account for the development of this ability in children. Listener-oriented speech production, as posited by H&H, would require the speaker to make greater effort in language production when talking to a peer with HI – this may be difficult for adolescents who can find it challenging to take their interlocutor's perspective, and who still display greater levels of egocentrism than do adults (e.g., Blakemore and Choudhury, 2006).

It seems likely that the development of the ability to make acoustic adaptations to a listener would also require the speaker to have adequate speech coordination and control, which may still be developing in older children. Children's speech production differs from that of adults until at least the early teenage years, likely as a result of increasing experience in and practice with speech production (Koenig et al., 2008; Lee et al., 1999). In general, children have been found to show much greater within-speaker variability than adults (e.g., Nittrouer, 1993; Lee et al., 1999; Nittrouer et al., 2005; Romeo et al., 2013), perhaps due to immature neuromotor control – typically, children's speech gestures are longer and slower than adults' (Smith and Goffman, 1998; Cheng et al., 2007), and motor control of the tongue continues to be refined during adolescence (Cheng et al., 2007). Even 9- to 14-year-olds' speech is not yet adult-like: with age, their speech displays a reduction of F0 range and mean energy at 1-3 kHz, and a faster articulation rate; by age 14, children are found to reach adult-like values for F0 range (Hazan et al., 2016; Pettinato et al., 2016).

Although children do not produce adult-like vocal characteristics or articulatory variability until late childhood, this does not necessarily imply that they will not have control over these characteristics of their speech until that age – in fact, older children may already be adept at making many acoustic adaptations, as they may be somewhat automated processes, similar to the Lombard effect (Lombard, 1911; Amazi and Garber, 1982; Summers et al., 1988). Few studies have investigated whether children are able to make acoustic speech adaptations. There is some evidence that 4-year-old children use a slower speech rate, but not an increased F0 range (Weppelman et al., 2003) when conversing with infants and toddlers compared to peers or adults. Three to 5-year-olds increased their F0 range, used longer and more intense vowels and a larger vowel space when teaching a puppet words compared to telling an adult about them (Syrett and Kawahara, 2014), but in another study, children of the same age did not produce any differences in vowel formants when asked to speak 'clearly' compared to speaking in a casual manner (Redford and Gildersleeve-Neumann, 2009). At age 16, adolescents were found to increase their F0 range and the duration of certain phonemes in infant-directed compared to adult-directed speech (Kempe, 2009). Hazan et al. (2016) and Pettinato et al. (2016) investigated 9- to 14-year olds' peer communication through a normal channel and through an intelligibility-reducing noise-excited vocoder in a problem-solving task, and compared their performance to that of adults. Children changed their articulation rate, F0 median and intensity to a similar extent to adults when talking to a friend through the vocoder compared to talking normally. Unlike adults, children also increased their F0 range in the vocoder condition, although this strategy is unlikely to be helpful to the listener, as the vocoder does not transmit F0 information. Similarly, only older children were found to increase the size of their vowel space in the vocoder condition compared to the no-barrier condition. These results suggest that although children are able to make speech modifications in adverse listening conditions, even at the age of 14 children's speech modifications are not as attuned to the listener as adults' are.

Thus, there is some evidence that the older children tested in this study may be able to make at least some acoustic adaptations to their listener, albeit not as extensively as adults. However, for a speaker to make linguistic adaptations to listener needs may require advanced cognitive processing, as well as well-developed pragmatic and linguistic awareness. Children would also need advanced enough vocabulary and syntax to be able to generate alternative ways of producing their utterances; these are also likely to develop with age and with greater linguistic experience (e.g., Nippold et al., 2005) even until the late teenage years. There is some evidence that 2- to 6-year-old children use shorter utterances in communication with infants and toddlers than with adults and peers (Shatz and Gelman, 1973; Sachs and Devin, 1976), but other research shows that linguistic adaptations may develop fairly late – studies on children's 'repair responses' after clarification requests have shown that 5- to 8-year-old children's ability to revise the form or content of their original message increases with age, and may not yet be adult-like even at age 8 (Peterson et al., 1972). Similarly, Brinton et al. (1986) discovered that children started to use revision strategies in response to clarification requests only after age 7.

Very little research has been done on the communicative adaptations made by children to peers with HI. Historically, studies have generally investigated few participants, usually of preschool age, and used mostly subjective discourse measures to assess communication (Vandell and George, 1981; Spencer et al., 1994; Arnold and Tremblay, 1979; Seewald and Brackett, 1984; Vandell et al., 1982). Some of these studies found evidence of modification of hand gestures by even young children when communicating with peers with HI (Spencer et al., 1994; Arnold and Tremblay, 1979). Seewald and Brackett's (1984) case study of a 6-year-old participant in interaction with an adult, a toddler, a peer with NH and a peer with HI revealed that she used less complex syntax and more directives when talking to the peer with HI than to the peer with NH, and she shortened her utterances to a similar extent to the peer with HI as to the toddler.

Although the studies reviewed above suggest that even preschool-aged children are able to adapt to a listener, adaptation does not seem adult-like even in later childhood. However, most previous studies use only limited measures to examine adaptations, and all but a few examined strategies used by younger children. Importantly, despite the fact that communication with peers with HI involves both an acoustic and linguistic barrier, no previous studies have examined both the global acoustic and linguistic modifications made by children speaking to HI peers. The aim of the current study is therefore to investigate, using interactive spontaneous speech and a comprehensive set of measures, whether children adapt their speech to peers with HI. We use a within-subjects design to enable us to explore the differences in communication strategies used by each child when talking to a friend with NH compared to a friend with HI. This is achieved using a new communicative problem-solving task which was developed for the elicitation of both global acoustic and linguistic measures of speech. We also explore whether the age of the speaker affects the extent to which modifications are made.

2. Method

2.1. Participants

All participants were students at mainstream primary and secondary schools in Southern England which included units for pupils with hearing-impairments. The study was carried out in pairs – each participant with NH participated in one session with a friend with NH (NH-directed condition) and one session with a friend with HI (HI-directed condition). Two participants with NH and two participants with HI were recruited per school.

Eighteen children with NH (age range: 9.0-14.4 years; *M*: 11.9 years; 11 female) participated in the study. Fifteen participants were monolingual Southern British English speakers, and three were bilingual. However, all had received their entire schooling in English, and English was their dominant language. One participant had mild additional needs not affecting language production, and the remaining participants were not reported to have any neurological or medical conditions. Two participants were excluded from analyses: one due to speech impediment and another due to equipment malfunction leading to unusable audio recordings in the study.

In the HI-directed condition, the participants with NH carried out the task with eighteen children with HI (age range: 9.7-15.2 years; *M*: 12.0 years; 10 female). Thirteen were monolingual native Southern British English speakers, four were bilingual, and one used British Sign Language (BSL) as her first language. Their level of hearing loss ranged from moderate (4 participants) to profound (7 participants); 7 participants used one or two CIs, while the others wore bilateral hearing aids. All had sensorineural hearing loss; one participant had a mixed loss. Three participants had mild additional needs. The participants with HI used mostly oral communication with their parents, although some also used total communication. Only five out of the 18 participants with HI were assessed by their teachers as having age-appropriate language and communication skills.

To make the study as ecologically valid as possible, and to control for degree of exposure to speech produced by children with

HI, we ensured that all participants in each pair were friends prior to the recordings. The participants with NH were generally in regular contact with peers with HI ('How often during a typical week do you talk to people who are deaf?'; M: 4.2, on a scale of 1-not at all to 5-all the time), and all had at least some friends with HI ('How many of your friends are deaf?'; M: 2.5, on a scale of 1-none to 5-all). Most participants with NH usually used spoken English to communicate with peers with HI, but some also reported combining speech and hand gestures, and/or British Sign Language (BSL) signs.

2.2. Materials

A new referential communication task, the Grid task, was designed to be a complex enough task to enable analyses of global acoustic, linguistic and phonetic measures of speech.

In the task, each participant is given a picture-grid, an empty grid with colour-number squares (see Fig.1) and a tray containing five different pictorial versions (see Fig.2) of 16 keywords. The aim of the task was for each conversational partner, without being able to see each others' grids or trays, to replicate the other's picture-grid in their empty grid, by finding (1) the correct keyword, (2) the correct version of the keyword, and (3) the correct location (i.e., the colour-number square) of the keyword. Each participant pair completed between two and four grids.

One of the aims of the task was to elicit productions of three segmental contrasts (the high front vowel contrast /i/-/ɪ/, the bilabial voicing contrast /p/-/b/ and the sibilant place distinction /s/-/ʃ/) which have typically been found to be difficult for children with HI to produce and perceive (e.g., Bernhardt et al., 2003; Kosky and Boothroyd, 2003; Uchanski and Geers, 2003; Stelmachowicz et al., 2004; Borg et al., 2007; Tsui and Ciocca, 2000). Encountering these contrasts will therefore likely lead to frequent misunderstandings between interlocutors, thus necessitating speech adaptation by the friends with NH in HI-directed compared to NH-directed speech. Although these analyses are not reported here, the frequent production of these contrasts in the task will enable us to investigate the participants' speech adaptation also on the phonetic level in the future.

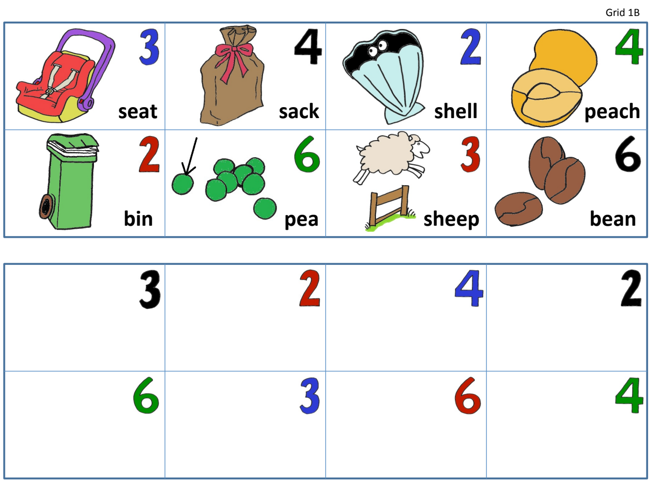


Figure 1: An example picture-grid and empty grid from the Grid task. These were coloured in multiple colours in the task.

Sixteen keywords forming minimal pairs were created (three per contrast, see Table 1). Keywords were chosen to be common vocabulary items which could be represented pictorially as concrete objects, and which 9-year-old children would know. No filler items were used in the experiment.

Five versions of each keyword were hand-drawn, scanned and coloured in (see Fig.2).

/p-b/	/s-ʃ/	/i-ɪ/
pin-bin	cell-shell	bean-bin
peach-beach	seat-sheet	peach-pitch
pea-bee	sack-shack	sheep-ship

Table 1: The sixteen minimal pair keywords used in the Grid task.

These versions differed from each other in representing either different types of a certain object (e.g., for 'pitch'; rugby, football, baseball, and cricket pitches), and/or in being the same object but differing in detail (i.e., for 'bee': bees with different numbers of stripes, faces or kinds of antennae and wings). A pack of five laminated cards was made for each keyword.

To enable the elicitation of several instances of different vowels, for use in vowel space analysis, colour-number words ('green', 'red', 'black', 'blue'; 'three', 'six', 'four', 'two') reflecting as wide a range of vowels as possible were chosen for use in the task. The colours and numbers were combined in all possible ways, leading to 16 colour-numbers used in the task.

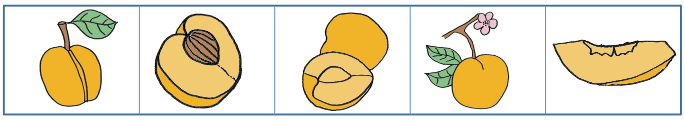


Figure 2: An example of the five versions of the keyword 'peach' in the task. These were coloured in multiple colours in the task.

The 16 keywords and 16 colour-numbers were used to build eight pairs of two-by-four grids (see Fig.1) (pictorial versions of all grid sets can be found in Granlund, 2015). The keywords and colour-number words were distributed between the grids so that each grid included keywords with several different target sounds. Several randomisation processes were applied to the keywords and colour-numbers in the grids, (1) to reduce the probability of a contrastive accent on the words, (2) to ensure that upcoming keywords would not be predictable, and (3) to evenly distribute the keywords and colour numbers within the grids to avoid position effects in the data. The keywords were written on the picture-grids (see Fig.1), but not on the five card versions made to each keyword (see Fig. 2). However, the five card versions were presented in a plastic tray, with the cards for each keyword taking up one slot – here, the keyword for each slot was written to enable the participants to find the correct keyword easily.

2.3. Procedure

Each participant took part in the study for approximately 3.5 hours across three sessions; two of these sessions involved communication tasks done with another participant, and one session consisted of speech production and perception tasks which were completed alone[1]. All sessions were recorded during school hours at the children's schools in a quiet room. In each group of four participants from each school, each participant completed the first session either in a 'same hearing status' pair (HI-HI and NH-NH pairs) or a 'different hearing status' pair (NH-HI pairs), and the order of 'same' and 'different' pairs was counterbalanced between groups. Note that although participants with HI completed a session with a friend with HI, the analyses of this condition are not reported in this paper.

In the communication tasks, each participant wore an Audiotechnica AT8531 lapel

[1]The data acquired in this session are not reported in this paper.

microphone which was connected to a Scarlett 2i2 USB audio interface which fed into a laptop running Audacity with a sampling rate of 44,100 Hz (16 bit). The participants were sitting at a table facing each other, approximately 1-1.5 metres away from each other. A video camera was positioned behind each participant, which enabled video recordings to be made of the participant sitting opposite.

In the two communication sessions, the pair of participants completed two sets of grids from the Grid task, followed by another communication task not reported here, and finally a further two sets of grids together. Due to time constraints, it was not always possible to complete all tasks as planned – however, all pairs completed at least two sets of grids. To enable analyses to be done on their spoken language skills, participants were asked not to use BSL signs in the communication sessions. However, due to our desire to maintain naturalistic communication between participants, the use of hand gestures was allowed.

In the Grid task, each participant was given a board which stood upright between the participants; the participants could see each others' faces but not the front of each others' boards. The board had a picture-grid and an empty grid on it, and the participant was also given a tray with the 16 Grid keywords.

Before the start of the task, the participants were shown a slideshow which instructed them on how to complete the task. They were asked to work from the top left hand corner horizontally, row by row, with the conversational partners taking it in turns to describe the squares on their picture-grid to each other. The participants completed a few squares of a pair of practice grids together – these contained another set of words which were not minimal pairs, and which therefore could be used to ensure that the participants understood the task.

The participants' familiarity with the task's keywords was also checked. Initial piloting showed that some 10-year-old pilot participants were not familiar with the keywords 'cell' and 'shack' and therefore, in the study, the experimenter gave a simple explanation of these two words along with the pictures to each participant pair. All participants readily used all the keywords in the task.

The order of the grids given to the participants was randomised, and no grids were presented more than once to each participant. The participants were shown the correct answer sheet for each of their completed grids at the end of each task. On average, 28 minutes of Grid task conversation was recorded and 3.5 grids were completed, per participant per condition.

The same procedure was used in both communication sessions, except that in the second communication session the participants were not given an introduction to the task or asked to complete a practice picture. Each communication session lasted between 1 and 1.5 hours. Although the communication sessions took time to complete, the participants enjoyed them as they were framed to them as a game to be played with their friend. Each pair of grids usually took between 5 and 15 minutes to complete, and participants were given breaks between pairs of grids and between tasks. Additionally, the pictures used in the task were made to be colourful, and even humorous, for the children's benefit.

2.4. Preliminary file processing

Each participant's speech was saved to a separate channel. In total, 14.9 hours (111 files) of single-channel recordings of the participants with NH were made.

The dual-channel recordings were transcribed orthographically by the first author using Praat (Boersma and Weenink, 2015). The transcription criteria followed the general guidelines of those used in Hazan and Baker (2011), Pettinato et al. (2016) and Hazan et al. (2016). In short, any speech that occurred during external noise or laughter was tagged as such, and words spoken only partially were given their own label; instances of laughter, whispering, external noise, or breaths in the microphone were also labelled. If a word was unintelligible to the transcriber, it was tagged as such. Any speech overlap between participants was also tagged. Within-speaker pauses over 500ms in length were marked as 'SIL'; 500ms was chosen as the minimum within-speaker duration as it is a typical silence threshold for automatic silence detectors (c.f., Heldner and Edlund, 2010).

The transcribed utterances from each speaker's tier were extracted using a Praat script and converted to text files. After further processing using Python scripts, the utterances were aligned to the single-channel waveform using automatic alignment software developed at UCL by Huckvale and Iverson based on the Hidden Markov Model Toolkit (HTK Team, 2012). The aligner created TextGrid files for the utterances with word- and phone-level tiers. Each single-channel wav file, and its corresponding TextGrid file with phoneme- and word-level tiers was then used for analysis.

2.5. Data processing

2.5.1. Transaction time

If children found the task more difficult when completing it with a friend with HI compared to a friend with NH, it is more likely that they would need to enhance their communication strategies in HI-directed compared to NH-directed speech. Therefore, to obtain a measure of task difficulty, the total time taken for each participant pair to complete each pair of grids, excluding the duration of any talk by the experimenter, was divided by the total number of correct picture-squares per grid per participant pair to obtain the mean time taken to find one correct picture in the Grid task. Any outliers below or above 3 standard deviations from the mean per each pair type were excluded from the analysis.

2.5.2. Global acoustic and linguistic measures

The acoustic measures (F0 range, mean energy 1-3 kHz, articulation rate, vowel space) and linguistic measures (number of words per phrase, lexical frequency of content words, lexical diversity) were calculated for each file separately to obtain one value per file per participant per condition. Any part of the speech signal containing the interlocutor's speech (either on its own or spoken simultaneously with the participant) or unintelligible words was not analysed. With the exception of the vowel space measures, any outliers over or below 3 standard deviations from the mean of each condition were excluded from analysis.

F0 range. To investigate whether speakers expanded their F0 range in HI-directed compared to NH-directed speech, as often found in clear speech studies (c.f., Smiljanić and Bradlow, 2009), the fundamental frequency (F0) range in all files was measured using a Praat script. The script calculated the interquartile range for each file in semitones re 1 Hz, using a time step of 150 values per second. As in Hazan and Baker

(2011) and Hazan et al. (2016), semitones were used as an attempt to normalise for speaker and sex across participants. The accuracy of the automatic F0 estimation was visually checked for a subset of the files, and no major problems were detected. F0 analyses were not separated by sex, as previous studies show that significant sex differences in F0 appear after approximately age 12 years (Hollien et al., 1994), and only one of the male speakers in this study was above that age. Additionally, sex did not significantly affect F0 range for 9-14-year-olds in Hazan et al. (2016).

Mean energy 1-3 kHz. As in Hazan et al. (2016), for each file, speech intensity was measured using a Praat script, which calculated the mean energy between 1 and 3 kHz. This frequency region was chosen as previous research (e.g., Krause and Braida, 2004) has demonstrated that speakers increase their relative energy in this frequency range in clear speech. To ensure that words which were clipped were not analysed, words for which portions of the signal were over 88 dB were excluded from analysis. The remaining words were concatenated and normalised for peak intensity (to 75 dB) before being band-pass filtered between 1 and 3 kHz. The mean energy per file between those frequencies was then calculated.

Articulation rate. To analyse whether speakers changed their articulation rate when talking to a friend with HI compared to a friend with NH, the duration of all words except agreement, hesitation and exclamation words (such as 'yeah', 'err' and 'ooh') was measured using a Praat script. Then, the number of syllables in the same words was calculated using the qdap scripts in R (Rinker, 2013). The total number of syllables was divided by the total duration of words to determine the number of syllables produced per second in each file. It would be possible that children with NH use complex words with more syllables when talking to a friend with NH than when talking to a friend with HI – articulation rate is therefore used as it is subject to less influence from the inherent length of words than is speech rate.

Vowel space measures. Vowel space measures were taken to investigate whether participants increased spectral distance between vowels in HI-directed compared to NH-directed speech, as typically found in clear speech studies (e.g., Bradlow, 2002). Thus, the midpoint F1 and F2 of /i/, /æ/ and /ɔ/ vowels in content words were measured using a Praat script. These vowels were chosen due to them being at the extremes of the vowel space for Southern British English speakers (e.g., Pettinato et al., 2016), and due to their high frequency of occurrence in the speech elicited in the Grid task.

For /æ/ and /ɔ/, the default formant tracking settings on Praat[2] were deemed to produce accurate formant values. However, for the /i/ vowel, these settings were found to produce erroneous F2 tracks because of the high /i/ F2 values produced by the participants. Therefore, for the vowel /i/, the Praat script was rerun with an F2 reference point of 3000Hz for female and 2800Hz for male speakers[3]. Due to the vari-

[2]Females - F1: 550Hz; F2: 1650Hz; males - F1: 500Hz, F2: 1485Hz

[3]This was done due to previous findings showing that male and female children's F1 and F2 ranges significantly differ from each other (c.f., Pettinato et al., 2016).

ation in values produced by potential erroneous tracking, for each vowel, F1 and F2 values above or below two standard deviations from the mean per person per condition were excluded from analysis. An R script was then used to transform the values to equivalent rectangular bandwidth (ERB) values. ERB is traditionally classified as a 'vowel-intrinsic' normalisation method (c.f., Adank et al., 2004), and as such, is a better reflection of human psychophysical perception than are Hz values. ERB was used for vowel normalisation because it does not require values from vowels in the entire vowel space, unlike 'vowel-extrinsic' normalisation methods. A mean of 29.6 /æ/, 18.3 /ɔ/ and 79.8 /i/ vowels per speaker per condition were used.

To obtain a value for vowel space area, the mean distances between /i-æ/ (a), /i-ɔ/ (b) and /ɔ-æ/ (c) (the vowel triangle perimeter) were calculated, and (a) (b) and (c) were summed (s). The vowel space area was calculated using Heron's method, by taking the square root of s*((s-a)*(s-b)*(s-c)). One measure of vowel space area was obtained per speaker per condition.

In addition to the vowel space area measure, F1 range was calculated to examine the extent to which speakers increase their vowel height in HI-directed speech. This was done by subtracting the mean F1 of /i/ (vowel with lowest F1 value) from the mean F1 of /æ/ (vowel with the highest F1 value in this vowel set). One F1 range measure per speaker per condition was obtained. F2 range was calculated to investigate whether speakers increased the front-back distance of their vowels in HI-directed speech. This calculation was done per speaker per condition by subtracting the mean F2 of /ɔ/ from the mean F2 of /i/, due to these vowels having the highest and lowest F2 values within the three analysed vowels.

Number of words per phrase. The mean number of words per phrase was used as a measure of the complexity of the sentences produced by the speakers. We investigated whether speakers decreased the mean number of words per phrase in HI-directed compared to NH-directed speech using a Praat script. Phrases were defined as words occurring between two silences, which were either within-speaker silences (SIL) over 500ms in length, or between-speaker silences (SILP) (i.e., silences in which the interlocutor takes the next turn). Agreement, hesitation and exclamation words, such as 'yeah','umm' or 'ooh', were not counted as part of a phrase, as the aim was to include only utterances with meaningful content. These criteria are similar to those used in Shatz and Gelman (1973). Analyses were done on the mean number of words per phrase of each file in each condition.

Lexical frequency of content words. The lexical frequency of words is here taken as a measure of the complexity of the vocabulary used by the speakers; we examined whether children decreased the complexity of words used when speaking to a friend with HI compared to a friend with NH. A Python script was used to calculate the mean lexical frequency of words per speaker per file. Lexical frequency was determined using standardised word frequency on the logarithmic Zipf scale (van Heuven et al., 2014) from a list of word frequencies derived from subtitles from UK primary school children's television programmes (lexical frequency of CBBC channel subtitles in $SUBTLEX_{UK}$) (van Heuven et al., 2014). Part-of-speech information was obtained from the same list, and only content words (nouns, adjectives, adverbs and verbs) were included in the calculation of mean lexical frequency. The lexical frequency of words which occurred several times in the file were only counted once. Of contractions, only the main part of the word was used for calculating lexical frequency. Any content words that were not included in the $SUBTLEX_{UK}$ word list were excluded from the frequency calculation (mean number of word exclusions per file: 0.3).

Lexical diversity (VOCD). In this study, lexical diversity is used to investigate whether speakers used lexical repetition (low diversity) or revision (high diversity) strategies in HI-directed compared to NH-directed speech. The measure has previously been used to assess the diversity of vocabulary used by children and second language learners (Lu, 2012; Richards, 1987). A speaker needs to use a range of vocabulary items and repeat the same words rarely to obtain a high lexical diversity score. Lexical diversity was calculated using the Lingua-Diversity package (Xanthos, 2011) on Perl. Agreement, hesitation and exclamation words were excluded from analysis. The package calculates lexical diversity using an implementation of VOCD from McKee et al. (2000). Traditional methods for measuring lexical diversity typically rely on the type-to-token ratio (TTR), i.e., the number of different words (types) divided by the number of total words. However, the TTR is strongly correlated with the total number of words in the file (McKee et al., 2000). Instead, VOCD is calculated by taking the TTR of several random subsets of words in the file, and then finding the value of the curve ('D') that best fits the generated TTR x sample size curve. D therefore reflects lexical diversity over the entire file, with higher values indicating high lexical diversity. The VOCD measure is reported to correlate with other child language measures (McKee et al., 2000).

3. Results

3.1. Statistical approach

A linear mixed effects model approach was used for all the statistical analyses reported in this paper. The lmer function in the lme4 package (Bates et al., 2015) for R (R Core Team, 2014) was applied to build each model with a bottom-up hierarchical approach, in which each predictor is added one-by-one to the baseline model. Comparison of model residuals was computed with Chi-square-tests using the likelihood ratio test, with α set at 0.05. For each model which produced a significant result, model summary information can be found in Tables 3 to 10 in the Supplemental materials.

3.2. Effect of pair type on task transaction time

First, to assess whether communicating with a friend with HI compared to a friend with NH was indeed more difficult for the participants, an analysis of task transaction time was made.

For this measure, 'pair type' (NH-NH and NH-HI), and the pair's mean age as a continuous variable, as well as the interaction between pair type and mean age, were included as fixed effects in the model, while 'pair' and 'grid number' were counted as random effects using random intercepts only.

There was a significant main effect of pair type (NH-NH, NH-HI)($\chi^2(5) = 9.11$; p=0.0025), with NH-HI pairs taking longer to complete each cell correctly in the Grid task (M: 40.10 s) than NH-NH pairs (M:

22.33 s) (see Fig.3). The pair's mean age had a near-significant effect on transaction time ($\chi^2(6) = 3.71$; p=0.054), with faster transaction times for older pairs (see Table 3 for model summary). The interaction between pair type and age ($\chi^2(7) = 0.31$; p=0.58) was not significant.

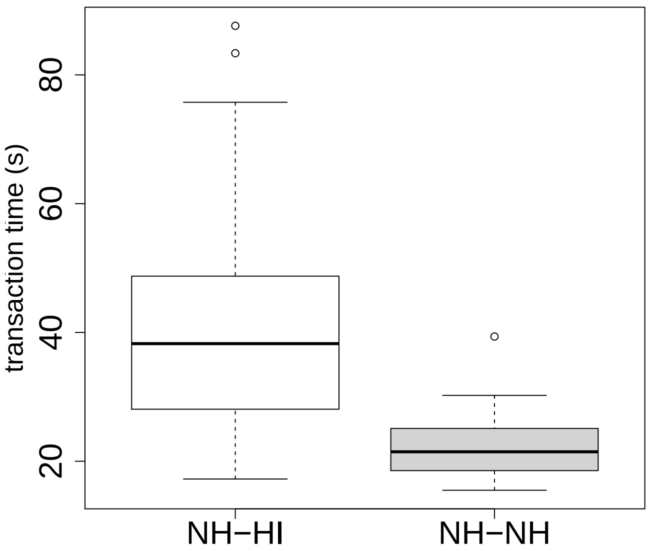


Figure 3: Mean time to find each picture correctly in the Grid task (in seconds) for NH-HI and NH-NH pairs.

This result implies that speakers have greater difficulty with the task when completing it with an interlocutor with HI compared to an interlocutor with NH. Therefore it is likely that the speakers were required to change their speech and language to adapt to the needs of their interlocutor. The following sections ascertain whether the speakers adapt the global acoustic and linguistic aspects of their speech in HI-directed compared to NH-directed speech.

3.3. Acoustic and linguistic adaptations

In the analysis of the global acoustic and linguistic measures, for each dependent variable, listener hearing status ('directed'; NH-directed, HI-directed), age as a continuous variable, and the interaction between 'directed' and age were included as fixed effects in the model.

As, for these measures, each speaker contributed two to four data points depending on the number of Grid pictures completed in each condition, 'speaker' and 'grid number' were treated as random effects, with by-speaker correlated random intercepts and slopes for the fixed effect 'directed', and by-grid number random intercepts. Although the order of the grids was not part of the statistical models reported here, no differences in the results were obtained when each model was rerun with 'grid order' rather than 'grid number' as a random effect.

For the vowel space measures, as for the other global acoustic measures, listener hearing status ('directed'), age as a continuous variable, and the interaction between 'directed' and age were included as fixed factors in the mixed effects models for each dependent variable. Because measures of vowel space were calculated over all completed grid pictures per condition, resulting in only one data point per condition per speaker, only 'speaker', without random slopes, was treated as a random effect.

The means and standard deviations of each measure in the NH-directed and HI-directed conditions are presented in Table 2, and boxplots of the global acoustic and linguistic measures in each condition are found in Figure 4.

measure	NH-directed	HI-directed
F0 range (st)	2.73 (0.69)	3.30 (0.82)
Articulation rate (sylls/s)	3.59 (0.52)	3.50 (0.56)
Mean energy 1-3k Hz (dB)	62.13 (2.11)	64.04 (2.39)
Words per phrase	4.54 (1.05)	4.18 (0.94)
Lexical frequency	5.16 (0.13)	5.19 (0.11)
Lexical diversity (VOCD)	34.61 (8.06)	38.10 (10.23)
Vowel space area (ERB^2)	20.74 (2.94)	22.67 (3.65)
F1 range (ERB)	4.92 (0.45)	5.20 (0.60)
F2 range (ERB)	8.88 (0.56)	9.17 (0.60)

Table 2: Means and standard deviations (in parentheses) for each global acoustic and linguistic measure in NH-directed and HI-directed conditions.

3.3.1. Acoustic adaptations

F0 range. The speakers significantly increased their F0 range in HI-directed speech (M: 3.30 st) compared to NH-directed speech (M: 2.73 st) ($\chi^2(7) = 8.86$; p=0.0029) (see Fig. 4a). There was also a near-significant main effect of age ($\chi^2(8) = 3.75$; p=0.053) – speakers' F0 range decreased with age (see Table 4 for model summary). The interaction of directed*age was, however, not significant ($\chi^2(9) = 0.19$; p=0.66).

Articulation rate. For articulation rate, there were no significant main effects of 'directed' ($\chi^2(7) = 0.96$; p=0.33) or age ($\chi^2(8) = 0.63$; p=0.43), and no significant interaction of directed*age ($\chi^2(9) = 0.37$; p=0.54), indicating that the speakers did not decrease their articulation rate when talking with a friend with HI, compared to a friend with NH.

Mean energy 1-3 kHz. The participants spoke with greater relative intensity at 1-3 kHz when talking with a friend with HI (M: 64.04 dB) than with a friend with NH (M: 62.17 dB) ($\chi^2(7) = 20.31$; p<0.00001) (see Fig.4c). The speakers also decreased their mid-frequency intensity with age ($\chi^2(8) = 7.13$; p=0.0076)(see Table 5 for model summary). There was no significant interaction of the two main effects ($\chi^2(9) = 0.48$; p=0.49).

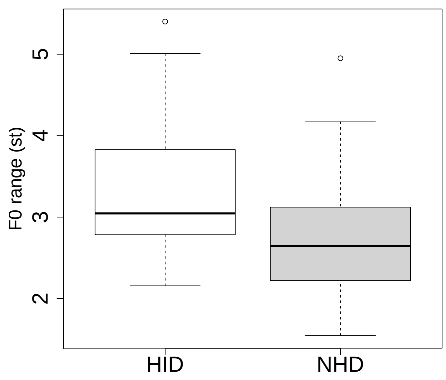
Vowel space area. Vowel space area was significantly expanded in speech directed to a friend with HI (M: 22.67) compared to speech directed to a friend with NH (M: 20.74) ($\chi^2(4) = 9.94$; p=0.0016) (see Fig. 5). Participants also produced smaller vowel spaces with increased age ($\chi^2(5) = 5.89$; p=0.015) (see Table 8 for model summary). The interaction of 'directed' and age was not significant ($\chi^2(6) = 1.00$; p=0.32).

F1 range. Speakers increased their F1 range in HI-directed (M: 5.20) compared to NH-directed speech (M: 4.92) ($\chi^2(4) = 6.98$; p=0.0082), and F1 range decreased with age ($\chi^2(5) = 7.04$; p=0.0080) (see Table 9 for model summary). However, there was no directed*age interaction ($\chi^2(6) = 1.14$; p=0.29).

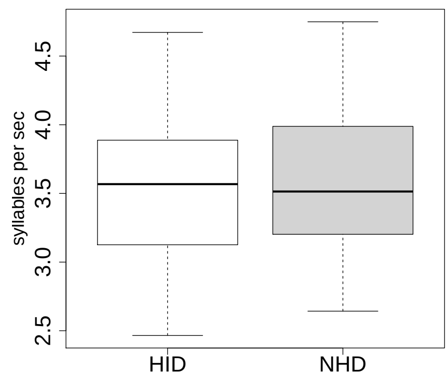
F2 range. Speakers increased their F2 range in HI-directed (M: 9.17) compared to NH-directed speech (M: 8.88) ($\chi^2(4) = 4.00$; p=0.046) (see Table 10 for model summary). The effect of age was only near-significant ($\chi^2(5) = 3.55$; p=0.060). There was no significant interaction of 'directed' and age ($\chi^2(6) = 0.052$; p=0.82).

3.3.2. Linguistic adaptations

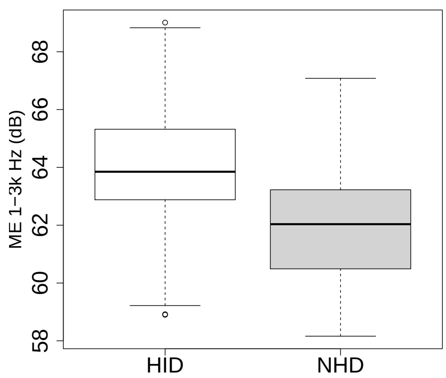
Number of words per phrase. The speakers slightly decreased the number of words used per phrase in the HI-directed condition (M: 4.18) compared to the NH-directed condition (M: 4.54) ($\chi^2(7) = 4.69$; p=0.030) (see Fig. 4d, and Table 6 for model summary). The main effect of age ($\chi^2(8) = 1.51$;



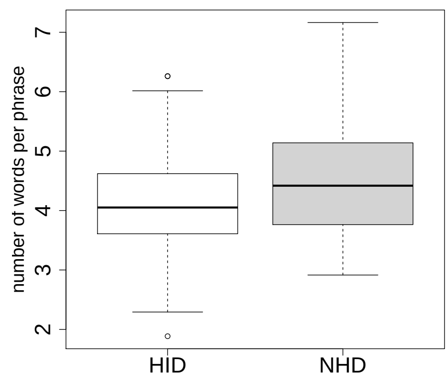
(a) F0 range (st)



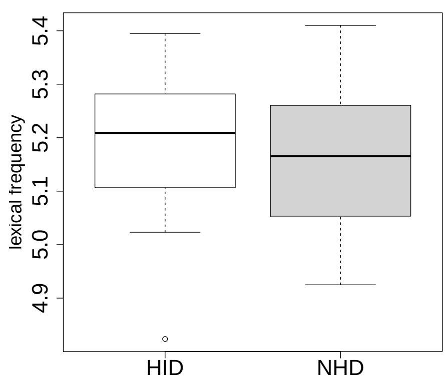
(b) Articulation rate (sylls/sec)



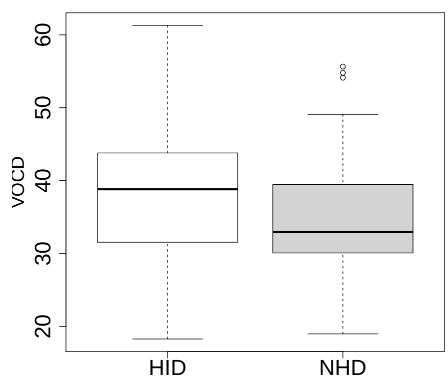
(c) Mean energy 1-3 kHz



(d) Number of words per phrase



(e) Lexical frequency



(f) Lexical diversity (VOCD)

Figure 4: Boxplots showing each acoustic and linguistic measure in HI-directed (white) and NH-directed (grey) conditions.

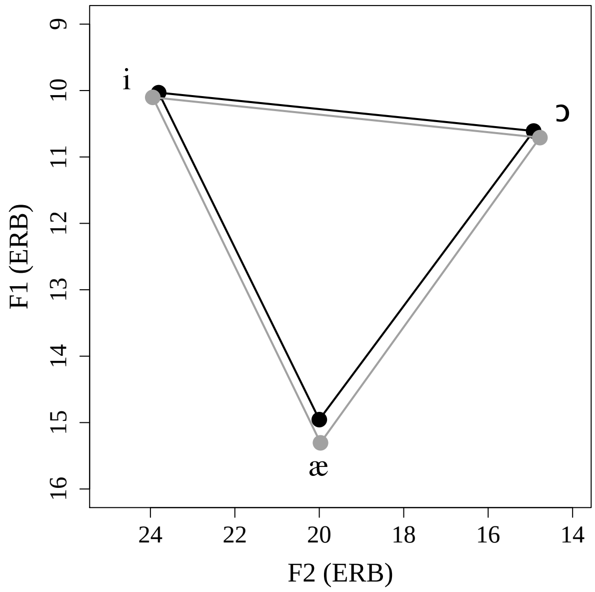


Figure 5: Mean vowel space area in the NH-directed (black points) and HI-directed (grey points) conditions.

p=0.22), and the interaction of directed*age ($\chi^2(9) = 1.52$; p=0.22), were not significant.

Lexical frequency of content words. For lexical frequency, there was no significant effect of either 'directed' ($\chi^2(7) = 2.03$; p=0.15) (see Fig. 4e) or of the interaction of directed*age ($\chi^2(9) = 0.91$; p=0.34). However, a significant main effect of age ($\chi^2(8) = 6.95$; p=0.0084) demonstrated that with increased age, speakers used less frequent content words in the Grid task (see Table 7 for model summary).

Lexical diversity (VOCD). No significant main effects of 'directed' ($\chi^2(7) = 2.23$; p=0.14) (see Fig. 4f) or age ($\chi^2(8) = 1.74$; p=0.19), and no significant interaction of directed*age ($\chi^2(9) = 1.96$; p=0.16), were found, implying that speakers did not produce more diverse language in the HI-directed compared to the NH-directed condition.

3.4. Summary

In summary, findings suggest that the speakers adapted their speech to an interlocutor with HI by increasing their F0 range, the mid-frequency intensity in their voice, and their vowel space in both F1 and F2 dimensions, and by decreasing the number of words used per phrase in the HI-directed condition compared to the NH-directed condition. Nonetheless, in HI-directed speech relative to NH-directed speech, no evidence was found for a decrease in articulation rate, or an increase in lexical frequency or diversity. With age, speakers decreased their F0 range, mid-frequency intensity, the size of their vowel space and the frequency of content words used in the task; however, the lack of significant interactions with age imply that the age of the speaker did not affect the extent to which speech adaptations were made for any of the investigated acoustic or linguistic measures.

4. Discussion

This study investigated whether 9- to 14-year-old children adapt their speech and language to benefit a listener with HI and, if so, which strategies are used. Unlike most previous studies, in which a clear speaking style is elicited by instructing adult speakers to read sentences as if speaking to a person with HI compared to a friend (e.g., Ferguson and Kewley-Port, 2007; Smiljanić and Bradlow, 2005), in this study, speech was elicited using a more ecologically valid method – participants interacted face-to-face with a friend with either HI or NH on a problem-solving task similar to those used in an everyday school setting. Children with HI experience both an acoustic and linguistic barrier in communication, and therefore, in this study, the Grid task was successfully devised to enable the analysis of a range of global acoustic and linguistic measures to be made from the participants' speech. The participants were not given instructions on how to communicate with each other on the task, but interacted spontaneously.

Previous literature indicates that children's global speech production is likely to differ from adults' until at least the early teenage years (Hazan et al., 2016; Lee et al., 1999), and displays greater within-speaker variability than adults', probably due to immature speech motor control and inexperience in speech production (e.g., Koenig et al., 2008; Lee et al., 1999). The age effects observed in the participants' speech in this study are in line with previous studies – as in Hazan et al. (2016), the children in the current study were found to decrease their F0 range and their mid-frequency intensity with age. Children's vowel spaces reduced with age, similarly to findings of Pettinato et al. (2016). As expected, due to their continued vocabulary development (e.g., Nippold et al., 2005), younger children also used higher frequency content words than did older children.

However, as in Hazan et al. (2016), older children's still-developing speech production systems did not appear to prevent them from making adaptations to the listener. Results demonstrate that children with NH had more difficulty in the task when completing it with a friend with HI than with a friend with NH, and were likely trying to counteract this difficulty by using certain acoustic and linguistic strategies to spontaneously adapt to the needs of their friend with HI – they increased their F0 range, the mid-frequency intensity of their speech and their vowel space area, and decreased the number of words used per phrase. These adaptations did not interact with the age of the speaker, thus suggesting that participants in this age range were able to enhance their speech using these measures regardless of the maturity of their speech production system.

Indeed, numerous studies have found that adults typically increase their F0 range and speech intensity when asked to speak clearly (Smiljanić and Bradlow, 2009). These two strategies also often feature in instructions for communication partners of persons with HI, due to their being able to improve the salience of speech cues (Cooke et al., 2014a). Decreasing the number of words used per phrase, indicative of the use of a message simplification strategy in HI-directed speech, is also similar to that found in foreigner-directed speech by adult speakers (e.g., Long, 1983), and in infant-directed speech even by preschool-aged children (Dunn and Kendrick, 1982; Sachs and Devin, 1976; Shatz and Gelman, 1973). Increasing the size of the vowel space area, and therefore increasing spectral distance between vowels, when speaking clearly compared to casually is also a typical finding in adult clear speech studies (Bradlow, 2002; Ferguson and Kewley-Port, 2007). The children in the current study thus seem to able to make more extreme articulatory gestures when needed, which likely aids the categorisation of vowels in their conversation partners with HI.

However, the child speakers in this study did not decrease their articulation rate significantly in HI-directed compared to NH-directed speech, although a decrease of between 26% and 48% is a main finding in typical adult clear speech studies (Picheny et al., 1986; Smiljanić and Bradlow, 2005).

Both adults and 9-14-year-old children were also found to reduce their articulation rate in spontaneous speech approximately 21-23% in a low-intelligibility vocoder condition compared to a normal listening channel when completing a different problem-solving task (Hazan et al., 2016). This discrepancy in articulation rate enhancement between the current study and previous studies may be due to task differences, or can at least partly be due to the interactive nature of the current study: perhaps a substantial decrease in articulation rate is not a useful strategy when talking to a peer with HI. Several sources which give instructions to communication partners of persons with HI ask them not to speak too slowly as it may affect the HI person's ability to lip read (Action on Hearing Loss, 2012; NDCS, 2012a). Indeed, there is evidence that a slower speech rate on its own does not increase speech intelligibility, at least to listeners with NH in noise (Cooke et al., 2014b). Alternatively, articulation rate was not decreased due to speakers using fairly short phrases even in the NH-directed speaking condition due to the nature of the task. It is also possible that, instead of decreasing articulation rate to aid their communication partner's speech processing, participants increased their pause frequency in HI-directed compared to NH-directed speech.

The participants did not produce lexically more frequent words when talking to a friend with HI compared to a friend with NH – although, in another linguistic barrier condition, foreigner-directed speech, adults are found to use this strategy (Long, 1983). Similarly, speakers were not found to produce differences between conditions in terms of the diversity of their vocabulary (VOCD), despite a revision strategy (i.e., higher lexical diversity) being potentially helpful when talking to a friend with HI. For a speaker to be able to use higher frequency vocabulary items or a more varied lexicon in their speech likely requires a wide vocabulary, and an awareness of the lexical items which may be easier for the interlocutor to comprehend – aspects of linguistic competence which require extensive linguistic experience, and which therefore are unlikely to be adult-like in late childhood. Alternatively, it may be that the restrictive nature of the referential communication task used in this study does not allow speakers to demonstrate lexical frequency effects, as the vocabulary used within the interaction mainly consists of the objects used in the tasks, many of which are fairly frequent words and may not have higher-frequency alternatives.

In Hazan et al. (2016), older children increased their F0 range when talking to a friend through a vocoder, although a vocoder does not transmit F0 information; thus it has been suggested that even older children are not necessarily very sensitive to the specific needs of a listener. However, in this study, children appear to be able to make adaptations to the needs of their listeners with HI. This implies that at least these children may be using similar monitoring of their listener as adults do – suggesting that, despite their potential egocentrism (Blakemore and Choudhury, 2006), they already produce speech and language from a listener-oriented perspective, in which speech effort is increased when communication is difficult (as hypothesised by Lindblom, 1990, in the H&H model), and speech is adapted to suit less competent interlocutors, as interactionist and usage-based theories posit (Vygotsky, 1978; Gallaway and Woll, 1994; Tomasello, 2003).

However, it is unclear how these children were able to learn to take their listener's needs into account. The H&H theory and the interactionist and usage-based theories have not, to our knowledge, proposed developmental theories of their models which explain children's abilities in this domain. Additionally, as the speech adaptation measures in the current study were taken from the participants' entire interaction, we do not know whether the adaptations were made dynamically according to listener feedback as proposed by H&H, or based on past experience. As each speaker interacted with only one interlocutor with HI, and the interlocutors were friends and had experience in interacting with each other, the speakers may have learned to use adaptation strategies only with their particular friend, but may not be able to generalise their strategies to unfamiliar HI peers (c.f., Lederberg et al., 1986). On the other hand, it is possible that the experience of frequently communicating with friends with HI may lead these participants to be able to adapt to other persons with HI to a greater extent than speakers with no such experience.

A referential communication task was used in this study to enable the elicitation of spontaneous, but controlled, interactive speech. Although children are likely to encounter such tasks frequently in problem-solving contexts at school and at home, such tasks are nonetheless somewhat artificial. In particular, referential communication tasks have a high understanding criterion (e.g., Skelt, 2011) – namely, the interlocutors require a high degree of mutual understanding for the purpose of completing the task successfully, and therefore, if miscommunications occur, they are likely to use frequent clarification requests and repair misunderstandings in their interaction. However, in everyday social situations, the understanding criterion may be much lower – interlocutors may allow some miscommunications to pass (Clark and Wilkes-Gibbs, 1986). Thus, the high understanding criterion in the task may have influenced the strategies speakers used – speakers may have been more motivated to make greater effort to maintain successful communication than in an everyday situation.

Also, although the study attempted to analyse the main communication strategies used by the speakers in their peer-to-peer interaction, many of the participants may have been using other communication strategies with their friend which were not analysed here. For example, when asked by the researcher about the strategies they use during miscommunications with their friends with HI, many participants mentioned using visual strategies, such as hand gestures or sign language (c.f. Appendix G in Granlund, 2015). As videos of the interactions were collected in addition to the audio data used here, future work can analyse the extent to which these additional strategies were used.

Similarly, as the Grid task elicited many miscommunications relating to minimal pair keywords, it may be valuable in future work to analyse the acoustic-phonetic properties of the participants' speech occurring before and after miscommunication in conjunction with repair strategies used. This would enable us to explore whether children modify their speech dynamically during the interaction as required by the interlocutor, as predicted by certain models of communication (e.g., Lindblom, 1990). It would also allow an examination to be done on whether acoustic-phonetic strategies interact with higher-level linguistic and pragmatic strategies used by the speakers.

This study analysed the adaptation strategies used in peer-to-peer interaction between NH and HI children, as effective communication with peers is likely to be very important to school-aged HI children's social and emotional development (Antia et al., 2011; Batten et al., 2014). However, it is unclear to what extent the results from this study are affected by the children's continued speech and language development, as no similar data has been collected from adult NH-HI interactions – a further potential focus of future work.

In summary, this study was conducted to discover the speech communication strategies being used by children in inclusive education, to ensure that children with HI obtain maximum benefit in this setting. The main conclusion of the study is that despite older children's continued speech and language development, they are able to make some global acoustic and linguistic adaptations to their speech when talking with a peer with HI compared to a peer with NH, even without having been given specific instruction to do so – a potentially useful finding for educators doing deaf-awareness training for children with NH. Furthermore, the findings of the current study highlight the need for the adaptation and clear speech literature to consider more holistic and realistic approaches to speaker-listener adaptation, which would integrate accounts of both the development of these interactive skills as well as the multiple levels of adaptation available to speakers when modifying their speech to their listener.

5. Acknowledgments

We thank all the participating schools and children. This project was funded by a PhD studentship from the UK Economic and Social Research Council (ESRC) linked to grant number RES-062-23-3106.

6. References

Action on Hearing Loss, 2012. Miscommunication? Tips when speaking to someone with a hearing loss. London, UK, leaflet.

Adank, P., Smits, R., van Hout, R., 2004. A comparison of vowel normalization procedures for language variation research. The Journal of the Acoustical Society of America 116 (5), 3099–3107.

Amazi, D. K., Garber, S. R., 1982. The Lombard sign as a function of age and task. Journal of Speech and Hearing Research 25, 581–585.

Antia, S. D., Reed, S., Shaw, L., Jan. 2011. Risk and Resilience for Social Competence: Deaf Students in General Education Classrooms. In: Resilience in Deaf Children: Adaptation Through Emerging Adulthood. Springer New York, New York, NY, pp. 139–167.

Arnold, D., Tremblay, A., 1979. Interaction of deaf and hearing preschool children. Journal of Communication Disorders 12 (3), 245–251.

Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. Journal of Statistical Software 67 (1), 1–48.

Batten, G., Oakes, P. M., Alexander, T., Jun. 2014. Factors Associated With Social Interactions Between Deaf Children and Their Hearing Peers: A Systematic Literature Review. Journal of Deaf Studies and Deaf Education 19 (3), 285–302.

Bernhardt, B., Gick, B., Bacsfalvi, P., Ashdown, J., Jan. 2003. Speech habilitation of hard of hearing adolescents using electropalatography and ultrasound as evaluated by trained listeners. Clinical Linguistics & Phonetics 17 (3), 199–216.

Blakemore, S.-J., Choudhury, S., 2006. Development of the adolescent brain: implications for executive function and social cognition. Journal of Child Psychiatry and Psychology 47, 296–312.

Blamey, P. J., Sarant, J. Z., Paatsch, L. E., Barry, J. G., Bow, C. P., Wales, R. J., Wright, M., Psarros, C., Rattigan, K., Tooher, R., 2001. Relationships among speech perception, production, language, hearing loss, and age in children with impaired hearing. Journal of Speech Language and Hearing Research 44 (2), 264–285.

Boersma, P., Weenink, D., 2015. Praat: doing phonetics by computer.
URL http://www.praat.org

Borg, E., Edquist, G., Reinholdson, A.-C., Risberg, A., McAllister, B., Jul. 2007. Speech and language development in a population of Swedish hearing-impaired pre-school children, a cross-sectional study. International Journal of Pediatric Otorhinolaryngology 71 (7), 1061–1077.

Bradlow, A. R., 2002. Confluent talker-and listener-oriented forces in clear speech production. Laboratory Phonology 7, 241–273.

Bradlow, A. R., Torretta, G. M., Pisoni, D. B., 1996. Intelligibility of normal speech I: global and fine-grained acousticphonetic talker characteristics. Speech Communication 20, 255–272.

Brinton, B., Fujiki, M., Loeb, D. F., Winkler, E., 1986. Development of conversational repair strategies in response to requests for clarification. Journal of Speech Language and Hearing Research 29 (1), 75–81.

Burnham, D., Kitamura, C., Vollmer-Conna, U., 2002. What's new, pussycat? On talking to babies and animals. Science 296 (5572), 1435–1435.

Caissie, R., Tranquilla, M., Jun. 2010. Enhancing Conversational Fluency: Training Conversation Partners in the Use of Clear Speech and Other Strategies. Seminars in Hearing 31 (02), 95–103.

Charles-Luce, J., 1997. Cognitive factors involved in preserving a phonemic contrast. Language and Speech 40 (3), 229–248.

Cheng, H. Y., Murdoch, B. E., Goozée, J. V., Scott, D., 2007. Electropalatographic assessment of tongue-to-palate contact patterns and variability in children, adolescents, and adults. Journal of Speech Language and Hearing Research 50 (2), 375–392.

Clark, H. H., Wilkes-Gibbs, D., 1986. Referring as a Collaborative Process. Cognition 22, 1–39.

Cooke, M., King, S., Garnier, M., Aubanel, V., Mar. 2014a. The listening talker: A review of human and algorithmic context-induced modifications of speech. Computer Speech & Language 28 (2), 543–571.

Cooke, M., Lu, Y., Oct. 2010. Spectral and temporal changes to speech produced in the presence of energetic and informational maskers. The Journal of the Acoustical Society of America 128 (4), 2059–2069.

Cooke, M., Mayo, C., Villegas, J., Feb. 2014b. The contribution of durational and spectral changes to the Lombard speech intelligibility benefit. The Journal of the Acoustical Society of America 135 (2), 874–883.

Costa, A., Pickering, M. J., Sorace, A., Jun. 2008. Alignment in second language dialogue. Language and Cognitive Processes 23 (4), 528–556.

CRIDE, 2015. CRIDE report on 2015 survey on educational provision for deaf children. London, UK, http://www.batod.org.uk/content/resources/survey/CRIDEE2015.pdf.

Cristia, A., 2010. Phonetic enhancement of sibilants in infant-directed speech. Journal of the Acoustical Society of America 128 (1), 424–434.

Dorman, M. F., Loizou, P. C., Spahr, A. J., Maloff, E., 2002. Factors that allow a high level of speech understanding by patients fit with cochlear implants. American Journal of Audiology 11 (2), 119–123.

Doyle, M., Dye, L., 2002. Classroom tips for working with deaf and hard of hearing students: A guide for professionals, teachers and parents.
URL http://www.handsandvoices.org/pdf/mainst_cal.pdf

Dunn, J., Kendrick, C., Sep. 1982. The speech of two- and three-year-olds to infant siblings: 'baby talk' and the context of communication. Journal of Child Language 9 (03), 579–595.

Ferguson, S. H., Kewley-Port, D., 2007. Talker differences in clear and conversational speech: Acoustic characteristics of vowels. Journal of Speech Language and Hearing Research 50 (5), 1241–1255.

Gallaway, C., Woll, B., Jan. 1994. Interaction and childhood deafness. In: Gallaway, C., Richards, B. J. (Eds.), Input and interaction in language acquisition. Cambridge University Press, Cambridge, pp. 197–218.

Garnier, M., Henrich, N., Dubois, D., 2010. Influence of sound immersion and communicative interaction on the Lombard effect. Journal of Speech Language and Hearing Research 53 (3), 588–608.

Geers, A. E., Sedey, A. L., 2011. Language and verbal reasoning skills in adolescents with 10 or more years of cochlear implant experience. Ear and Hearing 32 (1 Suppl), 39S–48S.

Granlund, S. C., 2015. Speech communication strategies in older children: acoustic-phonetic and linguistic adaptations to a hearing-impaired peer. Ph.D. thesis, University College London,

(UCL).
URL http://discovery.ucl.ac.uk/1470858/

Hazan, V., Baker, R., 2011. Acoustic-phonetic characteristics of speech produced with communicative intent to counter adverse listening conditions. The Journal of the Acoustical Society of America 130 (4), 2139–2152.

Hazan, V., Markham, D., 2004. Acoustic-phonetic correlates of talker intelligibility for adults and children. The Journal of the Acoustical Society of America 116 (5), 3108–3118.

Hazan, V., Tuomainen, O., Pettinato, M., 2016. Suprasegmental characteristics of spontaneous speech produced in good and challenging communicative conditions by talkers aged 9-14 years. Journal of Speech, Language and Hearing Research 59, S1596–S1607.

Heldner, M., Edlund, J., Oct. 2010. Pauses, gaps and overlaps in conversations. Journal of Phonetics 38 (4), 555–568.

Hollien, H., Green, R., Massey, K., 1994. Longitudinal research on adolescent voice change in males. Journal of the Acoustical Society of America 96, 2646–2654.

HTK Team, 2012. The Hidden Markov Model Toolkit.
URL http://htk.eng.cam.ac.uk/

Imaizumi, S., Hayashi, A., Deguchi, T., 1993. Planning in Speech Production: Listener Adaptive Characteristics. The Japan Journal of Logopedics and Phoniatrics 34 (4), 394–401.

Imaizumi, S., Hayashi, A., Deguchi, T., 1995. Listener adaptive characteristics of vowel devoicing in Japanese dialogue. The Journal of the Acoustical Society of America 98 (2), 768–778.

Kempe, V., Oct. 2009. Personality and Individual Differences. Personality and Individual Differences 47 (6), 610–615.

Koenig, L. L., Lucero, J. C., Perlman, E., 2008. Speech production variability in fricatives of children and adults: Results of functional data analysis. The Journal of the Acoustical Society of America 124 (5), 3158–3170.

Kosky, C., Boothroyd, A., Feb. 2003. Perception and Production of Sibilants by Children with Hearing Loss: A training study. The Volta Review 103 (2), 71–98.

Krause, J. C., Braida, L. D., 2004. Acoustic properties of naturally produced clear speech at normal speaking rates. The Journal of the Acoustical Society of America 115 (1), 362–378.

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., Sundberg, U., Lacerda, F., 1997. Cross-language analysis of phonetic units in language addressed to infants. Science 277 (5326), 684–686.

Kuo, Y.-C., Rosen, S., Faulkner, A., 2008. Acoustic cues to tonal contrasts in Mandarin: Implications for cochlear implants. The Journal of the Acoustical Society of America 123 (5), 2815–2824.

Lederberg, A. R., Ryan, H. B., Robbins, B. L., 1986. Peer interaction in young deaf children: The effect of partner hearing status and familiarity. Developmental Psychology 22 (5), 691–700.

Lee, S., Potamianos, A., Narayanan, S., Mar. 1999. Acoustics of children's speech: Developmental changes of temporal and spectral parameters. The Journal of the Acoustical Society of America 105 (3), 1455–1468.

Lind, C., Campbell, K., Davey, A., Rodgers, L., Seipolt, C., Akins, C., Jun. 2010. Some Prosodic Characteristics of Repeated Talk following Conversation Repair Requests by Adults with Hearing Impairment. Seminars in Hearing 31 (02), 127–139.

Lindblom, B., 1990. Explaining phonetic variation: A sketch of the H&H theory. In: Hardcastle, W. J., Marchal, A. (Eds.), Speech production and speech modelling. Kluwer Academic Publishers, pp. 403–439.

Lombard, E., 1911. Le signe de l'elevation de la voix. Annals maladiers oreille, Larynx, Nez, Pharynx 37, 101–119.

Long, M. H., 1983. Linguistic and conversational adjustments to non-native speakers. Studies in second language acquisition 5 (2), 177–193.

Lu, X., Dec. 2012. The Relationship of Lexical Richness to the Quality of ESL Learners' Oral Narratives. The Modern Language Journal 96 (2), 190–208.

Marschark, M., Hauser, P. C., 2012. How deaf children learn: what parents and teachers need to know. Oxford University Press, Oxford.

McKee, G., Malvern, D., Richards, B., 2000. Measuring vocabulary diversity using dedicated software. Literary and linguistic computing 15 (3), 323–338.

Moore, B. C., 2007. Cochlear hearing loss: physiological, psychological, and technical issues. John Wiley and Sons, Chichester.

NDCS, 2012a. Look, Smile, Chat. London, UK, http://youngpeople.ndcsbuzz.org.uk/looksmilechat/.

NDCS, 2012b. Tips for communicating with deaf children. London, UK, leaflet.

Nippold, M. A., Hesketh, L. J., Duthie, J. K., Mansfield, T. C., 2005. Conversational versus expository discourse: A study of syntactic development in children, adolescents, and adults. Journal of Speech Language and Hearing Research 48 (5), 1048–1064.

Nittrouer, S., 1993. The emergence of mature gestural patterns is not uniform: evidence from an acoustic study. Journal of Speech Language and Hearing Research 36, 959–972.

Nittrouer, S., Estee, S., Lowenstein, J. H., Smith, J., 2005. The emergence of mature gestural patterns in the production of voiceless and voiced word-final stops. The Journal of the Acoustical Society of America 117, 351–364.

Peterson, C. L., Danner, F. W., Flavell, J. H., 1972. Developmental changes in children's response to three indications of communicative failure. Child Development, 1463–1468.

Pettinato, M., Tuomainen, O., Granlund, S., Hazan, V., Jan. 2016. Vowel space area in later childhood and adolescence: Effects of age, sex and ease of communication. Journal of Phonetics 54, 1–14.

Picheny, M. A., Durlach, N. I., Braida, L. D., Dec. 1986. Speaking Clearly for the Hard of Hearing II. Journal of Speech Language and Hearing Research 29 (4), 434–446.

R Core Team, 2014. R: A language and environment for statistical computing.
URL `http://www.R-project.org/`

Redford, M. A., Gildersleeve-Neumann, C. E., 2009. The development of distinct speaking styles in preschool children. Journal of Speech Language and Hearing Research 52 (6), 1434–1448.

Richards, B., Jun. 1987. Type/Token Ratios: what do they really tell us? Journal of Child Language 14 (02), 201–209.

Rinker, T. W., 2013. qdap: Quantitative Discourse Analysis Package. University at Buffalo/SUNY, Buffalo, New York, version 2.2.0.
URL `http://github.com/trinker/qdap`

Romeo, R., Hazan, V., Pettinato, M., 2013. Developmental and gender-related trends of intratalker variability in consonant production. The Journal of the Acoustical Society of America 134 (5), 3781–3792.

Ruffin, C. V., Kronenberger, W. G., Colson, B. G., Henning, S. C., Pisoni, D. B., 2013. Long-Term Speech and Language Outcomes in Prelingually Deaf Children, Adolescents and Young Adults Who Received Cochlear Implants in Childhood. Audiology and Neurotology 18 (5), 289–296.

Sachs, J., Devin, J., Feb. 1976. Young children's use of age-appropriate speech styles in social interaction and role-playing. Journal of Child Language 3 (01), 81–98.

Scarborough, R., Brenier, J., Zhao, Y., Hall-Lew, L., Dmitrieva, O., 2007. An acoustic study of real and imagined foreigner-directed speech. Proceedings of the International Conference of the Phonetic Sciences (ICPhS) XVI, 2165–2168.

Seewald, R. C., Brackett, D., 1984. Spoken language modifications as a function of the age and hearing ability of the listener. The Volta Review 86 (1), 20–35.

Shannon, R., 2002. The relative importance of amplitude, temporal and spectral cues for cochlear implant processor design. American Journal of Audiology 11, 124–127.

Shatz, M., Gelman, R., 1973. The development of communication skills: Modifications in the speech of young children as a function of listener. Monographs of the society for research in child development, 1–38.

Skelt, L., 2011. Dealing with misunderstandings: The sensitivity of repair in hearing-impaired conversation (chapter 7). In: Egbert, M., Deppermann, A. (Eds.), Hearing Aids Communication: Integrating Social Interaction, Audiology and User Centered Design to Improve Communication with Hearing Loss and Hearing Technologies. Verlag für Gesprächsforschung, pp. 56–66.

Smiljanić, R., Bradlow, A. R., 2005. Production and perception of clear speech in Croatian and English. The Journal of the Acoustical Society of America 118 (3), 1677–1688.

Smiljanić, R., Bradlow, A. R., Jan. 2009. Speaking and Hearing Clearly: Talker and Listener Factors in Speaking Style Changes. Language and Linguistics Compass 3 (1), 236–264.

Smith, A., Goffman, L., 1998. Stability and patterning of speech movement sequences in children and adults. Journal of Speech Language and Hearing Research 41, 18–30.

Spencer, P., 2004. Individual differences in language performance after cochlear implantation at one to three years of age: child, family, and linguistic factors. Journal of Deaf studies and deaf education 9, 395–412.

Spencer, P., Koester, L. S., Meadow-Orlans, K., 1994. Communicative Interactions of Deaf and Hearing Children in a Day Care Center: An Exploratory Study. American Annals of the Deaf 139 (5), 512–518.

Stelmachowicz, P. G., Pittman, A. L., Hoover, B. M., Lewis, D. E., Moeller, M. P., 2004. The importance of high-frequency audibility in the speech and language development of children with hearing loss. Archives of Otolaryngology–Head & Neck Surgery 130 (5), 556–562.

Summers, W. V., Pisoni, D. B., Bernacki, R. H., Pedlow, R. I., Stokes, M. A., 1988. Effects of noise on speech production: acoustic and perceptual analyses. J. Acoust. Soc. Am. 84 (3), 917–928.

Syrett, K., Kawahara, S., Nov. 2014. Production and perception of listener-oriented clear speech in child language. Journal of Child Language 41 (06), 1373–1389.

Tomasello, M., 2003. Constructing a language: a usage-based theory of language acquisition. Harvard University Press, Cambridge, MA.

Tsui, I. Y. H., Ciocca, V., 2000. Perception of aspiration and place of articulation of Cantonese initial stops by normal and sensorineural hearing-impaired listeners. International Journal of Language & Communication Disorders 35 (4), 507–525.

Uchanski, R. M., Geers, A. E., Feb. 2003. Acoustic Characteristics of the Speech of Young Cochlear Implant Users: A Comparison with Normal-Hearing Age-Mates. Ear and Hearing 24 (Supplement), 90S–105S.

Uziel, A. S., Sillon, M., Vieu, A., Artieres, F., Piron, J.-P., Daures, J.-P., Mondain, M., 2007. Ten-year follow-up of a consecutive series of children with multichannel cochlear implants. Otology & neurotology 28 (5), 615–628.

van Heuven, W. J. B., Mandera, P., Keuleers, E., Brysbaert, M., 2014. SUBTLEX-UK: A new and improved word frequency database for British English. Quarterly Journal of Experimental Psychology 67, 1176–1190.

Vandell, D. L., Anderson, L. D., Ehrhardt, G., Wilson, K. S., 1982. Integrating hearing and deaf preschoolers: An attempt to enhance hearing children's interactions with deaf peers. Child Development, 1354–1363.

Vandell, D. L., George, L. B., 1981. Social interaction in hearing and deaf preschoolers: Successes and failures in initiations. Child Development, 627–635.

Vygotsky, L. S., 1978. Mind in Society: The development of higher psychological processes. Harvard University Press, Cambridge, MA.

Weppelman, T. L., Bostow, A., Schiffer, R., Elbert-Perez, E., Newman, R. S., 2003. Children's use of the prosodic characteristics of infant-directed speech. Language & Communication 23 (1), 63–80.

Xanthos, A., 2011. Lingua::Diversity - measuring the diversity of text units.
URL `http://search.cpan.org/~axanthos/Lingua-Diversity-0.06/lib/Lingua/Diversity.pm`

Supplemental materials

Fixed effects	**Estimate**	**SE**	**t**
(intercept)	80.92	18.88	4.29
pair type (NH-NH)	-19.50	5.52	-3.54
pair age	-3.17	1.57	-2.02

Table 3: Fixed effects for the best-fit model for task transaction time. Number of observations: 82; number of different pairs: 25

Fixed effects	**Estimate**	**SE**	**t**
(intercept)	5.03	0.87	5.79
directed (NHD)	-0.52	0.15	-3.47
age	-0.15	0.07	-2.07

Table 4: Fixed effects for the best-fit model for F0 range. Number of observations: 107; number of participants: 16

Fixed effects	**Estimate**	**SE**	**t**
(intercept)	72.19	2.66	27.10
directed (NHD)	-1.80	0.27	-6.62
age	-0.69	0.22	-3.12

Table 5: Fixed effects for the best-fit model for ME 1-3 kHz. Number of observations: 111; number of participants: 16

Fixed effects	**Estimate**	**SE**	**t**
(intercept)	4.12	0.20	20.33
directed (NHD)	0.42	0.18	2.34

Table 6: Fixed effects for the best-fit model for words per phrase. Number of observations: 110; number of participants: 16

Fixed effects	**Estimate**	**SE**	**t**
(intercept)	5.53	0.11	49.92
directed (NHD)	-0.040	0.024	-1.67
age	-0.027	0.009	-2.99

Table 7: Fixed effects for the best-fit model for lexical frequency. Number of observations: 111; number of participants: 16

Fixed effects	**Estimate**	**SE**	**t**
(intercept)	34.02	4.31	7.89
directed (NHD)	-1.92	0.52	-3.71
age	-0.96	0.36	-2.67

Table 8: Fixed effects for the best-fit model for vowel space area. Number of observations: 32; number of participants: 16

Fixed effects	**Estimate**	**SE**	**t**
(intercept)	7.14	0.66	10.81
directed (NHD)	-0.28	0.09	-2.96
age	-0.16	0.06	-2.98

Table 9: Fixed effects for the best-fit model for F1 range. Number of observations: 32; number of participants: 16

Fixed effects	**Estimate**	**SE**	**t**
(intercept)	9.17	0.14	65.47
directed (NHD)	-0.29	0.14	-2.13

Table 10: Fixed effects for the best-fit model for F2 range. Number of observations: 32; number of participants: 16