# Speech communication strategies in older children: acoustic-phonetic and linguistic adaptations to a hearing-impaired peer

Sonia Catherine Granlund

Department of Speech Hearing and Phonetic Sciences

University College London (UCL)

A thesis submitted for the degree of

*Doctor of Philosophy*

2015

I, Sonia Granlund, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Sonia Granlund

# Abstract

This thesis examines the communication strategies used by both hearing (NH) and hearing-impaired (HI) children when interacting with a peer with hearing loss, focusing on the acoustic-phonetic and linguistic properties of their speech. To elicit frequent repetitions of segmental contrasts in HI children's spontaneous speech in interaction, a new task was developed using minimal pair keywords in a communicative game context. In addition, another referential communication task, the 'spot the difference' Diapix task (Van Engen *et al.*, 2010), was used. Eighteen NH and eighteen HI children between 9 and 15 years of age performed the two tasks in pairs, once with a friend with normal hearing (NH-directed speech) and once with a friend with a hearing-impairment (HI-directed speech). Task difficulty increased in interactions involving a HI interlocutor, implying a need for speaker-listener adaptations.

Participants' global acoustic-phonetic (articulation rate, F0 median and range, speech intensity and pausing), segmental (/p/-/b/, /s/-/ʃ/, and /i/-/ɪ/) and linguistic (phrase length, lexical frequency, lexical diversity and speech overlap) adaptations to a HI interlocutor were explored. Although HI speakers were found to differ from NH speakers in many aspects of their speech and language, the two groups used similar, mostly global and linguistic, strategies to adapt to the needs of their HI friend – and the HI children's ability to adapt did not seem to be related to their own speech level. Only a subset of speakers was found to increase the discriminability of phonetic contrasts in speech, perhaps partly due to speakers using segmental and linguistic strategies as alternative methods in adaptation. Both NH and HI speakers appeared to adjust the extent of adaptations made to the specific needs of their HI interlocutor, therefore implying surprising sensitivity to listener needs. Implications to models of speech communication are discussed.

# Acknowledgements

My heartfelt gratitude goes to my primary supervisor, Valerie Hazan, whose support and guidance during this PhD was unparalleled. Her expertise, encouragement and insightful comments on my work always pushed me forward. She has been a great inspiration to me throughout my PhD.

Many thanks also to my secondary supervisor, Merle Mahon, who advised me on deafness, and supported and guided me through the many practical aspects of testing children with hearing loss. In particular, I would like to thank her for taking on project students to analyse the video data that I collected – Merle's thoughtful discussion on the data during project meetings provided me with an invaluable complementary perspective on the analysis of conversation.

I was also very lucky in that my PhD was linked to a 'sister' project, which Michèle Pettinato and Outi Tuomainen worked on as postdoctoral fellows. They were always ready to discuss ideas, give helpful feedback, or just to go for a tea and a chat. I am very grateful for their friendship and support. Special thanks to Outi for her expert advice on statistics.

Thank you also to all those who helped me with different aspects of this research. Hannah Pimperton, Lucy Carroll, Katrin Skoruppa and Ian Noon for their help and advice on participant recruitment, Steve Nevard and Dave Cushing for all their technical help with getting my testing equipment ready, Stuart Rosen for his assistance with setting up the WiNiCS task, Andy Faulkner for providing the materials for the VCV test, and Paul Iverson and Mark Huckvale for their indispensable help with the aligner. Anita Auszmann and Erica Gold were a vital help in alignment checking and segmentation. Thank you to Joy Chu, Ellie Dunn, Rukhaiya Ebrahim, Hannah Harris and Maelíosa Ní Almhain for their insightful analyses and discussion on the video data.

Thank you also to all the PhD students, past and present, who made Room 326 such a warm and enjoyable environment to work in. In particular, I would like to thank Yasna Pereira, Tim Schoof and Wafa'a Alshangiti for their friendship and support throughout these years. A big thank you also to Joaquín Atria and Mauricio Figueroa for their advice and support on all things Praat. My desk neighbour Mark Wibrow was also a great source of technical knowledge and help. Thank you to my spontaneous speech buddies, Jieun Song and Gise Tomé Lourido, for their continuous encouragement and enthusiasm, and Albert Lee and Yasu Shinohara, who were great colleagues as fellow tutors for the courses we taught on together.

My family has always been my greatest source of support and encouragement. In particular, my parents first sparked my interest in language, and never stopped believing in me. My husband, Yan, provided me with second-to-none support, understanding and, at times, much-needed perspective and insight, during this PhD.

I would like to thank all the schools that made this study possible. I am indebted to the teachers who took considerable time out of their busy schedules to find me suitable participants, and to discuss, plan and organise testing with me. A very special thank you to all the children who participated in this study with such good humour, patience and enthusiasm.

Finally, I am very grateful to the UK Economic and Social Research Council (ESRC) whose funding, which was linked to grant number RES-062-23-3106, made it possible for me to pursue this PhD.

# Contents

# List of Figures

# List of Tables

*"What do you do if your friend can't understand you?"*

*"[I would] explain it in more child language like words that they would use instead of bigger words"*

Girl, normally-hearing, 10 years

*"I'd do it really slowly like slowly but loudly"*

Boy, hearing-impaired, 12 years

*"You repeat it for them so that they understand it and say it a lot clearer this time"*

Girl, normally-hearing, 12 years

*"Sometimes we need to sound it out for each other so we can make it easier"*
*"Yeah I can help you"*
*"Yeah and I can help you sometimes can't I"*

Boys, normally-hearing and hearing-impaired, 9 years

# Chapter 1

# Introduction

The fundamental role of speech communication in human interaction is that it enables interlocutors to achieve mutual understanding. However, this is not always an easy feat – not only does communication often occur in less than ideal listening environments, such as noisy schools or reverberant classrooms, but people may also differ from each other in their background knowledge, in their native language, or in their ability to formulate and 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 (Foster, 1990).

The development of the ability to adapt to a listener's needs is especially important for children who are in frequent contact with hearing-impaired (HI) peers, as these children may have difficulty in comprehending the speech of others due to delays in receptive speech and language. Although increasing numbers of HI children in the UK attend mainstream schools (CRIDE, 2014), many normally-hearing (NH) children report not knowing how to communicate with their HI peers (NDCS, 2012a). A recent campaign by the UK National Deaf Children's Society (NDCS) called 'Look, Smile, Chat' (NDCS, 2012a) encouraged NH children to adopt various strategies for talking with HI peers – but very little research has explored the speech communication strategies used by NH and HI peers when interacting with each other. On the other hand, with recent advances in amplification devices and early intervention of hearing loss, increased numbers of HI children rely on an oral-only communication mode (CRIDE, 2014). Therefore, despite many HI children having delays in their speech production skills, they themselves will also need to acquire competence in adapting their speech communication to the needs of their listener, for example to the needs of another HI child. Most research on HI children assesses their performance in speech production or speech perception, but few explore their competence in real peer interaction.

This thesis examines whether 9- to 15-year-old HI children, some wearing hearing aids and others cochlear implants, and their NH peers change their speech and language according to the hearing status of their interlocutor, with the study focusing on the acoustic-phonetic and linguistic aspects of their speech. In the study, each child, whether HI or NH, took part in two 'communication' sessions: one with a NH friend (to elicit NH-directed speech) and one with a HI friend (to elicit HI-directed speech). Each communication session involved participants engaging in collaborative problem-solving tasks to elicit spontaneous speech. The HI-directed speech is compared to the NH-directed speech produced by each child to investigate the differences in their speech and language in the two conditions. We also relate communication strategies to the effectiveness of the interaction, and investigate the factors affecting the amount of adaptation shown by the interlocutors. Although most studies on speech interaction have examined either young infants or adults, this study assesses peer interaction in older children, as peers become increasingly important for a child's social and emotional development in late childhood and early adolescence, heightening the importance of robust interaction strategies being used in communication (e.g., Antia, Reed, and Shaw, 2011; Batten, Oakes, and Alexander, 2014).

This chapter aims to review literature on NH and HI children's abilities in adapting to the needs of a HI interlocutor. First, section 1.1 gives an overview of the current trends in intervention, amplification and education for HI children in the UK. Then, section 1.2 explores two main reasons behind potential difficulties in HI and NH children's peer communication – speech perception (1.2.1) and speech production (1.2.2) development. Section 1.3 relates HI children's speech perception difficulties to the potential strategies that may be useful in enhancing speech to HI children (1.3.1). It also investigates the skills that may be needed for a speaker to be able to make adaptations to listener needs, and reviews the speech adaptations that previous studies have found adults (1.3.2), NH children

(1.3.3) and HI children (1.3.4) to be able to make. Finally, section 1.4 summarises the findings of this chapter, and gives an overview of the structure and research questions of this thesis.

## 1.1 Children and hearing loss in the UK

In 2014, there were approximately 40,000 hearing-impaired[1] children in England (CRIDE, 2014). Estimates show that a hearing-impairment is present at birth in between 1 and 2 babies out of every 1000 born in the UK (Bamford, Fortnum, Bristow, Smith, Davies, Taylor, Watkin, Fonseca, Davis, and Hind, 2007; Bamford, Uus, and Davis, 2005) - a number which approximately doubles by the age of 9 due to acquired hearing loss (Fortnum, Summerfield, Marshall, Davis, and Bamford, 2001). Genetic factors likely contribute to more than half of all incidences of childhood hearing-impairment (Nadol and Merchant, 2001), with the remaining cases often due to craniofacial abnormalities, premature birth, complications or viral infections during pregnancy, or due to trauma or illness, such as meningitis, in childhood (Davis, Davis, and Mencher, 2009; Fortnum and Davis, 1997). Between 20% and 40% of HI children have additional or complex needs (CRIDE, 2013; Fortnum and Davis, 1997). Children from certain ethnic backgrounds also have a high risk of hearing-impairment (Bajaj, Sirimanna, Albert, Qadir, Jenkins, Cortina-Borja, and Bitner-Glindzicz, 2009).

All of the participants in the current study, and most children with a permanent hearing-impairment, have sensorineural hearing loss (SNHL) (Fortnum and Davis, 1997), i.e., impairments of the cochlea or the cochlear nerve. It is usually caused by outer hair cell damage in the cochlea (Moore, 2007; Zeng and Djalilian, 2010). This leads to reduced frequency resolution and selectivity, and the loss of sensitivity to sounds, i.e., being unable to hear quiet sounds. The dynamic range of sounds from audible to painfully loud is much more restricted in those with SNHL than those with normal hearing, leading to 'loudness recruitment'– moderately loud sounds are heard as quiet sounds, while sounds at high intensities still sound very loud. Because of the nature in which the outer hair cells

[1]The term 'hearing-impaired' is used throughout this thesis to denote people who have a permanent bilateral hearing loss of at least 40 dB or more in the better ear.

process sounds in the cochlea, the higher frequencies are usually more affected by SNHL (Halliday and Moore, 2010). However, temporal processing in the cochlea is usually unaffected (Jerger, 2007; Zeng and Djalilian, 2010). Alternatively, or additionally, the hearing-impairment can be conductive, in which the middle or outer ears do not function adequately (Davis *et al.*, 2009), and it can be progressive, where hearing thresholds deteriorate over several years (Davis *et al.*, 2009). In England, approximately 35% of HI children have moderate hearing losses; 11% of cases are severe and 12% are profound[1] (CRIDE, 2013).

During the last 20 years, several major advances in technology in the UK and worldwide have lead to HI children having better and earlier access to spoken language than ever before. One of the most important advances was the introduction of universal newborn hearing screening (UNHS), which started in late 2001, and became standard practice throughout England by 2006 (Action on Hearing Loss, 2011; Bamford *et al.*, 2005). Prior to UNHS, the median age of identification for congenital hearing loss was 22 months. However, there was increasing evidence that the first six months of an infant's life are important for the development of language and communication skills (Davis, Bamford, Wilson, Ramkalawan, Forshaw, and Wright, 1997). Early identification, coupled with early intervention, could therefore significantly increase early access to speech and language, and provide more positive outcomes in HI children. Now, due to UNHS, the average age of identification in the UK may be as low as 2 months (Kennedy, McCann, Campbell, Law, Mullee, Petrou, Watkin, Worsfold, Yuen, and Stevenson, 2006; Young and Tattersall, 2007).

Another important recent advancement has been in hearing aid technology. Hearing aids (HAs) contain a microphone and a sound processor, which sends amplified acoustic information to the ear canal via a loudspeaker; the device is connected to the ear via an ear mould (Moore, 2007). In the mid-1990s, digital signal processing was introduced to HA sound processors, to enable the use of multi-channel processing to amplify sounds selectively at different frequencies, and automatic gain control to compress the intensity of sounds to the restricted

[1] The level of hearing-impairment is classified according to the average hearing threshold of the better ear at 0.5, 1, 2 and 4kHz as follows: mild - 20-39 dB HL; moderate - 40-69 dB HL; severe - 70-94 dB HL; profound - 95 dB+ (Davis *et al.*, 2009)

dynamic range of the impaired cochlea (Zeng and Djalilian, 2010). They provide several advantages over older analogue aids, such as enabling greater precision in adjusting the aid's frequency response, providing processing strategies for background noise reduction as well as directional microphones (Moore, 2007; Taylor and Paisley, 2000). HAs are, however, currently unable to rectify losses in frequency resolution and selectivity in the cochlea (Moore, 2007). They are also mostly unable to amplify higher frequencies (Stelmachowicz, Pittman, Hoover, Lewis, and Moeller, 2004). Because of this, some modern HAs contain frequency compression processing strategies, in which either the high frequency components are transposed to lower frequencies (transposition strategy) or the bandwidth of higher frequencies is compressed to cover a smaller than normal frequency range (non-linear frequency compression strategy), to make as much use of residual hearing in lower frequencies as possible (c.f., Ellis, 2012). Modern digital HAs are typically used to treat people with mild to severe SNHL.

Simultaneously, the establishment of over 15 cochlear implant (CI) programmes in the UK since approximately 1995 (Raine, 2013) has enabled many children with severe-to-profound hearing impairments greater access to spoken language than was possible through analogue or digital hearing aids alone, especially if implantation occurs before age 5 (e.g., Stacey, Fortnum, Barton, and Summerfield, 2006). In a CI, a microphone worn externally detects sound in the environment, and passes the signals to a digital sound processor which separates the sounds into 12 to 22 different channels, and sends the frequency information to an array of electrodes implanted within the cochlea. It therefore bypasses the outer and middle ears and provides direct stimulation of the cochlear nerves through electrical impulses (Moore, 2007; Zeng, 2004). CIs are fairly accurate in transmitting the amplitude and temporal envelope information of speech, but provide reduced spectral resolution and temporal fine structure information compared to a normally functioning cochlea (Dorman, Loizou, Spahr, and Maloff, 2002; Zeng and Djalilian, 2010), as well as potentially introducing shifts in the correspondences between frequency and place in the cochlea (Shannon, 2002).

In the UK, CIs are currently being received by approximately 74% of eligible children with severe and profound losses by age 3, and by 93% of them by age 17 (Raine, 2013). CRIDE (2013) estimates that, in 2013, nearly 3,000 children in England had at least one CI (approximately 8% of the total population of HI children in England). The participants in the current study were born between 1998 and 2004, and therefore they were among the first generation of HI children to grow up with the new technological advancements as part of their everyday lives[1].

Partly as a result of the recent technological advancements leading to better speech and language outcomes, HI children are increasingly educated in mainstream schools with their NH peers. Currently, in England, 76% of HI children attend mainstream schools with no specialist provision, while 8.5% attend mainstream schools with specialist resource provision. Only 3% of HI children attend special schools for HI children (CRIDE, 2013). Most HI children use only spoken English at school, with less than 1 in 10 using spoken English with sign language, and only 2% of HI children using British Sign Language on its own (CRIDE, 2014).

The above figures demonstrate that the vast majority of HI children attend classes with NH peers, and use spoken English as the primary communication mode at school. However, it is unclear how well HI children are integrated into mainstream schools (for a recent review, see Xie, Potmesil, and Peters, 2014). Several studies report on relative social isolation of HI children in mainstream environments (e.g., Keating and Mirus, 2003; Martin and Bat-Chava, 2003), likely partly caused by NH and HI children's difficulties in communicating with each other (Bat-Chava and Deignan, 2001). The next section explores in greater detail some of the causes of communication difficulty faced by both NH and HI children in interaction.

## 1.2 Causes of speech communication difficulty in peer interaction

When communicating with peers using spoken language, intelligibility may be compromised for both an HI child and their interlocutor: HI children may have

[1]Most of the HI children in the study were probably not, however, identified using UNHS; only 8 of the 18 HI participants were diagnosed prior to six months of age.

problems understanding the interlocutor due to their own perception deficits, and the interlocutor may find the speech produced by the HI child to be difficult to understand due to the HI child's speech production deficits. Additionally, even NH child peers themselves may still be refining their speech perception and production skills, which may also add to communication difficulties. This section will review NH and HI children's speech perception and production development to explore the reasons behind possible communication difficulties, focusing on research conducted on prelingually deafened HI children after many of the major advancements in screening and hearing aid technology since 1995.

### 1.2.1 Speech perception development

**Normally-hearing children** To develop adult-like speech perception skills, a child needs to be able to (1) receive and process sensory information about speech sounds, (2) classify speech sounds into different phonetic categories, and (3) recognise words and use linguistic contextual knowledge in comprehension (Aslin and Smith, 1988; Nittrouer and Lowenstein, 2010).

Foetuses that are known to have normal hearing after birth already respond to auditory stimuli by 25-29 weeks of gestation (Birnholz and Benacerraf, 1983), and therefore, if born at full term, they are exposed to approximately 2 months of hearing prenatally (Houston, 2011). Foetuses are mostly exposed to frequencies below 1000Hz in utero (Lecanuet, Gautheron, Locatelli, Schaal, Jacquet, and Busnel, 1998) and, in late pregnancy, are able to discriminate the gender of a voice (Lecanuet, Granier-Deferre, Jacquet, Capponi, and Ledru, 1993), and different languages (Kisilevsky, Hains, Brown, Lee, Cowperthwaite, Stutzman, Swansburg, Lee, Xie, Huang, Ye, Zhang, and Wang, 2009), likely using suprasegmental features in speech.

Newborn normally-hearing infants display a predisposition for learning about speech – they prefer listening to speech over non-speech sounds (Vouloumanos and Werker, 2004), are sensitive to phoneme category boundaries (Eimas, Siqueland, Jusczyk, and Vigorito, 1971), are able to distinguish between most phonological contrasts which occur in the world's languages (e.g., Kuhl, Conboy, Coffey-Corina, Padden, Rivera-Gaxiola, and Nelson, 2008) and, by 4.5 months of age, prefer to look at congruent rather than incongruent audiovisual stimuli (Kuhl and Meltzoff, 1982). After approximately 6 months of age, infants start to display the effects of language exposure – they become less sensitive to phonetic contrasts which do not occur in the ambient language (Werker and Tees, 1984), and by about 7.5 months of age, use language-specific segmentation cues to recognise frequently occurring words in speech (Jusczyk and Aslin, 1995). By 1 year of age, they are able to exploit the stress patterns (Jusczyk, Cutler, and Redanz, 1993), phonotactic cues (Mattys and Jusczyk, 2001), and allophonic information (Jusczyk, Hohne, and Bauman, 1999) of their native language for segmenting speech and for learning new words and grammar through a statistical learning mechanism (Saffran, Aslin, and Newport, 1996).

Although, even in the second year of life, normally-hearing children are already able to process auditory stimuli, classify speech sounds according to their native language and recognise words and phrases, all aspects of the speech perception process mature and develop at least until late childhood. In terms of auditory acuity, normally-hearing children become adult-like by age 4-6 years for intensity discrimination (Jensen and Neff, 1993), and age 9 or 10 years for duration and frequency discrimination (Allen and Wightman, 1992; Elfenbein, Small, and Davis, 1993; Jensen and Neff, 1993). Young children may also be more affected by background noise (Nittrouer and Boothroyd, 1990) and may need more intense acoustic signals for recognising speech (Nabelek and Robinson, 1982) than adults.

With auditory-linguistic experience, NH children also learn to use specific cues in their native language for identifying and discriminating speech sounds. Categorical perception, which is the ability to discriminate only between across-category sounds rather than within-category sounds (Liberman, Harris, Hoffman, and Griffith, 1957), is adult-like by age 6 (Hoonhorst, Medina, Colin, Markessis, Radeau, Deltenre, and Serniclaes, 2011), but the consistency with which children identify speech sounds into phonetic categories (categorical precision) develops even into adolescence (Hazan and Barrett, 2000; Hoonhorst *et al.*, 2011; Pursell, Swanson, Hedrick, and Nabelek, 2002; Simon and Fourcin, 1978). Children may also use different perceptual cue weightings than adults to distinguish certain phonetic contrasts, such as fricatives and vowels (Morrongiello, Robson, and

Best, 1984; Nittrouer and Lowenstein, 2009; Nittrouer and Studdert-Kennedy, 1987), and are less able to recognise phonemes on the basis of limited acoustic cues, therefore demonstrating less flexible perception (Eisenberg, Shannon, Martinez, Wygonski, and Boothroyd, 2000; Hazan and Barrett, 2000). They are also less able to cope with speaker variability (e.g., Jacewicz and Fox, 2014). Children continue to increase their receptive vocabulary during adolescence (Duncan, Rhoades, and Fitzpatrick, 2014), and are less able to use semantic context and word frequency in speech recognition than adults (e.g., Elliott, Clifton, and Servi, 1983; Nittrouer and Boothroyd, 1990). Background noise may also affect the linguistic processing of speech to a greater extent in children than in adults (Elliott, 1979).

**Hearing-impaired children** Hearing-impaired children often demonstrate delays in learning spoken language (Blamey and Sarant, 2011), likely due to HI children's reduced quality and quantity of auditory and linguistic experiences compared to NH children (see Moeller, Tomblin, Yoshinaga-Itano, Connor, and Jerger, 2007c, for a review). Unless identified using UNHS and provided with very early intervention, many children with congenital hearing loss experience auditory deprivation in the first few months or even years of life, as well as in utero (Houston, Pisoni, Kirk, Ying, and Miyamoto, 2003). Researchers postulate that there is a 'sensitive period' for the development of language (c.f., Knudsen, 2004), and therefore the longer the period of auditory deprivation, the more unlikely it seems that the child will be able to develop normal speech and language. Indeed, there is evidence of cortical reorganisation in instances of longer auditory deprivation (Ponton and Eggermont, 2001), which may affect language learning later in life (Houston *et al.*, 2003). The other main reason for the differences observed between NH and HI children's language learning is the continued poorer quality auditory input received by HI children. Although most modern digital HAs and CIs are able to greatly improve hearing thresholds for HI children, HAs are unable to rectify the poor frequency resolution in a cochlea, and CIs are also unable to provide very accurate frequency resolution (see section 1.1). Even children with milder losses are unlikely to receive similar auditory experiences to NH peers, due to difficulties perceiving speech in background noise (Nittrouer, Caldwell-

Tarr, Tarr, Lowenstein, Rice, and Moberly, 2013), inconsistencies in early HA use (Jamieson, 2010), and the limited amplification of higher frequencies in HAs (Stelmachowicz *et al.*, 2004). All of the above factors may affect not only lower-level sensory processing but also higher levels of linguistic organisation (Jerger, 2007).

Much of the recent literature on speech perception concentrates on children with CIs, with less research done on children with moderate to severe losses wearing hearing aids (see Eisenberg, 2007; Jerger, 2007, for reviews). Research seems to show, however, that those with milder losses perform better than those with more severe losses (Eisenberg, 2007; Hennies, Penke, Rothweiler, Wimmer, and Hess, 2012). Already in infancy, children with moderate losses prefer congruent rather than incongruent AV stimuli, similarly to NHs, but those with more severe losses do not (Bergeson, Houston, and Miyamoto, 2010). Similarly, infants with moderate to profound losses wearing hearing aids who have received early intervention already lag behind NH peers in receptive language skills at 12 and 16 months of age, while those with mild hearing loss do not (Vohr, Jodoin-Krauzyk, Tucker, Johnson, Topol, and Ahlgren, 2008). These studies are indicative of language delays for children with moderate and severe hearing losses even very early in life. Additionally, early identification and/or early age at amplification (Sininger, Grimes, and Christensen, 2010; Yoshinaga-Itano, Sedey, Coulter, and Mehl, 1998) and early intervention (Moeller, 2000; Vohr *et al.*, 2008; Yoshinaga-Itano *et al.*, 1998) are associated with better receptive language outcomes in HI children wearing HAs.

As discussed in section 1.1, currently only children with severe-to-profound hearing losses are eligible for cochlear implantation in the UK. These children are particularly vulnerable to the effects of early auditory deprivation. For example, before being implanted, infants with profound hearing loss do not prefer congruent to incongruent audiovisual stimuli in a preferential looking paradigm (Bergeson *et al.*, 2010). Similarly, using visual habituation, Houston *et al.* (2003) found that pre-implantation, children with profound hearing loss do not prefer speech over silent trials. However, gradually, within 2 to 6 months after being implanted, infants began to prefer the speech trials and were able to discriminate between the sounds ‘ah’ and ‘hop hop hop’ (Houston *et al.*, 2003). Children with CIs who were implanted before age 14 months seem to show similar word-learning skills to NH peers, but later-implanted children demonstrate deficits in this skill even as infants (Houston, Carter, Pisoni, Kirk, and Ying, 2005). Indeed, early age at implantation, which is usually associated with a shorter duration of deafness pre-CI, is one of the most important factors associated with positive receptive language outcomes in prelingual CI users (e.g., Geers and Sedey, 2011; Niparko, Tobey, Thal, Eisenberg, Wang, Quittner, Fink, and Team, 2010; Rotteveel, Snik, Vermeulen, Cremers, and Mylanus, 2008; Ruffin, Kronenberger, Colson, Henning, and Pisoni, 2013; Uziel, Sillon, Vieu, Artieres, Piron, Daures, and Mondain, 2007). However, receptive language outcomes are very variable in children with CIs. Other positive outcome factors seem to be a greater amount of residual hearing prior to implantation (Geers, Tobey, Moog, and Brenner, 2008; Niparko *et al.*, 2010; Ruffin *et al.*, 2013), higher maternal education level (Niparko *et al.*, 2010), better quality of language input (Szagun and Stumper, 2013), higher non-verbal intelligence (Geers and Sedey, 2011), parents' higher socioeconomic status (Geers and Sedey, 2011; Niparko *et al.*, 2010; Ruffin *et al.*, 2013), and an oral, rather than total, communication mode (Geers, 2002; Ruffin *et al.*, 2013; Sarant, Blamey, Dowell, Clark, and Gibson, 2001).

Indeed, recent studies demonstrate the great benefit that CIs can bring to profoundly hearing-impaired children – after several years of implant use, the receptive language scores of children with CIs are roughly equivalent to those of children with severe losses of 70 to 85dB HL (Blamey, Sarant, Paatsch, Barry, Bow, Wales, Wright, Psarros, Rattigan, and Tooher, 2001; Lovett, Vickers, and Summerfield, 2015; Rotteveel *et al.*, 2008), with some studies even reporting that many early-implanted CI users are able to perform within NH norms when tested on receptive language in a quiet environment (Geers and Sedey, 2011). Impressively, most long-term CI users are able to use a telephone with a familiar speaker, even though before the implant, they were not even aware of environmental sounds (Beadle, McKinley, Nikolopoulos, Brough, O'Donoghue, and Archbold, 2005; Uziel *et al.*, 2007). However, CI users display significant difficulties in speech perception in background noise (e.g., Svirsky, Teoh, and Neuburger, 2004), possibly because of the poor frequency resolution and reduced temporal fine structure provided by the CI (Moore, 2007) (see section 1.1). CIs also transmit F0 information only weakly (Kuo, Rosen, and Faulkner, 2008), and therefore to be able to perceive pitch and intonation, CI users need to use secondary cues such as temporal information (Chatterjee and Peng, 2008; Kuo *et al.*, 2008). Accordingly, children with CIs have been found to have difficulty in identifying Cantonese tones (Ciocca, Francis, Aisha, and Wong, 2002) and suprasegmental characteristics of speech (Most and Peled, 2007). They are also less able to distinguish voices from each other (Vongpaisal, Trehub, Schellenberg, van Lieshout, and Papsin, 2010).

The reduced quality of moderate to severe HI children's auditory input, especially in the higher frequencies and in terms of frequency resolution, may particularly affect the perception of fine phonetic detail in speech (Skoruppa and Rosen, 2014). Historically, there is ample evidence of HA users' reduced phoneme discrimination skills (e.g., Boothroyd, 1984; Hazan, Fourcin, and Abberton, 1991; Johnson, Whaley, and Dorman, 1984). Recent research shows that many current HA users also have difficulty in identifying and discriminating phonemes, with at most those with moderate losses being able to perform on par with NH children (Borg, Edquist, Reinholdson, Risberg, and McAllister, 2007; Eisenberg, 2007; Halliday and Moore, 2010), and children with moderate losses performing better than those with severe losses (Borg *et al.*, 2007; MacArdle, Hazan, and Prasher, 1999). The majority of studies have shown that consonant recognition is usually more affected than vowel identification. Mild to severe HA users find high-frequency and low-amplitude fricatives, especially /s/ (Borg *et al.*, 2007; Hennies *et al.*, 2012; Stelmachowicz, Pittman, Hoover, and Lewis, 2002; Stelmachowicz *et al.*, 2004), as well as place of articulation (MacArdle *et al.*, 1999; Tsui and Ciocca, 2000), which relies on higher-frequency formant transitions, especially difficult to perceive. There is mixed evidence on whether frequency compression hearing aids assist in consonant recognition (c.f., Ellis, 2012; Glista, Scollie, Bagatto, Seewald, Parsa, and Johnson, 2009; Simpson, Hersbach, and McDermott, 2005, 2006), but studies show that the processing strategy may lead to greater confusion between /s/ and /ʃ/ in phoneme identification, at least for NH adults with simulated sloping hearing loss (Ellis, 2012), likely due to the two fricatives being spectrally closer together when frequency compression is activated. Some difficulties may also be evident in disciminating stop voicing contrasts (Borg *et al.*, 2007; MacArdle *et al.*, 1999; Tsui and Ciocca, 2000), possibly due to the unavailability of visual cues to voicing (Kishon-Rabin, Taitelbaum, Muchnik, Gehtler, Kronenberg, and Hildesheimer, 2002). HA users also usually perform better on audiovisual than audio-only test stimuli (Halliday and Moore, 2010). There is some evidence that HI children with severe hearing losses using HAs use different cue weightings in discriminating stop consonants than do NH peers (Tsui and Ciocca, 2000), and may need greater distinctions between voiced and voiceless stop consonants to distinguish them (Holden-Pitt, Hazan, Revoile, Edward, and Droge, 1995). Even children with mild and moderate hearing losses may have poor phonological processing compared to NH peers (Briscoe, Bishop, and Norbury, 2001; Gilbertson and Kamhi, 1995; Jerger, Martin, and Damian, 2002).

On the other hand, the differences in the acoustic features transmitted by CIs compared to the normal hearing mechanism also affect the acoustic cues that CI users can use to identify and discriminate speech sounds. Accordingly, child CI users perform significantly worse than NH peers on vowel and consonant perception (Medina and Serniclaes, 2009), with studies showing that 8- to 9-year-old children implanted by age 5 achieve approximately 40% of minimal pairs correct (Geers, Brenner, and Davidson, 2003), although CI users can achieve up to 75% correct phonetic contrast perception with several years' experience of CI use (Kishon-Rabin *et al.*, 2002). The temporal envelope transmitted by the CI can be used to identify manner and voicing features in consonants, but the missing temporal fine structure information may be important for perceiving place of articulation and nasal contrasts. The poor frequency resolution of the CI, on the other hand, may affect detailed vowel contrast perception, which is mostly dependent on good spectral resolution (for a discussion, see Giezen, Escudero, and Baker, 2010). Child CI users generally find vowels easier to perceive correctly than consonants (Geers *et al.*, 2003; Kishon-Rabin *et al.*, 2002; Mildner, Šindija, and Vrban Zrinski, 2006), although some studies report that discriminating vowels differing in height may be more difficult, possibly due to F1 frequencies falling within the same electrodes in the implant (Kishon-Rabin *et al.*, 2002). For consonants, manner is indeed easier to perceive than other features (Geers *et al.*, 2003; Mildner *et al.*, 2006), and place of articulation is more difficult (Geers *et al.*, 2003; Kishon-Rabin *et al.*, 2002; Medina and Serniclaes, 2009; Mildner *et al.*, 2006). Differences in voicing can be difficult to discriminate initially (Kishon-Rabin *et al.*, 2002; Mildner *et al.*, 2006), likely due to profoundly hearing-impaired children's lack of prior experience of this distinction which is not available visually, but the perception of voicing can improve greatly over time with CI use (Kishon-Rabin *et al.*, 2002). Other studies report voicing to be an easier feature to perceive (Bouton, Serniclaes, Bertoncini, and Cole, 2012; Tye-Murray, Spencer, and Woodworth, 1995) compared to fricatives (Tye-Murray *et al.*, 1995). Visual information is likely to enhance perception of many of these contrasts for CI users (e.g., Geers *et al.*, 2003; Lachs, Pisoni, and Kirk, 2001). Child CI users are likely to have similar categorical perception to NH peers (Bouton *et al.*, 2012; Medina and Serniclaes, 2009), but may differ in categorical precision, at least in place of articulation perception (Medina and Serniclaes, 2009) and nasality (Bouton *et al.*, 2012), but less so in consonant manner and voicing (Bouton *et al.*, 2012; Medina and Serniclaes, 2009). Adult CI users have also been found to have perceptually less precise categories for the /s/-/ʃ/ contrast than NH adults (Lane, Denny, Guenther, Hanson, Marrone, Matthies, Perkell, Stockmann, Tiede, and Vick, 2007).

A great deal of variability in outcomes of word and sentence comprehension has been found in HI children with moderate to severe hearing loss. For example, in a test of moderate to severe HI children's comprehension of vocabulary, morphology and syntax, Yoshinaga-Itano, Baca, and Sedey (2010) found that half of the 38 children tested who had had early intervention had age-appropriate skills at age 7. HA users seem to have particular trouble in acquiring vocabulary, possibly due to them having less access to 'incidental' word learning (Löfqvist, Sahlén, and Ibertsson, 2010) – they have been shown to progress at about 65% of the rate of NH peers in receptive vocabulary acquisition (Blamey *et al.*, 2001), with studies showing a 2- to 3-year delay in receptive vocabulary development in school-aged children with moderate to severe hearing loss (Blamey *et al.*, 2001; Briscoe *et al.*, 2001; Pittman, 2011). Gilbertson and Kamhi (1995) found that even school-aged children with mild to moderate losses who were 'high-performers' within the participant group scored in the low-average range on receptive vocabulary, while many scored significantly below average, compared to NH children. Moeller's

(2000) research demonstrates that age at intervention and family involvement may be key factors in accounting for the variability found – mild to profound children given early intervention had similar receptive vocabulary skills to NH peers by age 5, but those who were late identified were over 1 standard deviation below average NH peers, with those with additional low family involvement performing even worse on average. However, receptive syntax abilities seem to be similar between children with moderate hearing loss and NH peers, both at preschool (Gilbertson and Kamhi, 1995) and school age (Briscoe *et al.*, 2001).

Child CI users vary greatly in their receptive vocabulary skills – some studies (e.g., Geers and Sedey, 2011; Ruffin *et al.*, 2013) report that over half of preschool and school-aged CI users score within age-appropriate receptive vocabulary norms, while other studies (e.g., Uziel *et al.*, 2007) show that most CI users are delayed in receptive vocabulary. In an open-set sentence comprehension task, long-term CI users attending primary and secondary schools have been found to achieve 70% words correct (Ruffin *et al.*, 2013). Children with CIs may have deficits in syntactic knowledge (Spencer, 2004), and there is some evidence that child CI users are unable to use sentence context to the same extent to NH peers to recognise words (Conway, Deocampo, Walk, Anaya, and Pisoni, 2014; Smiljanić and Sladen, 2013) (although see Eisenberg, Martinez, Holowecky, and Pogorelsky, 2002).

In summary, although child HA and CI users are able to develop speech and language remarkably well, they still display deficits in speech and language perception, especially in receptive vocabulary and in perceiving fine phonetic detail. NH children are also less consistent in their perception of segmental contrasts, and are still learning aspects of vocabulary and syntax of their native language. There is also likely to be a great deal of individual variability in outcomes on these measured in HI children. These factors together are likely to contribute to difficulties in HI and NH children's communication with peers.

### 1.2.2 Speech production development

**Normally-hearing children** To be able to produce adult-like speech, children must be able to (1) control vocal features such as F0, intensity, and duration accurately, (2) produce fine phonetic detail such as vowels and consonants correctly, and (3) use the lexicon and syntax of their native language to plan and produce spoken utterances (Stoel-Gammon, 2011). The vocal tract and the articulators themselves grow and develop until late adolescence (Vorperian, Kent, Lindstrom, Kalina, Gentry, and Yandell, 2005), and therefore a particularly challenging aspect of speech production in childhood is the need to accommodate to this growth over the course of development.

Newborn babies' vocal tracts are still underdeveloped for producing adult-like speech sounds, but particularly rapid growth occurs during the first 18 months of life (Vorperian *et al.*, 2005). Until about 6 months of age, babies mostly make 'primitive' vocalisations, such as crying or cooing, but gradually begin to make vowel-like sounds (see Oller, 2000, for a review). From approximately 6 to 14 months of age, infants engage in 'canonical babbling': producing reduplicated consonant-vowel (CV) syllables. The sounds produced in early babbling are similar across languages (Gildersleeve-Neumann, Davis, and MacNeilage, 2011); usually stops and nasals are the first to be produced due to their relatively simple production (McCune and Vihman, 2001). Gradually, babbling becomes more language-specific, with the consonants and vowels produced becoming more similar to those in the ambient language (de Boysson-Bardies and Vihman, 1991; Vihman and de Boysson-Bardies, 1994). From 12 to 15 months of age, infants start to add suprasegmental features to their babbling, and produce their first words. They gradually increase the number of words acquired until they reach approximately the 50-word stage, after which a rapid increase in expressive vocabulary skills occurs (see Stoel-Gammon, 2011, for a review). Vowels are usually acquired by age 3, while most consonants are produced correctly by age 4 – however, some fricatives, affricates and approximants can produce errors up to about age 7 (Dodd, Hua, Crosbie, Holm, and Ozanne, 2002).

Even once a child has learned to produce the phonemes of their native language correctly, their speech production still differs from that of adults until at least the early teenage years. In general, children have been found to show much greater within-speaker variability than adults, both in acoustic (e.g., Lee, Potamianos, and Narayanan, 1999; Nittrouer, 1993; Nittrouer, Estee, Lowenstein, and Smith, 2005) and kinematic (Smith and Goffman, 1998; Walsh and Smith, 2002) studies of speech. Many researchers have postulated that the greater variability in children and adolescents's productions is due to their immature neuromotor control – typically, children's speech gestures are longer and slower (Cheng, Murdoch, Goozée, and Scott, 2007b; Smith and Goffman, 1998), and their articulators are less synchronised (Cheng, Murdoch, Goozée, and Scott, 2007c; Grigos, 2009) than adults'. Motor control of for example the tongue continues to be refined during adolescence (Cheng *et al.*, 2007b), but jaw movements may become adult-like earlier (Nittrouer, 1993). Alternatively, or additionally, children may simply lack practice in speech production; greater practice with age may lead to faster and more accurate gestures (Koenig, Lucero, and Perlman, 2008; Lee *et al.*, 1999).

To produce pitch (F0) accurately, a speaker needs to be able to control the tension of their vocal folds (Stathopoulos, Huber, and Sussman, 2011). Studies examining both individual vowel production as well as conversational speech samples have demonstrated that the mean F0 produced by children generally declines with age, with mean F0 still higher in 13- to 14-year-old children compared to adults, and greater changes seen during puberty, especially for males (Hazan, Tuomainen, and Pettinato, submitted; Lee *et al.*, 1999; Stathopoulos *et al.*, 2011). The decline with age, as well as the change during puberty, is mainly due to the increase in mass and length of the vocal folds (Stathopoulos *et al.*, 2011). F0 variability is also found to be greater in children than in adults, decreasing with age (Hazan *et al.*, submitted; Pedersen, Møller, Krabbe, Bennett, and Svenstrup, 1990; Stathopoulos *et al.*, 2011). In spontaneous speech, Hazan *et al.* (submitted) found that 9- to 10-year-olds, but not 11- to 14-year-olds, used a wider F0 range than adults – this may be due to young children being less able to control the tension in their vocal folds, or because of differences in vocal fold physiology (Stathopoulos *et al.*, 2011). To manage the intensity of their voice, a speaker needs to use respiratory, laryngeal and neural control (Finnegan, Luschei, and Hoffman, 2000). Again, children may have less control over these articulators – children's voices tend to be more intense than adults' (Hazan *et al.*, submitted; Stathopoulos and Sapienza, 1997, although see Stathopoulos *et al.* (2011)), with Hazan *et al.* (submitted) showing that 9- to 12-year olds, but not older children, produced speech with greater vocal intensity than adults. Similarly, speech rate increases with age (Cheng, Murdoch, and Goozée, 2007a; Hazan *et al.*, submitted;

Lee *et al.*, 1999; Nittrouer, 1993), and has been found to reach adult-like proportions by age 10 or 11 years (Farantouri, Potamianos, and Narayanan, 2008; Hazan *et al.*, submitted), using spontaneous speech, or 13 to 14 years (Lee *et al.*, 1999) using read speech. Walsh and Smith (2002) found greater speech rate variability even in later adolescents compared to adults. Although a fast speech rate could be interpreted as a speaker's fine control over their articulation (Redford, 2014), faster speech rates may also be produced as a result of more effective cognitive and linguistic processing with age (Walsh and Smith, 2002).

Children also produce fine phonetic detail differently to adults, at least until the teenage years. As well as being due to the growing vocal tract, immature motor control or inexperience with speech production, differences in phonetic production between children and adults may also be due to children's internal word representations or phoneme categories being less stable or precise than adults' (Ertmer and Goffman, 2011), as seen in section 1.2.1. This may lead to children's segmental articulatory targets being less robust than adults' (Lee *et al.*, 1999). To produce vowel categories contrastively, a speaker needs to have fine control of the tongue, jaw and lips. Indeed, children's first and second vowel formants (F1 and F2) are more variable than adults', and the variability reduces with age (Lee *et al.*, 1999; Vorperian and Kent, 2007). Vowel formants also become lower in frequency with age, presumably due to the growing vocal tract (Vorperian and Kent, 2007). Children's vowel spaces have been found to be larger than adults' (Flipsen and Lee, 2012; Pettinato, Tuomainen, Granlund, and Hazan, 2016; Vorperian and Kent, 2007), even up to the age of 14 years, perhaps due to children making larger movements to reach targets than adults (Pettinato *et al.*, 2016). Accordingly, it has been suggested that with age, children learn to make more efficient speech gestures and therefore learn to reduce vowels when needed (Redford, 2014). Similarly, vowels become shorter with increasing age (Fletcher, 1989; Lee *et al.*, 1999), which is consistent with articulatory timing being an important feature of more mature articulations (Vandam, Ide-Helvie, and Moeller, 2011).

Children are also more variable than adults in consonant production. In analysing the token-to-token variability of the /p/-/b/ distinction using voice-onset-time (VOT), the primary cue in English for the voiced-voiceless distinction in stop consonants (Lisker, 1978), Romeo, Hazan, and Pettinato (2013) discovered that 9- to 14-year-olds' stop voicing categories in a picture-naming task were more variable than adults'. However, older children's /p/-/b/ categories were also further apart than adults', making the discriminability of the contrast similar to adults'. Similar results were obtained for the spectral /s/-/ʃ/ contrast. Indeed, both of these contrasts require fine control over speech – in producing voicing, speakers must adequately coordinate the timing of the onset of vocal fold vibration with that of the release of the stop closure (Koenig *et al.*, 2008). The fricative place-of-articulation contrast, on the other hand, requires the control over precise tongue movements. Other studies have also found greater variability in children's productions of both voicing and fricatives compared to adults (Koenig *et al.*, 2008; Munson, 2004; Whiteside, Dobbin, and Henry, 2003).

For learning the lexicon of their native language, children are much less constrained by physiology, and new words are used throughout a person's lifetime (Nippold and Duthie, 2003). Utterance length increases from childhood to adolescence and into young adulthood, and more complex syntax is used with age (Nippold, Hesketh, Duthie, and Mansfield, 2005). Likely due to the greater variability present in their speech, children's speech is less intelligible than at least female adults' (Markham and Hazan, 2004), and the intelligibility of children's speech increases with age (Baudonck, Buekers, Gillebert, and Van Lierde, 2009). Redford's (2014) study showed that children's speech intelligibility is likely related to better temporal motor control over speech – listeners rated children with faster articulation rates as being more intelligible than their peers with slower articulation rates.

**Hearing-impaired children** The above review of normally-hearing children's speech production demonstrated that the differences in speech production observed between adults and children can be attributed to immature speech motor control, the growth of the vocal tract, inexperience with speech production, as well as the developing linguistic knowledge of the phonemes and words in the speaker's native language. HI children's speech production is affected not only by these factors, but also by poorer speech perception skills, which are strongly correlated with their speech production skills (e.g., Eisenberg, 2007; Tye-Murray *et al.*, 1995). Their deficit in speech perception reduces their ability to compare the acoustics of their own speech to that of adults (Niparko, 2009), and the reduced auditory feedback makes it difficult for HI children to monitor their own speech and to learn to map their articulation onto acoustics (Halliday and Moore, 2010; Rahilly, 2013). Instead, HI infants and children need to rely more on visual, kinesthetic and tactile cues, which do not provide as accurate feedback as an acoustic signal does (Niparko, 2009). Similarly, due to their poorer perception skills, HI children are likely to have less linguistic experience with their native language, and may therefore be delayed in linguistic knowledge compared to NH children.

During the first 6 months of life, NH and HI infants' primitive vocalisations tend to be similar, but HI infants with severe losses may continue to use these simple vocalisations for longer than NH infants (Oller, Eilers, Bull, and Carney, 1985). Even HI infants with early amplification and moderate to severe hearing loss show less canonical babbling behaviour than NH infants (Nathani, Oller, and Neal, 2007; von Hapsburg and Davis, 2006), and their babbling may differ from that of NH infants – for example, moderate to profound infant HA users are slower to acquire more complex syllables than NH infants (Iyer and Oller, 2008; Moeller, Hoover, Putman, Arbataitis, Bohnenkamp, Peterson, Wood, Lewis, Pittman, and Stelmachowicz, 2007b). Even those with mild-to-moderate hearing losses are delayed in the onset of phonemic speech development (Yoshinaga-Itano and Sedey, 2000). However, although phonetic and phonological development is considered to be delayed in HI children with milder losses, it is usually qualitatively similar to NH children's speech development (Moeller, Hoover, Putman, Arbataitis, Bohnenkamp, Peterson, Lewis, Estee, Pittman, and Stelmachowicz, 2007a).

Cochlear implants can make a large difference to the speech production outcomes of severely and profoundly hearing-impaired children. Infants with profound hearing loss who receive cochlear implants before the age of 30 months have been found to rapidly acquire accuracy in consonant production post-implantation (Connor, Craig, Raudenbush, Heavner, and Zwolan, 2006). Ertmer, Jung, and Kloiber (2013) showed that by 12 months of CI hearing experience, infants who were implanted by 3 years of age were producing a majority of speech-like utterances – a feat achieved by NH children only at 18 months of age – therefore demonstrating rapid early speech production development after implantation.

However, although the stages of development that CI infants go through are generally similar to those of NH infants, fricatives, affricates and liquids may still be acquired later in hearing age than in NH infants (Ertmer and Goffman, 2011), and their babbling and speech behaviour may still differ qualitatively from NH infants' (c.f., Lederberg and Beal-Alvarez, 2011).

For HI children wearing HAs, better speech production outcomes are more likely to be achieved with milder degrees of hearing loss (Yoshinaga-Itano and Sedey, 2000) and early identification and intervention (Yoshinaga-Itano *et al.*, 1998). As with the speech perception literature, the recent speech production literature has, however, placed greater emphasis on studying children with severe-to-profound hearing loss wearing cochlear implants (Elfenbein, Hardin-Jones, and Davis, 1994; Wake, Hughes, Poulakis, Collins, and Rickards, 2004). For children wearing CIs, better speech production outcomes are associated with early implantation (Connor *et al.*, 2006; Montag, AuBuchon, Pisoni, and Kronenberger, 2014; Tye-Murray *et al.*, 1995), longer CI use (Nicholas and Geers, 2006), and a more oral communication mode (Tobey, Geers, Douek, Perrin, Skellet, Brenner, and Toretta, 2000; Tobey, Geers, Sundarrajan, and Lane, 2011; Tye-Murray *et al.*, 1995; Uchanski and Geers, 2003). CI users make much greater gains in speech production accuracy than would be expected of their hearing level with hearing aids, and typically achieve similar speech production skills to HI children with severe losses wearing HAs (Blamey *et al.*, 2001).

In terms of voice characteristics, the speech production of HI children with milder losses is unlikely to be greatly affected by their hearing loss – however, HI children with more severe losses wearing HAs have been found to exhibit problems in speech motor control and the coordination of the articulators (Kosky and Boothroyd, 2003; McGarr and Campbell, 1995; Niparko, 2009), as well as to show greater variability in acoustic measures of speech production (Allen and Arndorfer, 2000; McGarr and Campbell, 1995), compared to NH peers. For example, children wearing hearing aids with moderate-to-severe to profound hearing loss may have difficulty in controlling and coordinating their laryngeal and respiratory muscles (O'Halpin, 2001), and are sometimes reported to produce higher mean F0 (Higgins, McCleary, Carney, and Schulte, 2003; Nakamura, Gilbert, and Robb, 2007; Ryalls and Larouche, 1992) or more monotonal F0 (Pratt and Tye-Murray, 2008), as well as greater F0 variability in their speech (Allen and Arndorfer, 2000; Hocevar-Boltezar, Radsel, Vatovec, Geczy, Cernelc, Gros, Zupancic, Battelino, Lavrencak, and Zargi, 2006) compared to NH controls. Similarly, Allen and Arndorfer (2000) found that adult listeners judging sentences produced by six 7- to 14-year-old HA users who had severe-to-profound hearing loss were often unsure whether the sentences were interrogatives or declaratives, showing that HI children may have difficulty in intonation production. Child HA users may also produce greater amplitude variability (Hocevar-Boltezar *et al.*, 2006), longer sentence durations and both hypo- and hypernasal speech (Fletcher, Mahfuzh, and Hendarmin, 2001). Although these findings are consistent with those of older studies, usually investigating children with severe-to-profound hearing losses wearing HAs (e.g., Gilbert and Campbell, 1980; Monsen, 1979; Osberger and McGarr, 1982), most of the research cited above tested few participants and elicited speech using read sentences or individual syllables rather than spontaneous speech, and therefore it is unclear whether problems with F0 production, intensity and speech rate occur in children with moderate to severe hearing losses using modern, more advanced HAs.

Due to the limited transmission of F0 information in the CI (see 1.2.1), it seems likely that HI children wearing CIs will also have trouble in controlling F0. Indeed, HI children wearing CIs have been found to exhibit poor tone production in tonal languages (e.g., Peng, Tomblin, Cheung, Lin, and Wang, 2004; Xu, Chen, Lu, Zhou, Wang, Liu, Li, Zhao, and Han, 2011), as well as poor intonation contour production for interrogative sentences in English (Chin, Bergeson, and Phan, 2012; Peng, Tomblin, Spencer, and Hurtig, 2007; Peng, Tomblin, and Turner, 2008; Snow and Ertmer, 2009). Similarly to HI children wearing HAs, HI children wearing CIs also tend to produce higher (Higgins *et al.*, 2003) and more variable F0 (Campisi, Low, Papsin, Mount, Cohen-Kerem, and Harrison, 2005) than NH peers, which does not diminish with experience with the implant – mean F0 may even increase post-CI (Campisi *et al.*, 2005). This demonstrates that the acoustic input of CI users in terms of F0 is not sufficient to restore normal speech motor control. Child CI users are also found to use greater amplitude variation, which however does decrease with longer use of the implant (Campisi *et al.*, 2005). Lenden and Flipsen's (2007) study of six prelingually deafened severe to profound CI users aged between 3 and 6 years of age in which the children's voice characteristics were rated by professionals, only a subset of CI users were found to have problems in producing speech at an adequate intensity level – most children were judged to be within normal range for this measure. Similarly to severe and profound HI children using HAs, child CI users tend to speak more slowly than their NH peers even 4 years post-implantation (Burkholder and Pisoni, 2003; Chuang, Yang, Chi, Weismer, and Wang, 2012; Uchanski and Geers, 2003), perhaps partly due to an increased number and length of pauses (Burkholder and Pisoni, 2003; Chuang *et al.*, 2012). This indicates that children with CIs may still be slower in their linguistic processing and planning of speech production. The speech and voice characteristics of CI users reviewed above demonstrate, therefore, that CI users continue to have problems with speech motor control and speech production planning in several voice dimensions, even after several years of CI use.

The production of phonological contrasts by HI children with mild to moderate losses is often found to be similar to that by NH peers (Elfenbein *et al.*, 1994). However, the production of segmental contrasts by HA users with severe to profound losses is likely to be more affected, especially if the hearing loss is identified late (Eriks-Brophy, Gibson, and Tucker, 2013). As evidenced in section 1.2.1, HI children's cue-weighting strategies for segmental contrasts may differ from those of NH peers due to their difficulties in perceiving some important acoustic cues (e.g., Tsui and Ciocca, 2000). Therefore it is possible that any productions of phonological contrasts by HI children which do not adhere to NH norms may have arisen due to HI children's phonological systems themselves differing from those of NH peers. Alternatively, HI children may have a similar phonological system to NH peers, but the articulatory mapping of the contrasts and speech motor control may differ from controls' (Monsen, 1979).

Although vowel categories are likely to be easier to perceive for HI children than consonants (c.f., 1.2.1), children with more severe hearing losses are likely to receive less tactile feedback from vowel production than from consonant production (Monsen, 1976), which may increase vowel production difficulty. HI children wearing HAs with mild-to-moderate to profound losses have been found to produce longer vowels with more variable durations (Allen and Arndorfer, 2000;

Nicolaidis and Sfakiannaki, 2007; Vandam *et al.*, 2011), and a restricted F2 range (Rahilly, 2013) compared to NH peers. Some studies also indicate a restricted F1 range for prelingual HI speakers (c.f., Rahilly, 2013). HI children with more profound losses and wearing HAs may also have a reduced vowel space area compared to NH peers (Horga and Liker, 2006; Nicolaidis and Sfakiannaki, 2007; Rahilly, 2013). Older studies on HI children with severe to profound hearing losses wearing hearing aids show similar results (e.g., Monsen, 1976).

In particular, certain fine vowel distinctions, such as the tense-lax distinctions in English, may be difficult for HI children with more severe losses to produce. Monsen's (1974) study of 12 HI children between 13 and 16 years of age with severe to profound hearing losses demonstrated that the speakers tended to produce the spectral-temporal /i/-/ɪ/ distinction using consistently greater durational distinctions between the two vowels than NH peers. Although Monsen (1974) did not measure whether HI speakers also used the primary spectral cue to distinguish the vowels from each other, he postulated that HI speakers' phonology may discriminate these two phonemes more in terms of durational rather than spectral cues, due to duration being an easier cue for them to perceive. Bernhardt, Gick, Bacsfalvi, and Ashdown's (2003) adolescent HI participants with severe-to-profound hearing losses also found vowel tense-lax distinctions difficult to produce accurately. Similarly, Harris, Rubin-Spitz, and McGarr (1985) report that six severe to profound HI adolescents' productions of vowel formants were significantly more variable than NH peers', which suggests that instead of having a deviant phonological system, HI speakers have a more immature speech motor control system than NH speakers.

In production, as with perception, consonants are likely to be more greatly affected by hearing loss. Elfenbein *et al.* (1994), in their study of 40 HI children between 5 and 18 years of age with mild to profound hearing loss completing a sentence-reading task, found that even children with only mild hearing loss produced fricatives erroneously. A similar result was reported by Wiggin, Sedey, Awad, Bogle, and Yoshinaga-Itano (2013) in their cohort study of 269 early-identified children with mild to profound losses. Another study examining the production of the /s/-/ʃ/ contrast by six moderate to profound HI children between the ages of 8 and 12 years discovered that only approximately 58% of their productions were perceived as being in the correct category by the experimenter, with more of the HI speakers' productions being perceived as belonging to the lower-frequency /ʃ/ category than the /s/ category (Kosky and Boothroyd, 2003). However, the pronunciation accuracy of consonants other than fricatives and affricates are generally not affected by hearing losses under 75 dB HL (Elfenbein *et al.*, 1994). Children with more severe losses are more likely to use processes such as consonant substitutions or omissions in their speech (Elfenbein *et al.*, 1994; Eriks-Brophy *et al.*, 2013). They may also exhibit problems with controlling vocal fold adduction, which may lead to errors in consonant voicing (Pratt and Tye-Murray, 2008). Postlingually deafened adults who have acquired the voicing categories before the onset of deafness have been found to produce a smaller VOT contrast between voiced and voiceless plosives than NH adults (e.g., Lane, Wozniak, and Perkell, 1994), which suggests that poor auditory feedback can affect the production of these categories also for prelingual HI children with more severe losses.

Although early-implanted CI users' segmental speech production greatly improves after implantation, they continue to display delays in acquiring speech sounds, and produce phonetic contrasts with reduced accuracy and greater variability compared to NH peers approximately 2 years post-implant (Ertmer and Goffman, 2011) (see also Wiggin *et al.*, 2013). After 4 years of implantation, most CI users implanted by age 5 managed to use 66% of all English speech sounds in word context correctly – although all monophthongs were acquired, speakers often produced diphthongs and consonants erroneously (Serry and Blamey, 1999). Tobey, Geers, Brenner, Altuna, and Gabbert (2003) showed that 108 8- to 9-year-old children implanted between 1.5 and 5.5 years of age produced, on average, only 62 % of all vowels and 68% of all consonants accurately, therefore demonstrating a significant delay in both vowel and consonant production compared to NHs. Children implanted between 2.5 and 7.5 years of age seemed to improve in their production of phonemes up to approximately 8 years post-implantation, after which accuracy plateaued (Tomblin, Peng, Spencer, and Lu, 2008). This is probably at least partly due to some of the acoustic cues to phonemes being limited by CI processing (see 1.2.1), therefore limiting the accurate acquisition of certain phonemes. Uchanski and Geers (2003) discovered that 8- to 9-year-old

CI users, who had worn their implants for at least 4 years, produced imitated vowels which were generally longer than those of NH peers. However, most of the CI users produced the /i/ and /ɔ/ vowels with similar F2 ranges to NH controls. Other research points towards a reduced F2 range in CI children compared to NH peers (Chuang *et al.*, 2012; Löfqvist *et al.*, 2010; Neumeyer, Harrington, and Draxler, 2010), and longer and more variable vowel durations in child CI users compared to peers (Vandam *et al.*, 2011). Löfqvist *et al.* (2010) also report that their adolescent CI users produced lower mean first formants but similar mean second formants to NH controls. Similarly, several studies suggest that child CI users produce smaller vowel space areas compared to those produced by NH peers (Chuang *et al.*, 2012; Liker, Mildner, and Šindija, 2007; Löfqvist *et al.*, 2010; Neumeyer *et al.*, 2010).

In terms of consonant production, Uchanski and Geers (2003) discovered that 70% of the CI users' /s/-/ʃ/ productions were within NH norms, although the spectral mean of the /s/ categories of the remaining 30% of tokens were often close to NH peers' /ʃ/ productions, and therefore the /s/-/ʃ/ categories were unlikely to have been produced distinctively by the speakers. Similarly, 70-80% of speakers' /t/-/d/ tokens had VOT discriminability within NH range, although within-speaker variability was found to be much higher for CI than for NH speakers. Another study (Grant, Bow, Paatsch, and Blamey, 2002) (cited in Stelmachowicz *et al.*, 2004) reported that CI children had higher accuracy in producing the /s/-/z/ distinction than did children with hearing aids – this shows that CI users are likely to benefit from input in the higher frequencies compared to HA users.

As with their receptive language skills (1.2.1), HI children wearing HAs are also likely to be delayed in their expressive language skills. In a cohort study of 89 HI children aged 7 to 8 years with congenital mild to profound hearing loss, the HI children scored below NH norms on a test of expressive language skills (Wake *et al.*, 2004), and greater hearing losses were associated with poorer expressive language skills (Wake, Poulakis, Hughes, Carey-Sargeant, and Rickards, 2005). Several studies on HI children wearing HAs have examined the 'mean length of utterance' (MLU), generally considered to measure the syntactic complexity of children's utterances (Koehlinger, Van Horne, and Moeller, 2013). Flipsen and

Kangas (2014) reviewed several older studies on MLU in HI children wearing HAs and concluded that most studies report that HI children have shorter MLUs than NH peers. Similarly, McGuckian and Henry (2007) found that 7-year-old HI children with moderate losses had similar MLUs to 3-year-old NH children, and Koehlinger *et al.* (2013) showed that 3- to 6-year-old HI children with mild to moderately severe hearing loss in spontaneous conversation produced 0.25 to 0.5 words less per utterance than NH peers. HI children with even mild to moderate losses have also been found to have specific deficits in producing bound morphemes and finite verb morphology (Elfenbein *et al.*, 1994; Koehlinger *et al.*, 2013) (although see Norbury, Bishop, and Briscoe, 2001).

Children using CIs also exhibit delays in expressive language scores compared to NHs, although they make greater gains with a CI with age than would be expected from their pre-implantation scores (Niparko *et al.*, 2010). Similarly to children with HAs, MLUs of children with CIs have generally been found to be lower than those of NH peers (Flipsen and Kangas, 2014; Geers *et al.*, 2003; Szagun, 2004). However, although Geers *et al.* (2003) report that their 8- to 9-year old CI group had significantly lower MLUs on average (mean: 4.0) than NH peers (mean: 6.5), half of the CI users were within 1 standard deviation of the mean of NH peers. MLUs in CI children are also likely to increase over time (Flipsen and Kangas, 2014).

As would be expected from the above review of the literature, the intelligibility of HI children's speech typically also varies according to the severity of their hearing loss. Elfenbein *et al.* (1994), in their study of 40 5- to 11-year old HI children with moderate to severe hearing losses wearing HAs, report that their group of mildly hearing-impaired children were nearly all rated as being within NH range in intelligibility, while only 55% of those with moderate losses, and 33% of those with moderately severe to severe losses were within NH range. In a review of the recent literature on the intelligibility of the speech of HI children wearing CIs, Flipsen (2008) notes that children wearing CIs achieve much higher speech intelligibility scores than historically was seen in children with severe-to-profound hearing impairments wearing hearing aids (see Gold, 1980, for a review). Flipsen's (2008) review also shows that HI children wearing CIs make rapid progress on their speech intelligibility after being implanted, especially those who were implanted early, and that the improvement in speech intelligibility continues until at least 10 years after implantation.

The generally lower intelligibility scores for HI children compared to NH peers are likely to be explained mainly by two factors: their reduced speech motor control and/or slower speech planning, as well as their less discriminable phonetic categories. As with NH children (Redford, 2014), HI children who speak at faster rates are considered more intelligible by naive listeners than those who speak more slowly (Metz, Samar, Schiavetti, Sitler, and Whitehead, 1985; Montag *et al.*, 2014; Pisoni and Geers, 2000), likely due to faster speech rates indicating greater speech articulatory control and faster speech planning (Burkholder and Pisoni, 2003). Metz *et al.*'s (1985) extensive study, which related 20 prelingually deafened HI young adults' speech production measures (including several different phonetic contrasts and speech rate) to intelligibility ratings of their speech, discovered that the main factor accounting for more intelligible speech in this population was temporal and spectral phoneme discriminability. Interestingly, Metz *et al.* (1985) reported that the mean difference between phoneme categories was more important for good speech intelligibility than less variable within-phoneme categories, which suggests that the HI child's internal phonological structure may be more important in determining their intelligibility to naive listeners than fine speech motor control. Monsen (1978) also found that the VOT difference between /t/-/d/, the F2 difference between /i/-/ɔ/, and the rated quality of liquid and nasal consonants accounted for 78% of the variation in intelligibility for HI adolescents' speech. Any differences between prosody produced by HI and NH children may not, on the other hand, affect intelligibility to a great extent – a study by McGarr and Osberger (1978) did not find strong associations between prosodic characteristics of 11- to 12-year-old HI children's speech and their speech intelligibility. HI children's speech intelligibility seems to be strongly correlated with other speech and language outcomes, but not with measures of non-linguistic outcomes – Montag *et al.*'s (2014) study of 63 prelingually deafened CI users demonstrated that the participants' speech perception scores in quiet, their receptive vocabulary measures and receptive and expressive language scores (CELF) were correlated with the intelligibility of their speech in read sentences, but their backward digit span and nonverbal IQ were not.

In summary, HI children have been found to have general difficulties with speech motor control, with likely slower linguistic processing and planning, as well as possible differences in internal phonological structure compared to NH children. Many of these differences to NH peers are similar to those produced by younger NH children whose speech production systems are still developing. All of these factors are likely to affect the decreased speech intelligibility of HI children with more severe hearing losses.

### 1.2.3 Summary

Altogether, therefore, communication difficulties between NH and HI children may arise from their still-developing speech perception and production systems, as well as directly from the HI child's hearing problem. NH children are likely to have more immature speech motor control, and therefore to produce more variable speech as well as speech at higher fundamental frequencies and with higher formant frequencies compared to adults – on the other hand, HI children, especially those wearing hearing aids, are likely to have limited audibility at higher frequencies, and to have deficits in discriminating between phonetic contrasts. In producing speech, HI children may exhibit problems with speech motor control, and produce greater speech variability than NH children. They may also produce less contrastive phonetic categories. These in turn may be especially problematic for NH children to perceive, as they exhibit less precise perception of phonetic categories and speaker variability until at least the early teenage years.

The above studies demonstrated the breadth of research examining the perception and production skills of both NH and HI children. However, in all but a few exceptions (i.e., Hazan *et al.*, submitted; Pettinato *et al.*, 2016), the speech production of the NH children was examined either in read speech or in conversation with an examiner – situations which seem closer to assessing the performance, rather than competence of a child. To our knowledge, the speech of HI children has not previously been examined in communication with a peer. As discussed in the introduction, peer interaction is an important aspect of communicative competence, especially for older children and adolescents. Therefore it is essential to move towards examining the speech produced by NH and HI children in peer interaction, as is done in the current study. Additionally, most of the literature reviewed above on HI children's production of segmental contrasts examine either the % correct of phonemes produced, with objective measures rarely taken, and elicitation often of only a few tokens per participant. Recent research has demonstrated the importance of using a large number of tokens per phonetic contrast to assess children's speech production (c.f., Romeo *et al.*, 2013) – an approach taken in this study. Finally, this study will determine the differences between NH and HI children's speech production using a comprehensive set of measures, from voice characteristics to segmental and linguistic, to give a better estimation of the similarities and differences between HI and NH children's everyday speech communication.

## 1.3 Speech adaptations to interlocutors' needs

Given the many potential problems that HI children have when listening to speech, HI children's communication partners – parents, professionals and peers alike – are often instructed on the strategies to use to make communication easier when speaking with the HI child (e.g., Action on Hearing Loss, 2012; Caissie and Tranquilla, 2010; Doyle and Dye, 2002; Marschark and Hauser, 2012; NDCS, 2012b). This section seeks to explore the potential strategies which may be helpful when interacting with a HI child, and to review literature on whether NH adults, NH children and HI children themselves may be able to use these and other strategies to adapt to the needs of their listener.

### 1.3.1 Suggested speech enhancement strategies

**Global acoustic-phonetic strategies** As the reduced audibility of speech is one of the main difficulties that HI children face, it would be likely that an important beneficial strategy for enhancing communication with HI children would be to increase the intensity of the speech produced. An increased intensity, especially in the mid-frequency 1-3k Hz range which provides important cues to formant transitions, could provide listeners with greater access to the phonetic cues in the speech (Cooke, King, Garnier, and Aubanel, 2014a). Indeed, some speakers who are inherently more intelligible to NH listeners seem to produce speech at greater intensities between 1 and 3k Hz (Hazan and Markham, 2004; Krause and Braida, 2004). However, as demonstrated by instructions to communication partners of HI children not to shout (e.g., Action on Hearing Loss, 2012), increasing intensity by shouting is likely to distort phonetic cues and lip patterns, and therefore to decrease intelligibility (Junqua, 1993).

An increase in intensity is often accompanied by an increase in mean F0 (Cooke *et al.*, 2014a), but this may occur due to its close relationship to increased subglottal pressure in the vocal folds (e.g., Titze, 1989), and may not contribute to a speech intelligibility increase per se (Bond and Moore, 1994; Bradlow, Torretta, and Pisoni, 1996). Communication partners are also often instructed to use "a full range of intonation" and to put "acoustic emphasis on keywords" (Caissie and Tranquilla, 2010, p.100), thus increasing the F0 variability in their speech to make parts of the speech signal more salient. Indeed, a few studies have shown that speakers with inherently more variable F0 tend to be more intelligible to NH listeners (e.g., Bradlow *et al.*, 1996).

Some instructions to communication partners suggest that decreasing speech rate can be beneficial for talking with people with a hearing-impairment (Caissie and Tranquilla, 2010; Marschark and Hauser, 2012), although others warn not to speak too slowly (Action on Hearing Loss, 2012; NDCS, 2012b). Indeed, a slower speech rate may give the HI child more time to process the input signal, but speech that is too slow compared to normal may distort lip-reading (Action on Hearing Loss, 2012; NDCS, 2012b). Some studies have failed to find an association between a speaker's inherent speech rate and their speech intelligibility (Bradlow *et al.*, 1996), although others show that speakers with inherently slower speaking rates are more intelligible than faster talkers (e.g., Hazan and Markham, 2004). However, even then, a slower speech rate may be associated with greater time to reach articulatory targets, and may therefore not contribute to speech intelligibility on its own (Hazan and Markham, 2004). Similarly, increasing the number and length of pauses at phrase boundaries may also be beneficial to HI listeners in easing utterance parsing (Cooke *et al.*, 2014a) – an increased number of pauses has been associated with greater intelligibility of speech in some studies (e.g., Liu and Zeng, 2006).

**Segmental contrast enhancement** Communication partners are also often asked to "speak clearly" (Action on Hearing Loss, 2012) or to "enunciate carefully" (Caissie and Tranquilla, 2010, p. 100) to HI persons. Indeed, making more extreme articulatory gestures and therefore increasing the mean distance between phonetic categories is likely to be a beneficial strategy when talking to HI children who have difficulty in perceiving phonetic contrasts accurately. Furthermore, enhancing phonetic contrasts by approximating articulatory targets is likely to increase the consistency of a speaker's phonetic categories, which in itself may increase intelligibility to HI listeners — NH listeners have been found to benefit from less within-category dispersion in speakers' phonetic categories (Hazan, Romeo, and Pettinato, 2013; Newman, Clouse, and Burnham, 2001). Other studies show that a speaker's intelligibility may instead be associated with the amount of spectral overlap between categories, at least in speakers with dysarthria (Kim, Hasegawa-Johnson, and Perlman, 2011). Clinical populations are indeed often found to have greater overlap in their phonetic categories than typical populations, which may account for their decreased intelligibility to NH listeners (Haley, Seelinger, Mandulak, and Zajac, 2010). Even in NH populations, both children and adults are found to have overlap between their phonetic categories (Romeo *et al.*, 2013). However, within-category dispersion and overlap of segments is rarely reported in investigations of HI children's segmental contrast production.

In vowels, an increased F1 and F2 range, leading to less vowel reduction and therefore a larger vowel space area and a greater spectral distance between vowels, has been found to be positively correlated with the inherent intelligibility of a speaker (Bond and Moore, 1994; Bradlow *et al.*, 1996; Hazan and Markham, 2004; Smiljanić and Bradlow, 2009). Longer vowels are also associated with greater intelligibility (Ferguson and Kewley-Port, 2007), presumably due to giving speakers more time to reach articulatory targets, and due to giving listeners more time to process the vowel. In one of the only studies examining the factors contributing to increased speech intelligibility for both NH and HI listeners, HI older adults' increased understanding of speech was associated more greatly with longer vowel durations and lower F1 vowel frequencies than NH younger adults'. On the other hand, elderly HI listeners' increased speech intelligibility was much less associated with F2 frequencies than NH younger listeners' (Ferguson and Quené, 2014).

Ferguson and Quené (2014) speculate that this difference was due to elderly HI adults' high frequency losses, which prevented them from using F2 information to a great extent when listening to vowels.

Similarly, it is likely that a strategy of increasing consonantal contrasts would enhance communication with a HI child. For example, Maniwa, Jongman, and Wade (2008) discovered that 'clearly spoken' sibilant fricatives were more discriminable to listeners with simulated sloping hearing losses than those that were spoken in a more 'casual' manner. However, more clearly spoken non-sibilant fricatives did not increase intelligibility for these listeners. Speakers with greater VOT distinctions between voiced and voiceless stops have also been shown to be more intelligible to NH listeners listening in noise than speakers who produce smaller VOT distinctions between categories (Bond and Moore, 1994).

**Linguistic and conversational enhancements** Many of the instructions given to communication partners emphasise the use of linguistic communication enhancements in conversation with a HI person. For example, when talking to a HI person, communication partners are asked to use shorter and less complex utterances (Marschark and Hauser, 2012) and increase sentence contextual information (NDCS, 2012b). In cases of misunderstanding, instructions also emphasise rephrasing rather than repeating the message (Action on Hearing Loss, 2012; Doyle and Dye, 2002; Marschark and Hauser, 2012). These linguistic simplification strategies are likely to be especially important in conversations with HI children, because of the delays they exhibit in receptive vocabulary and language skills (see section 1.2.1). The provision of sentential context has indeed been found to be beneficial for NH listeners (van Rooij and Plomp, 1991). Communication partners are also often asked to use visual cues and gestures when talking with a HI interlocutor (Action on Hearing Loss, 2012; Doyle and Dye, 2002; Marschark and Hauser, 2012; NDCS, 2012b).

**Summary** The studies cited above demonstrate that many of the suggested strategies for communicating with HI children may indeed enhance the intelligibility of the speaker. Most of the studies, nonetheless, employed NH adults as listeners in intelligibility tests, and therefore it cannot be confirmed whether the same cues enhance intelligibility for those with hearing loss (Ferguson and Quené, 2014). For example, Krause and Braida's (2009) study, which used signal processing algorithms to mechanically enhance intensity in the speech, found some similarities, but some differences, in the way in which three HI listeners perceived the processed speech, compared to NH controls. Unfortunately, it is not well understood which factors contribute well to the intelligibility of a speaker even to NH listeners (see Cooke *et al.*, 2014a; Smiljanić and Bradlow, 2009, for comprehensive reviews).

Importantly though, it is unclear whether HI children's communication partners do indeed change their communicative behaviour when talking with a HI listener (Pichora-Fuller, Goy, and van Lieshout, 2010). Adult HI persons report that if they ask their interlocutor to change their speech, for example to speak more slowly, the interlocutors often only do so for a few utterances (Pichora-Fuller *et al.*, 2010). Additionally, although there is evidence that NH adults are able to make adaptations to the needs of an interlocutor, NH children and HI children themselves may not be able to do so. The following sections will review the kinds of adaptations that NH adults, NH children and HI children have been found to make, to explore whether they may be able to use these strategies to adapt to the needs of a HI interlocutor.

### 1.3.2 Normally-hearing adults

**Clear speech studies** A large number of studies examining the acoustic-phonetic properties of 'clear speech' ask adult participants to 'speak as if to a hearing-impaired listener', as compared to speaking 'casually, as if to a friend' (or similar instruction) while reading sentences on a screen – speakers are found to decrease their speech rate, insert pauses, and increase their mean F0, F0 range and speech intensity, especially in the 1-3kHz range, in 'clear' speech, compared to 'casual' speech (Ferguson and Kewley-Port, 2002, 2007; Ferguson, Poore, and Shrivastav, 2010; Ferguson and Quené, 2014; Picheny, Durlach, and Braida, 1986; Smiljanić and Bradlow, 2005, 2008b; Van Engen, Chandrasekaran, and Smiljanić, 2012) (for a recent review, see Smiljanić and Bradlow, 2009). As well as these 'global' speech modifications, clear speech studies present some evidence that speakers also enhance the phonological contrasts in their language in clear speech (Bradlow and Bent, 2002) – for example, the spectral (Ferguson and Kewley-Port, 2002) or durational (Granlund, Hazan, and Baker, 2012) distance between tense and lax vowels is increased, and vowel space is expanded (e.g. Bradlow, 2002), the difference between the VOT of voiced and voiceless stops becomes larger (Granlund *et al.*, 2012; Smiljanić and Bradlow, 2008b), and word-final stops are produced more frequently (Picheny *et al.*, 1986). Perception studies on read clear speech have shown that, when listening in noise, clear speech benefits both NH adults (Ferguson, 2004; Ferguson and Kewley-Port, 2002; Helfer, 1997; Schum, 1996) and HI adults between 9 and 30 percentage points (Ferguson, 2012; Helfer, 1998; Liu, Del Rio, Bradlow, and Zeng, 2004; Payton, Uchanski, and Braida, 1994; Picheny, Durlach, and Braida, 1985; Zeng and Liu, 2006) (although see Ferguson and Kewley-Port, 2002; Ferguson and Lee, 2006). Children wearing cochlear implants have also been found to improve their speech recognition when listening to clear speech compared to casual speech in noise, although less so than NH children (Smiljanić and Sladen, 2013). Other populations, such as children with learning disabilities (Bradlow, Kraus, and Hayes, 2003) and non-native listeners, have also been found to benefit from clear speech changes, although non-native listeners benefit from them less than do native listeners (Bradlow and Alexander, 2007; Bradlow and Bent, 2002). Clear speech also increases sentence recognition memory in noise (Gilbert, Chandrasekaran, and Smiljanić, 2014; Van Engen *et al.*, 2012). Despite the overall finding of speakers modifying the acoustic-phonetic properties of their speech when speaking clearly, a large amount of between-speaker variability has been observed, both in the strategies used (e.g., Ferguson and Kewley-Port, 2007) and in the amount of intelligibility benefit that clear speech produces (e.g., Ferguson, 2004).

However, studies on read clear speech suffer from several shortcomings. There is a great deal of evidence that read and spontaneous speech have different acoustic-phonetic properties (see Wagner, Trouvain, and Zimmerer, 2015, for a review). For example, fewer phonetic reductions are produced in read than in spontaneous speech (Ernestus, Hanique, and Verboom, 2015) – an important consideration in clear speech research, in which the changes in vowel formant values in different conditions are often investigated. If vowels are reduced less in read speech in general, there may be less scope for speakers to enhance vowel formants in read clear speech. Indeed, Hazan and Baker (2011) found that the acoustic-phonetic properties of clear speech produced in read and spontaneous conditions differed significantly. Similarly, read speech does not allow a speaker to use any other kinds of speech adaptations which may occur in spontaneous speech – for example, linguistic modifications to speech may occur frequently in cases of communication breakdown, and may interact with the acoustic-phonetic modifications made (Lind, Campbell, Davey, Rodgers, Seipolt, and Akins, 2010). Recent research has also shown that speakers are very sensitive to the type of instructions given to them – in Lam, Tjaden, and Wilding's (2012) study, speakers produced different acoustic-phonetic speech modifications when asked to 'speak as if to a hearing-impaired listener' rather than to 'speak clearly'. In clear speech studies involving read speech, because speakers are given limited control in terms of the aspects of the speech which they can modify, and specific instructions on the type of speech they are asked to produce, higher-level planning of the utterance may not be required; speakers may be able to concentrate their cognitive resources on changing the acoustic-phonetic aspects of their speech to a greater extent than in spontaneous clear speech. For example, Volden, Magill-Evans, Goulden, and Clarke (2006) found that participants with autistic traits were better able to make typical clear speech modifications to imaginary listeners when specifically instructed to do so, compared to without instruction. An essential aspect of speech communication is being able to make these adaptations spontaneously when required by the speaking situation (Hazan *et al.*, submitted). Therefore, it is possible that the findings of clear speech research exaggerate the strategies used by speakers in their everyday lives.

Most importantly, however, a significant shortcoming of clear speech research is that it considers speech production and perception as two distinct processes which can be investigated separately. However, taking a listener's needs into account is always likely to occur in an interactive situation, in which each interlocutor is continuously contributing to a shared understanding of the dialogue (Pickering and Garrod, 2004, 2013), and in which both the listener's characteristics and feedback from the listener influence the speech produced. Therefore communicative intent is vital for the ecological validity of speech in studies investigating adaptations to listener needs (Hazan and Baker, 2011). Accordingly, previous research has shown that the presence of a listener influences the speech produced (Charles-Luce, 1997; Garnier, Henrich, and Dubois, 2010; Scarborough, Brenier, Zhao, Hall-Lew, and Dmitrieva, 2007) – this may be an especially important confounding factor in clear speech studies, as, despite being asked to 'talk as if to a hearing-impaired listener', most speakers are reported not to be familiar with talking to HI persons. They may, therefore, be less sensitive to the type and extent of adaptations required to ensure successful communication with a HI interlocutor, and it is unclear what basis they use to modify their speech when faced with these instructions. Recent studies have begun to use more naturalistic interactive situations to investigate intelligibility-enhancing speaking styles (e.g., Cooke and Lu, 2010; Hazan and Baker, 2011). The current study follows this trend by using communicative tasks to elicit HI-directed speech.

**Evidence for adaptation** There is ample evidence that speakers use strategies specific to the communication barrier encountered by the listener to enhance the intelligibility of their speech (see Cooke *et al.*, 2014a, for a review). If the listener is experiencing difficulty due to a poor acoustic listening environment, such as a competing talker, multi-talker babble, speech-modulated noise, or a great distance between interlocutors, speakers are found to increase their speech intensity, especially the energy in frequencies between 1 and 3k Hz, and to increase their mean F0 (e.g., Cooke and Lu, 2010; Garnier *et al.*, 2010; Hazan and Baker, 2011; Pelegrín-García, Smits, Brunskog, and Jeong, 2011). The greater speech intensity in these conditions likely increases the signal-to-noise ratio (SNR), i.e., the audibility of the speech in the masker (Cooke *et al.*, 2014a), and the enhanced energy between 1 and 3k Hz probably 'boosts' the audibility of the particular frequencies which contain important information, such as second formant frequency changes, in speech (Krause and Braida, 2004). The magnitude of speech production changes increases with greater energetic masking of the speech (Cooke and Lu, 2010). In fluctuating noise maskers, speakers seem to actively attempt to produce more speech during the 'dips' or pauses in the masker, to avoid temporal overlap between the speech and the masker, implying that speakers employ a 'listening while talking' strategy (Aubanel and Cooke, 2013; Cooke and Lu, 2010). Speakers also adapt to the specific communication barrier experienced by listeners, even if the speakers are themselves unable to hear the distortion that listeners are subjected to – Hazan and Baker (2011) found that speakers who had previously been exposed to vocoded speech increased their F0 range when listeners heard them through multi-talker babble, but this strategy was not adopted when listeners heard speakers through a vocoder, which does not transmit F0 information. Subsequent listening tests confirmed that when listening through multi-talker babble, listeners were quicker to process speech produced in that masker than the speech produced through the vocoder (Grynpas, Baker, and Hazan, 2011). Thus, in tailoring their speech to listener needs, speakers seem to make use not only of their own current experience of the noise masker, but also of either their own previous experience of the masker, or of implicit feedback provided by the interlocutor.

Speakers also adapt their speech and language due to specific characteristics of the listener. If the listener lacks sufficient linguistic knowledge due to being a non-native speaker of the language, speakers are found to decrease the cognitive effort required by listeners to understand them (Cooke *et al.*, 2014a), in simplifying their speech by speaking more slowly, using shorter and syntactically less complex utterances, a more restricted range of vocabulary, lexically more frequent content words and frequent repetition of words (see Costa, Pickering, and Sorace, 2008; Larsen-Freeman and Long, 1991; Long, 1983, for reviews). They may also enhance the linguistic information in the signal by increasing prosodic and speech-segmentation cues with exaggerated intonation contours which produce a wider F0 range, and using longer and more frequent pauses (Larsen-Freeman and Long, 1991) (although for differing results, see Hazan, Uther, and Granlund, 2015). There is also some evidence that speakers' 'foreigner-directed-speech' (FDS) attempts to enhance the segmental cues in a language by increasing the distance between the point vowel categories (expanded vowel space) (Hazan *et al.*, 2015; Scarborough *et al.*, 2007; Uther, Knoll, and Burnham, 2007). However, other findings suggest that some cues to phonological contrasts in FDS are in fact attenuated (Sankowska, García Lecumberri, and Cooke, 2011). Interestingly, the FDS effect may be driven mostly by feedback from the listener – while a prior expectation as to the proficiency level of a non-native listener was only found to affect speakers' mean length of utterance (MLU) at the beginning of a conversation, the effect of positive or negative feedback from the non-native interlocutor affected the MLU of speakers during the entire conversation (Warren-Leubecker and Bohannon, 1982).

Infants also lack linguistic knowledge due to their age, and 'infant-directed-speech' (IDS) research has identified several typical cues of speaking with infants. As with FDS, IDS elicits shorter and syntactically less complex utterances with frequent lexical repetitions and pauses (Cristia, 2010; Snow, 1972) as compared to adult-directed speech (ADS). Production of IDS also leads to an expansion of the vowel space – this is generally thought to help infants in learning the specific vowel categories of the ambient language (Kuhl, Andruski, Chistovich, Chistovich, Kozhevnikova, Ryskina, Stolyarova, Sundberg, and Lacerda, 1997; Uther *et al.*, 2007). However, some recent evidence suggests that vowel space expansion may be a side-effect of other IDS-related changes, such as speech rate reduction, as the acoustic difference between minimal pair vowel (Cristia and Seidl, 2014) and consonant (McMurray, Kovack-Lesh, Goodwin, and McEchron, 2013) contrasts does not seem to differ between IDS and ADS. Recent evidence also shows that the perceived clarity in vowel and consonant contrasts (Martin, Schatz, Versteegh, Miyazawa, Mazuka, Dupoux, and Cristia, 2015) does not increase in IDS compared to ADS. Similarly, the variability within vowel contrasts has been found to increase in IDS compared to ADS, which seems unlikely to be helpful to an infant language learner (McMurray *et al.*, 2013). Unlike FDS, IDS is critically thought to include an affective/attentional component, which adults typically display using heightened F0 mean and F0 variability in their speech (Burnham, Kitamura, and Vollmer-Conna, 2002; Fernald and Kuhl, 1987; Uther *et al.*, 2007). The extent of the F0 changes, as well as the syntactic complexity and lexical variety of sentences, seems to be adapted according to infant age (Henning, Striano, and Lieven, 2005; Hoff-Ginsberg, 1986; Kitamura and Burnham, 2003; Vosoughi and Roy, 2012). There is evidence that, as the child develops, the linguistic level of the child influences the caregivers' word frequency and MLU (Roy, Frank, and Roy, 2009). The amount of linguistic simplification and F0 variability in IDS may also depend on the perceived maturity of the child's face (Zebrowitz, Brownlow, and Olson, 1992).

On the other hand, speaking to a pet ('pet-directed-speech'; PDS) seems to activate only attentional and affective speech adaptations – Burnham *et al.* (2002) found that although the F0 changes present in IDS occurred in PDS to a similar extent, the 'linguistic/didactic' component of IDS, vowel space expansion, did not. The extent of vowel space expansion seems to depend on the linguistic competence of the listener, with findings showing that parrot-directed-speech elicits greater vowel space expansion than dog-directed-speech (Xu, Burnham, Kitamura, and Vollmer-Conna, 2013); it may be that the perceived intelligence of the animal, rather than its responsiveness to speech, affects the type of speech produced (Sims and Chin, 2002).

As outlined in section 1.2, the difficulties experienced by prelingually deafened HI listeners occur both due to the limited acoustic input they receive and due to a developmental delay in their linguistic knowledge. Therefore, as discussed in 1.3.1, we would expect HI-directed speech to display both audibility-related speech modifications (as in environmentally degraded conditions) and message simplification/linguistic enhancement of cues (as in FDS and IDS).

Several studies, mostly examining syntactic and pragmatic language use, have investigated the adaptations made by NH mothers when talking to an infant HI listener in a communicative context (see Spencer, 2003, for a review), with many studies reporting that NH mothers have similar maternal interactions with language-level matched HI and NH children (e.g., Gallaway, Hostler, and Reeves, 1990; Lederberg and Everhart, 2000). A few studies have also investigated the acoustic-phonetic aspects of mothers' talk to their HI infants. Kondaurova, Bergeson, and Dilley (2012) demonstrated that there were no differences in mothers' use of infant-directed speech to HI and age-matched NH infants in terms of enhancing the spectral and durational aspects of tense and lax vowels. Lam and Kitamura's (2010) case study of a mother speaking to her hearing-impaired infant found that, surprisingly, the mother decreased her vowel space when talking to her HI infant compared to when speaking to his NH twin. In a follow-up experiment, Lam and Kitamura (2012) placed 48 mothers and their NH infants in separate rooms so that the mother and child could interact through a video camera. They manipulated whether the mother knew the infants could hear them, and whether the infants could in fact hear the mother. Mothers' vowel hyperarticulation was not found to be affected by the mother's knowledge of infant hearing status, but instead simply by whether their infant could hear them - less hyperarticulation was observed in conditions in which the infant could not hear the mother. A similar experimental set-up in Smith and Trainor (2008) demonstrated that mothers' F0 variation depended on infant feedback. These findings imply that, at least in IDS, listener feedback drives F0 variability and hyperarticulation - perhaps due to the increased effort needed to maintain an infant's attention when they cannot hear, vowel hyperarticulation is sacrificed. Other studies show that mothers are sensitive to the amount of hearing experience that the child has had and adjust the acoustic-phonetic properties of their speech accordingly (Bergeson and McCune, 2006; Kondaurova and Bergeson, 2011).

Although the above studies offer valuable insight into the speech input of young HI children, they are unlikely to be representative of HI-directed speech to older children, especially with their confounding of infant-directed and HI-directed speech. Very few other studies have been conducted on HI-directed speech. Imaizumi, Hayashi, and Deguchi (1993, 1995) recorded Japanese teachers of the deaf playing a game with HI children with profound hearing losses and NH children. Their findings indicated that, in HI-directed speech, teachers simplified their utterances, and also devoiced certain segments, to a greater extent than in NH-directed speech, probably to aid the HI children's speech segmentation and understanding. Lind *et al.* (2010) compared non-repair and repair sequences of speech produced by a NH 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 (but not duration) of pauses in repair sentences – all of these changes would be compatible with the speaker using both acoustic and linguistic enhancements in speech. In case studies using Conversation Analysis (Sacks, Schegloff, and Jefferson, 1974), Skelt (2010, 2013) discovered that partners of postlingually deafened adults seem to be sensitive to the visual needs of their HI interlocutors and only initiate or continue talk when their interlocutor is gazing at them. However, the HI listeners in these case studies were postlingually deafened, and therefore may not have had difficulties with delayed linguistic knowledge as prelingually deafened children typically do. Most other work on communication between NH and HI adults concerns the types of repair strategies used in interaction (e.g., Caissie and Gibson, 1997; Sparrow and Hird, 2010), rather than specific communication strategies used in HI-directed speech.

Most of the above studies demonstrate that speakers are adept at adapting to the particular needs of their listener, tailoring their speech in response to different types of noise, different ages, different language proficiency levels and, at least to some extent, to the hearing status of the listener. Although the adaptation literature seems to imply that speakers use information about the listener's current listening environment (e.g., Aubanel and Cooke, 2013; Cooke and Lu, 2010), their assessment of a listener's characteristics (e.g., Sims and Chin, 2002; Zebrowitz *et al.*, 1992), and listener feedback (e.g., Lam and Kitamura, 2012; Smith and Trainor, 2008; Warren-Leubecker and Bohannon, 1982) to guide them in making these adaptations – and clear speech research shows that, when faced with an imagined interlocutor, speakers also use their understanding of the listener's particular listening difficulty in determining the strategies they use in producing speech (e.g., Lam *et al.*, 2012) – it is unclear what the mechanisms behind these specific adaptations might be. However, to be able to understand the development of this skill in children, and to predict whether populations other than typical adults may be able to adapt to the needs of listeners, it is vital to consider the underlying mechanisms of adaptation. The next section will investigate theories which have been postulated to account for speaker-listener adaptation.

**Adaptation theories** Many clear speech studies use the notions given in Lindblom's (1990) 'Hyper-Hypo' (H&H) theory to explain speakers' abilities in adapting the acoustic-phonetic properties of their speech. According to Lindblom (1990), 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, whenever a listener's needs differ from that of the speaker's, the listener's needs "win" (Bard and Aylett, 2005). Although Lindblom's (1990) theory was specifically postulated to explain acoustic-phonetic variation and is mostly used in that domain, it could be applied to all levels of communication. However, the theory has been criticised for being cognitively too demanding for a speaker; a speaker needs to continuously monitor the listener to detect signs of miscomprehension, update their internal listener model if misunderstandings occur, and apply the updated listener model simultaneously to the speech output, even on the articulatory level in which consultation of the listener model would need to be done frequently while planning the articulation of each individual segment (Bard and Aylett, 2005).

Less computationally costly alternatives have also been hypothesised, mainly in the field of Audience Design. Brown and Dell (1987) and Dell and Brown's (1991) 'Monitoring' hypothesis postulates that, at the beginning of a conversation, the speaker does not pay attention to listener needs, and only takes them into account when provided with explicit feedback. Thus, adapting to listener needs is a gradual process. Similarly, according to the 'Copresence' hypothesis (Fussell and Krauss, 1992; Horton and Keysar, 1996; Schober, 1993), the speaker makes an initial assessment of the listener's needs according to known listener characteristics, and speakers then infrequently attend to listener feedback and adjust their speech as necessary. Finally, Bard and colleagues (Bard, Anderson, Sotillo, Aylett, Doherty-Sneddon, and Newlands, 2000; Bard and Aylett, 2005) suggest in their 'Dual-process' approach that a speaker only models a listener's needs if sufficient cognitive resources are available, i.e., if other task demands are not too high. Even if a speaker has enough cognitive resources at their disposal to model the listener, they propose that speakers will only adapt their speech and language on higher linguistic levels, rather than on the fine phonetic level. This is because, as pointed out in the criticism of Lindblom's (1990) hypothesis, having to consult a listener model while planning the articulation of each segment in an utterance would be too computationally costly – however, higher linguistic levels, such as the lexical level, are slower processes which would incur less of a cost if a listener model was to be consulted during speech production planning (Bard and Aylett, 2005).

Notably, the Dual-process approach and the H&H hypothesis give very different accounts of the ability of speakers to adjust their speech to the listener on the fine phonetic level. Although a great deal of evidence reviewed above shows that speakers adapt their speech on the utterance, word and global acoustic levels – by using strategies such as simplifying their speech, using lexically more frequent words, increasing the intensity of their speech and adjusting F0 – there is much less research demonstrating that modifications are made on the segmental level.

In the clear speech literature, evidence does exist that some aspects of phonology are enhanced – for example, when a speaker is asked to speak clearly in American English, word-final stop bursts are released more often and alveolar flapping is less frequent (Bradlow *et al.*, 2003; Picheny *et al.*, 1986). Additionally, many studies show that both consonant and vowel durations increase in clear speech (Bradlow *et al.*, 2003; Krause and Braida, 2004; Picheny *et al.*, 1986). However, these modifications, while affecting the segmental level, do not necessarily enhance segmental contrasts as such, rather than provide more redundant cues to particular segments to assist the listener in phoneme identification. The literature provides less evidence for segmental contrast enhancement in clear speech. For consonants, the clear speech literature has mainly examined the voiced-voiceless distinction in word-initial plosives – although speakers generally demonstrate an enhancement of the contrast when raw VOT is measured (e.g., Granlund *et al.*, 2012; Smiljanić and Bradlow, 2008a), the contrast does not appear to be enhanced when calculated relative to the speech rate decrease (Smiljanić and Bradlow, 2008a) (although see Kang and Guion, 2008). Numerous clear speech studies have also shown that the point vowel space is enhanced in clear speech compared to casual speech (e.g., Bradlow *et al.*, 2003; Ferguson and Kewley-Port, 2002, 2007) – however, counterintuitively, this enhancement has been shown to be of a similar size in languages with fewer and therefore less confusable vowels (such as Spanish and Croatian) as well as languages with a more crowded vowel space (such as English) (Bradlow, 2002; Smiljanić and Bradlow, 2005). The tense and lax vowel distinction is also enhanced similarly in different languages despite the languages differing in the primary and secondary cues to the distinction (Granlund *et al.*, 2012; Smiljanić and Bradlow, 2008a; Wassink, Wright, and Franklin, 2007), and in Granlund *et al.* (2012) it was shown that late bilinguals' English /i/-/ɪ/ vowels became spectrally *closer* together in clear speech compared to casual speech. Additionally, one of the seminal clear speech studies discovered that the English tense vowel space expanded less than the lax vowel space in clear speech (Picheny *et al.*, 1986) – although not discussed by the authors, this would suggest that the tense and lax vowels for their speakers were in fact spectrally closer to each other in clear speech than in casual speech. Together, these findings imply that both consonant and vowel enhancement in clear speech is not necessarily driven by segmental contrast enhancement – changes between clear and casual speech may occur for example as a side effect of other, more global clear speech changes, such as an increased effort or speaking rate changes (see, for example, Cristia and Seidl, 2014, for a similar argument regarding IDS). However, there is some evidence that the tense-lax vowel distinction in English and other languages is enhanced *durationally*, by lengthening the tense vowel more than the lax vowel, in clear speech (Granlund *et al.*, 2012; Picheny *et al.*, 1986; Uchanski, Choi, Braida, Reed, and Durlach, 1996) (although see Smiljanić and Bradlow, 2008a), and therefore some evidence for contrast enhancement in clear speech remains.

In the adaptations literature, few studies have examined segmental enhancement beyond vowel space measures. In an interactive task, Sankowska *et al.* (2011) discovered that the VOT distinction between voiced and voiceless plosives was maintained in FDS and when speaking through noise compared to a casual condition in which native speakers completed the task together in quiet conditions, while the effect of vowel shortening before voiced codas was *attenuated* in both FDS and through noise. Similarly, Cristia and Seidl (2014) compared several tense-lax vowel contrasts in both ADS and IDS, and found that most contrasts were either maintained or attenuated in IDS compared to ADS, and not increased, as has previously been hypothesised. However, in a similar experiment, the VOT contrast between /p/ and /b/ and the spectral contrast between /s/ and /ʃ/ was enhanced in IDS directed to 12- to 14-month-old infants, but not to younger 4- to 6-month-old ones (Cristia, 2010). Studies (e.g., Garnier *et al.*, 2010) have also demonstrated that visual cues, such as lip protrusion, may be used by speakers to enhance the distinction between /i/ and /u/ vowels.

Additionally, a few studies have investigated segmental contrast enhancement directly by feedback through a speech recogniser. For example, Schertz (2013) asked participants to read a list of words to a computer, which in turn would give the speaker feedback on which word it thought the speaker had produced. The feedback, which was in fact controlled by the experimenter, was either the correct target word, an erroneous word which differed either in consonant voicing or in vowel identity to the target word, or a general error message asking the speaker to repeat. Speakers were only allowed to respond to the feedback by repeating the same word. Results showed that vowels were not enhanced spectrally in either the open-ended question condition nor in the minimal pair error condition. The voicing distinction between initial plosives was enhanced, but only in the minimal pair error condition, and only in the instance that the minimal pair error differed specifically in its voicing to the target word. Similar results of enhancement after minimal pair misrecognition were obtained for fricative distinctions by Maniwa, Jongman, and Wade (2009), although Ohala's (1994) study on voicing and vowel enhancement failed to find any such differences between conditions. These studies demonstrate that speakers are able to make specific changes to the phonetic contrasts in their speech, but may do so only when confronted with specific local misrecognition errors – and situations in which no alternative enhancement strategies are available.

Altogether, therefore, there is mixed evidence for segmental contrast enhancement when adult speakers adapt to a listener's needs, with few studies other than those in the clear speech literature exploring the topic. However, studies which give speakers specific misrecognition feedback on their productions of minimal pair contrasts have shown that, unlike hypothesised by the Dual-process approach (Bard and Aylett, 2005), speakers can and do adapt, at least at times, to the needs of a listener even on the segmental level.

In summary, each of the models reviewed in this section proposes that speakers, at least at some point in the conversation, model listener characteristics, take the listener's feedback into account, and apply it to, at least some parts, of their output.

### 1.3.3 Normally-hearing children

The previous section demonstrated that adults are adept at adapting to the needs of their interlocutor when faced with interlocutors of different ages, linguistic knowledge and hearing status, as well as with communication through different physical barriers. It also showed that several different types of adaptations, such as global acoustic-phonetic enhancements, linguistic enhancements, and possibly even segmental contrast enhancements, may be employed by adult speakers. However, few studies have examined listener adaptations in children, and, to our knowledge, none of the theories reviewed in section 1.3.2 include a hypothesis on the development of the ability to adapt to an interlocutor.

Piaget (1926) is one of the only psychologists to suggest a developmental model of listener adaptation. According to Piaget (1926), children's speech communication progresses from being mostly egocentric at 3-5 years old, consisting mostly of repetition and monologue, to being fully socialised, and therefore adapted to listener needs, after age 7. However, nowadays, Piaget's (1926) view is considered overly simplistic, with critics pointing out that, given the multidimensional nature of communication, the development of the ability to adapt to listener needs is likely to involve a complex set of developmental changes (Schmidt and Paris, 1984). This section will, therefore, review the development of some of the basic skills which the adult theories in section 1.3.2 imply are required for a speaker to be able to adapt their communication to a listener's needs. In particular, following Flavell (1974) (cited in Asher and Wigfield, 1980), it is hypothesised that a speaker would need to develop the knowledge of others' perspectives being different from one's own (perspective-taking), the ability to infer the listener's perspective and to infer the modifications which may be helpful for the listener (detecting and responding to feedback) and the flexibility to apply the information on listener needs to modify speech and language output appropriately (flexibility in speech and language).

**Perspective-taking** Most of the theories on speaker adaptation reviewed in the previous section require that, during an interaction, the interlocutors engage in some type of listener modelling. To be able to do so, speakers would presumably need to be skilled in taking the listener's perspective. For adult-like perspective-taking, a child would first need to have knowledge of the existence of others' perspectives, and also of the need to take them into account in certain situations.

'Theory of mind' is defined as the understanding that each person's mind does not reflect reality, but is made up of mental states such as beliefs, knowledge, wants and feelings, which are related to the person's experience, and which lead to a diverse range of behaviours (Baron-Cohen, Leslie, and Frith, 1985). It can be defined as the ability to think about what another person might be thinking (Marschark and Hauser, 2012). Theory of mind is thought to develop through observation, listening to others' conversations, effective early interaction, play and pretend play, all of which lead a child to learn about mental states and how they relate to behaviour (Courtin, Melot, and Corroyer, 2008; Duncan *et al.*, 2014; Marschark and Hauser, 2012; Russell, Hosie, Gray, Scott, Hunter, Banks, and Macaulay, 1998). Having a theory of mind enables a child to explain and predict social behaviour, as well as to understand others' emotions (Russell *et al.*, 1998), and is therefore essential for social interaction.

In the theory of mind literature, most studies explore what is termed by Miller (2000) as 'situational' belief, i.e. an understanding that different experiences in a situation cause people to hold different beliefs. A classic experiment used in the literature to test situational belief is the 'false-belief' task. In one variant of the task (Wimmer and Perner, 1983), a participant is shown a scenario of a boy putting a bar of chocolate in a cupboard, and then leaving the room. After the boy has left, his mother takes the chocolate from the cupboard, and puts it in the drawer instead. The participant is then asked to tell the experimenter whether the boy will look for the chocolate in the cupboard or in the drawer. If the participant possesses theory of mind, they will know that the boy will look for the chocolate in the cupboard, although the chocolate is in fact in the drawer. Typically, 3-year-olds answer that the boy will look for the chocolate in the drawer, as they presume that the boy's mind reflects reality, whereas most 4- and 5-year olds know that the boy will look for the chocolate in the cupboard. Psychologists use these outcomes to claim that by the end of preschool, children have an understanding that people's behaviour is governed by their beliefs, which depend on their experience of the world, and that their own mental states may differ from others' (Wellman, Cross, and Watson, 2001) – essential components for being able to take another person's perspective into account in a conversation.

Very little research has been conducted on situational belief beyond preschool age but, if more difficult tasks are used, there is evidence that theory of mind develops further even into late adolescence. In Keysar, Lin, and Barr's (2003) Director task, a message director has a different view of a grid than the participant. The director instructs the participant to move objects around in the grid, but, to move the correct object, the participant has to take into account the director's differing perspective. Using a variant of the Director task with participants aged between 7 and 27, Dumontheil, Apperly, and Blakemore (2010) found that although participants of all ages demonstrated that they knew that the director had a different perspective to themselves, task accuracy improved with age – younger participants seemed to operate more egocentrically, without taking the director's view into account in their task moves. Dumontheil *et al.* (2010) suggest this may be related to the continued development of executive functions and working memory throughout adolescence, as well as possibly a simple tendency for younger age groups, despite knowing the other's perspective, not to take it into account. These findings suggest, therefore, that, at least in more difficult task situations, older children may not use the listener's perspective to guide their behaviour in social interaction, even if they know that the listener's perspective of a situation is different to their own.

As well as 'situational' belief, children need to become aware of 'individual' belief – the understanding that, regardless of the situation, different people have different pre-existing knowledge (Miller, 2000). This type of understanding is essential for children to be able to realise that, for example, they need to use more simple talk to babies than to their peers because the knowledge and abilities between the two groups differ. Generally, children seem to be better at judging others' 'global' abilities, rather than any specific type of knowledge (see Miller, 2000, for a short review). The few studies conducted on this topic have shown that even 3- to 4-year-olds realise that age is associated with a person's abilities, but their understanding of age-related development in ability is limited and less accurate than 7-year-olds' (Dowker, Hart, Heal, Phillips, and Wilson, 1994, cited in Miller (2000)). Similarly, both 6- and 12-year-olds understand that intelligence increases with age, but 12-year-olds have a more sophisticated understanding of the development of intelligence (Montangero, 1996, cited in Miller (2000)).

For a child to be able to model a hearing-impaired listener in an interaction, they also need to have acquired individual belief in understanding that peers with disabilities may be less competent in some aspects than themselves, and in knowing the deficits associated with the particular disability. Typically-developing preschool children seem to be able to understand only more salient physical disabilities (Diamond and Hestenes, 1996; Smith and Williams, 2001, 2005), and they tend to generalise disabled peers' limited competencies to unaffected areas – for example, by claiming that children with emotional/behavioural problems have problems also in physical domains (Smith and Williams, 2001). They do not seem to have an understanding of the consequences of sensory impairments, with Diamond and Hestenes (1996) reporting that preschool children were confused as to why a hearing-impairment might affect a peer's language abilities. However, this may also be due to limited experience – there is evidence that regular contact with peers with disabilities increases preschoolers' understanding of disabilities (Hong, Kwon, and Jeon, 2013).

Throughout middle and late childhood there is an increase in the understanding of disability, its causes, its irreversibility and its effects (Lewis, 1993; Smith and Williams, 2001, 2005), with older children being better able to specify the consequences of certain disabilities, rather than generalising across domains. They also have an increased understanding of the psychological impact of disability, although less salient disabilities such as severe learning difficulties can still be somewhat difficult to comprehend for 11-12-year-olds (Lewis, 1993). Some of the problems associated with understanding disabilities may be related to the lack of experiencing the disability themselves – primary school students' positive attitudes towards peers with hearing impairments has been found to increase if they are given a chance to try the disability in a simulation, perhaps due to their increased understanding of the disability and, as a consequence, being better able to relate to hearing-impaired peers (Hurst, Corning, and Ferrante, 2012).

In summary, by late preschool age, children have the knowledge that, in a certain situation, other people's perspectives may differ from their own, although the ability to use this knowledge in situational tasks continues to develop throughout childhood and adolescence, possibly due to the continued development of executive functioning and working memory. Preschool-aged children also have an understanding of different people having different pre-existing knowledge and abilities, for example due to disability or differences in age. However, younger children's knowledge of these differences is likely to be crude, and it is not until late primary school age that children begin to have a more advanced understanding of the specific consequences of disabilities. Older children are, therefore, likely to possess the knowledge needed to model a hearing-impaired listener's perspective in an interaction.

**Detecting and responding to feedback** As well as acquiring the knowledge that the listener may have a different perspective to the speaker, the theories on adaptation to listener needs suggest that the speaker must be able to infer the listener's perspective online through feedback, to know what may be helpful for the listener, and to respond adequately to the feedback.

Clarification requests, i.e., a listener's indication to the speaker that the speaker's message has not been understood, are the most common type of listener feedback investigated in the literature. Clarification requests can be non-specific, such as 'What' or 'I don't understand', which only tell the speaker that the message was not understood, or they may be specific requests, which convey the particular part of the message which the listener did not understand (for example, if the speaker has described a colourful toy, a listener may ask 'Sorry, what colour is the toy?') (Levy, 1999). Clarification requests may be 'stacked' when a listener has to repeat the clarification request several times in consequent turns before obtaining an adequate repair response from the speaker. Feedback can also be more implicit and non-verbal, such as a frown or silence after a message which requires a response (Peterson, Danner, and Flavell, 1972).

Even 2-year-olds seem to be able to understand that a response is required if the listener asks a clarification request (Gallagher, 1981; Wellman and Lempers, 1977), although their responses tend to be inappropriate. For example, Gallagher (1981) found that although 2-year-olds consistently responded to confirmation requests, they mostly answered them affirmatively. With age, inappropriate responses rapidly decrease – in a picture-description task, 3-year-olds were found to give inappropriate responses, such as off-topic or discontinued responses, approximately 33% of the time to an initial non-specific clarification request, but 5-year-olds did so very rarely (Brinton, Fujiki, Loeb, and Winkler, 1986). Children also learn to distinguish between different types of clarification requests to infer the exact informational needs of a listener – Ferrier, Dunham, and Dunham (2000) showed that 33-month-olds, but not 27-month-olds, were able to differentiate between non-specific and specific clarification requests by giving the listener only the particular information that was needed in response to a specific clarification request, and repeating their previous utterance after a listener's non-specific clarification request. Similar results were also reported by Gallagher (1981) and Anselmi, Tomasello, and Acunzo (1986). By preschool age, therefore, children likely start to be able to respond appropriately to simple listener feedback, and are able to infer some of the basic informational needs of an interlocutor based on the type of clarification request given. However, Schmidt and Paris (1984) and Levy (1999) note that non-specific clarification requests are essentially questions to which speakers can provide answers without having to necessarily refer back to the original utterance. Message repetition in response to non-specific requests, on the other hand, shows that the speaker is aware of the need to reply to the clarification request, but does not necessarily imply any greater awareness of the listener's needs (Schmidt and Paris, 1984).

A key issue, therefore, is the age at which children begin to use so-called 'adaptive' repair strategies in response to non-specific clarification requests. Adaptive responses are messages in which the form or content of the original utterance has been revised, or in which additional information has been added, or a cue (such as defining terms or giving background information) is provided. Use of such repair strategies is indicative of a more sophisticated understanding both of the listener's needs and of potentially beneficial interaction strategies (Schmidt and Paris, 1984). Indeed, 3- to 4-year-old children seem to require specific clarification requests to be able to produce responses which are not simply repetitions (Nilsen and Mangal, 2011). Spilton and Lee (1977) examined repair sequences in 16 4-year-olds' play with a peer and found that children rarely produced adaptive repair responses when given non-specific clarification requests. In a referential communication task, Peterson *et al.* (1972) gave 24 4- to 5-year-old and 24 7- to 8-year old children who were describing drawings implicit non-verbal feedback, a non-specific clarification request ('I don't understand') and a more explicit specific clarification request ('Look at it again. What else does it look like? Can you tell me anything else about it?') on different occasions. They found that both age groups responded adaptively to the specific clarification request, but that only the 7- to 8-year-old group used adaptive repair responses to the non-specific clarification request. Interestingly, neither group responded adaptively to the implicit non-verbal clarification request, indicating that even 7-year-olds may not be completely mature communicators in this regard. Similarly, in Brinton *et al.* (1986), 40 2- to 9-year-old children described pictures to an investigator who sat behind a screen. At certain time points during the participants' picture descriptions, the investigator initiated a stacked non-specific clarification sequence. Children over 7 years of age were found to use adaptive strategies more than younger children, and only the 9-year-old group seemed to be able to provide cue responses. Although most of the above studies induced communication breakdown fairly artificially, on the basis of these studies it seems that the ability to adaptively respond to listener feedback develops fairly late into middle childhood.

Referential communication tasks have also been used to investigate whether speakers are able to infer a listener's needs without reference to listener feedback. Typically, in these studies, a speaker and listener are divided by a screen, and each of them is given an array of cards which differ from each other in certain attributes. The speaker's task is to describe a specific card to the listener so that the listener is able to choose the correct card. Lloyd, Mann, and Peers (1998) tested 578 children between 5 and 11 years of age as speakers on a referential communication task, and found that speakers' use of unambiguous item descriptions increased with age – only approximately 15% of cards described by 5-year-olds were readily identifiable in the array from the participant's initial description, while about 65% of 11-year-olds' descriptions were perfectly unambiguous. Similar age-related trends were also observed by Whitehurst (1976). Therefore it seems that, at least in this type of fairly demanding task, even 11-year-olds are still learning about the type of information that may be beneficial to a listener. However, approximately 65% of 5-year-olds' descriptions did contain some attributes which were contrastive between cards in the array, showing that the 5-year-olds did have an understanding that they needed to make the descriptions clear for the listener. It also seems that younger children are able to learn to produce less ambiguous descriptions if given adequate training – Matthews, Lieven, and Tomasello (2007) found that 2-, 3- and 4-year-olds benefited from being trained to produce adequate initial descriptions in a 'sticker game', similar to a referential communication task. Specifically, the children who had been trained to produce less ambiguous messages by practicing being the speaker benefited significantly more from the training than did those who were trained either by being the listener or as an observer. These results suggest that even younger children benefit from listener feedback and are able to learn to take the listener's needs into account to a greater extent.

Interestingly, the ability to monitor one's own comprehension also seems to be a fairly late development – in Lloyd *et al.*'s (1998) study, participants also acted as listeners in the referential communication task, and the investigator provided the participants with either an ambiguous or an unambiguous description. He then asked the participant whether or not they knew which card he had. Similar age-related trends as in the speaker-condition were observed, with 5-year-olds only detecting message ambiguity in 10% of cases, but 11-year-olds able to do so in approximately 70% of cases (see also Lloyd, Camaioni, and Ercolani, 1995). Therefore, as even in early primary school children are found to have trouble in monitoring their own comprehension, it seems unlikely that they would implicitly be able to do so for a listener. The ability to monitor both one's own and an interlocutor's understanding, in being able to take a listener's feedback into account and to be able to respond to it adaptively, have been linked to theory-of-mind capabilities (Reches and Pérez Pereira, 2007; Roberts and Patterson, 1983) and linguistic competence (e.g. Deutsch and Pechmann, 1982; Nilsen and Mangal, 2011). Altogether, then, the ability to detect and respond appropriately to feedback is likely to have developed fairly well into late childhood, although some skills, such as detecting message ambiguity, may well continue to develop even into adolescence.

**Flexibility in speech and language** Even if a speaker has modelled their listener accurately, and is aware of the speech and language modifications that would be helpful to the listener, they may still be unable to modify their speech and language appropriately due to an inflexible or inexperienced speech production system. For children, modifying their speech and language output may be particularly difficult due to their still-maturing speech production system (discussed in detail in section 1.2.2), which is mainly characterised by (1) immature speech motor control, such as longer and slower articulatory gestures (e.g., Smith and Goffman, 1998), (2) greater variability in the production of fine phonetic detail (e.g., Lee *et al.*, 1999), likely due to both the still-developing speech motor control and less stable internal phoneme categories, and (3) inexperience with the linguistic system of the native language, as evidenced by the continued acquisition of vocabulary and syntax in the language (e.g., Nippold *et al.*, 2005).

To be able to make global acoustic-phonetic changes, such as those involving F0 or intensity, to speech, a speaker would presumably need to have a flexible speech motor control system. In spontaneous speech, children have been found to become adult-like by age 11 in their production of speech rate and F0 range, and by age 13 in vocal intensity (Hazan *et al.*, submitted). By age 14, children are still producing higher median F0 than adults, presumably due to physiological differences to adults (Hazan *et al.*, submitted), and articulatory variability may continue to be greater even in adolescents compared to adults (Walsh and Smith, 2002). However, the fact that children do not produce adult-like vocal characteristics or articulatory variability until late childhood or adolescence does not necessarily imply that they will not have any control over these characteristics of their speech until that age – it is, nonetheless, possible that they will not be as precise in these adaptations to a listener's needs as adults.

For enhancing the segmental aspects of speech, such as making segmental contrasts more discriminable, a speaker would need to have specific internal articulatory targets, and a knowledge of the cues which could be useful for contrast enhancement. Additionally, a speaker would need to be able to adequately control their speech production to be able to make the specific cue enhancements. As children are less able to use limited acoustic cues to recognise phonemes, and their perceptual categories are less precise than adults' even until adolescence (e.g., Hazan and Barrett, 2000) (see section 1.2.1 for details), it seems likely that the segmental level may be less flexible in children for making speech production changes. On the other hand, Romeo *et al.* (2013) discovered that, although their within-category variability was still higher than adults', children above the age of 10 years were able to produce as discriminable /s/-/ʃ/ and /p/-/b/ categories as adults. Therefore, in practice, older children may be able to make adequate changes to the segmental level of their speech.

For a speaker to make linguistic changes to their speech according to listener needs, they would need to have an adequate amount of linguistic knowledge, vocabulary and syntax to be able to generate alternative ways of producing their utterances. As evidenced from the review above on the ability of children to make adaptive repairs to their speech, children may reach adequate levels of linguistic experience to make these changes from approximately age 7 onwards (e.g., Brinton *et al.*, 1986), although these skills are likely to develop further with greater linguistic experience with age (e.g., Lloyd *et al.*, 1998). Some linguistic changes may be easier to make than others – for example, speakers may be able to use shorter utterances more easily than changing the difficulty of the vocabulary that they use.

**Evidence for adaptation** Based on the literature reviewed in this section so far, it seems likely that, by the time they reach late primary or early secondary school age, NH children may be able to make at least some adaptations to the needs of a listener. However, only a limited number of studies have examined adaptations made by children, especially in response to a HI peer.

Most studies investigating the interactions between NH and HI children examine either the quality of peer relationships (c.f., Antia *et al.*, 2011; Martin, Bat-Chava, Lalwani, and Waltzman, 2010) or the type and duration of play between young children of differing hearing status (e.g., Antia, 1982). Alternatively, NH-HI pairs' communication is investigated without comparison to the same NH child's communication with a NH partner (e.g., Lederberg, Ryan, and Robbins, 1986; Lloyd, 2003). The little research done on the communicative adaptations made by NH children to HI peers is mostly over 30 years old, and generally investigates few participants, usually of preschool age, and uses mostly subjective discourse measures to assess participants' communication (Arnold and Tremblay, 1979; Seewald and Brackett, 1984; Spencer, Koester, and Meadow-Orlans, 1994; Vandell, Anderson, Ehrhardt, and Wilson, 1982; Vandell and George, 1981). Some of these studies found evidence of modification of some basic communication strategies by even young NH children when in contact with HI peers – for example, 2- to 5-year-old NH children used more gestures and signs, and fewer vocalisations, when playing with their HI peers as compared to their NH peers in integrated preschools (Arnold and Tremblay, 1979; Spencer *et al.*, 1994). Seewald and Brackett's (1984) case study of a 6-year-old NH girl in interaction with an adult, a toddler, a NH peer and a HI peer revealed that the participant used less complex syntax and more directives when talking with the HI peer than with the NH peer, and she shortened her utterances to a similar extent to the HI peer as to the toddler. On the other hand, Vandell and George (1981) found that NH 4-year-olds playing with severe to profound HI peers did not change the modality of their interaction to be more visual, but continued to use vocalisations. Unfortunately, an attempt to train the NH children to use the visual and sensory modalities with HI peers failed – in play with HI peers in a post-test, NH children used the trained strategies *less* than in the pre-test (Vandell *et al.*, 1982).

The studies above offer a very limited view of whether NH children are able to adapt to their HI interlocutor. However, more research has examined whether NH children modify certain linguistic aspects of their speech to children of different ages and/or linguistic levels. Guralnick and Paul-Brown (1986) found that 4- to 6-year-old NH children increased their use of simple syntactic phrases, repetition and gestures when talking with peers with moderate to severe developmental delays, as compared to peers without disabilities. Two- to 3- (Dunn and Kendrick, 1982), and 4- to 5-year olds (Sachs and Devin, 1976; Shatz and Gelman, 1973) use shorter utterances, more repetitious speech and more attentional language with infants and toddlers compared to peers, and both 3- and 5-year-olds used shorter utterances when conversing with a doll compared to an adult (Warren-Leubecker and Bohannon, 1983).

A few more recent studies have investigated age- and language-level -based speech modifications in children from an acoustic-phonetic point of view. Weppelman, Bostow, Schiffer, Elbert-Perez, and Newman (2003) asked 4-year-olds to show both an adult and an infant how a toy works – the children spoke more slowly to the infant than to the adult but, surprisingly, they did not change their F0 variability in infant-directed speech compared to adult-directed speech, even though high F0 variation is a defining characteristic of IDS in adult speech. The authors suggest that 4-year-olds may not be able to change the F0 of their speech in this situation, as they may lack the cognitive flexibility to do so – less cognitive resources may be required for lengthening words. However, results from Syrett and Kawahara (2014) seem to point against this view – 3- to 5-year-old children increased their F0 range, and used longer and more intense vowels and had a larger vowel space when teaching a puppet new words as compared to telling an adult about the same words. By 16 years of age, adolescents increase their mean F0 and F0 range, and increase the duration of certain phonemes, when talking with an infant compared to an adult (Kempe, 2009).

Although most of the above studies seem to suggest that even young children control some of their speech and language production for their listener's needs, none of them directly compared the children's speech modifications to those of adults, making it difficult to evaluate how adult-like children's speech modifications are. Hazan *et al.* (submitted) and Pettinato *et al.* (2016), however, investigated 9- to 14-year olds' peer communication through a normal channel and through an intelligibility-reducing 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. 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, whose vowel spaces were more adult-like, were found to increase the size of their vowel space in the vocoder 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. Interestingly, the results of Hazan *et al.* (submitted) suggest that the ability to make global acoustic-phonetic modifications to speech is not only associated with speech motor control – although their results showed that children's F0 range in the 'normal' speaking condition was adult-like by age 10, all age groups still increased their F0 range in the vocoder condition, despite this strategy being unhelpful to the listener. Similarly, although articulation rate and mean F0 were only adult-like by age 10, and speech intensity was adult-like by age 12, these adjustments were made to a similar extent to adults in the vocoder condition. Even 13- to 14-year-olds' vowel space areas were larger than adults', but they were still able to increase their vowel spaces in the vocoder condition, albeit less so than adults (Pettinato *et al.*, 2016).

There is some evidence that the above modifications made by children are done on the basis of listener feedback – Warren-Leubecker and Bohannon (1983) showed that if a baby doll gave preschoolers feedback signalling non-comprehension, the children decreased MLU in speech directed to the doll, but when given signals of comprehension, the children increased their MLU. However, the children did not modify the length of their utterances according to adult feedback. Similarly, 4-year-olds who were matched with either 'high-verbal' 2-year-olds or with 'low-verbal' 2-year-olds initially modified their MLUs to match that of their interlocutor (Masur, 1978). However, during the interaction, the 4-year-olds' MLU was affected by the amount of feedback given by the 2-year-olds – 4-year-olds whose interlocutors were highly responsive began using more complex utterances to the 2-year-old compared to those whose interlocutors were less responsive. Even preschoolers, therefore, seem to be able to apply certain modifications to their speech based on feedback.

Other evidence suggests that children also gradually begin to be able to take the perspective of an imagined interlocutor – only children over the age of 10 seem to be able to modify the information content of their message to an imagined infant interlocutor (Flavell, Botkin, Fry, Wright, and Jarvis, 1968; Sonnenschein, 1988). Redford and Gildersleeve-Neumann (2009) found that 5-year-olds started to use some acoustic-phonetic modifications to their speech, such as producing more final stop releases, when asked to speak 'clearly', but younger children did not.

Although the studies reviewed above suggest that even preschool-aged children are able to make some adaptations to a listener, it seems clear that adaptation is not yet adult-like even in later childhood. Most of the studies use only very limited measures, such as MLU, to examine adaptations (e.g., Masur, 1978; Shatz and Gelman, 1973; Warren-Leubecker and Bohannon, 1983). Very few studies have examined global acoustic-phonetic strategies used by children (e.g., Hazan *et al.*, submitted; Syrett and Kawahara, 2014; Weppelman *et al.*, 2003), and all but one of them examined strategies used by younger children. The only previous studies to have examined segmental enhancement in children involved either using a 'clear speech' instruction (Redford and Gildersleeve-Neumann, 2009) or have used only limited segmental measures, mostly examining vowel space area (Pettinato *et al.*, 2016; Syrett and Kawahara, 2014). Importantly, no previous study has examined the acoustic-phonetic modifications made by NH children to HI peers. This study will offer a comprehensive investigation of the strategies used by NH children when speaking with their HI friends using measures of linguistic, global acoustic-phonetic and segmental contrast enhancement.

### 1.3.4 Hearing-impaired children

The above sections demonstrate that some research has focused on HI-directed speech and the adaptations made by both NH children and adults. However, very little research has been conducted on whether HI individuals themselves are able to adapt to the needs of their interlocutor.

**Perspective-taking** There is a much less extensive literature on the development of perspective-taking and theory of mind in HI children than in NH children. Of the studies conducted, most demonstrate that, although native signing HI children of HI parents show no delays (e.g., Schick, De Villiers, De Villiers, and Hoffmeister, 2007), HI children of NH parents are significantly delayed in theory of mind development compared to NH peers. Research using false-belief tasks on prelingual severe to profound hearing-aid users using oral and total communication as their preferred communication mode show that 7-year-old HI children perform significantly worse than 5-year-old NH children (Schick *et al.*, 2007), 9- to 12-year old HI children obtain similar results to 6-year-old NH children (Levrez, Bourdin, Le Driant, D'Arc, and Vandromme, 2012), and that even 13-year-old HI children achieve lower scores than 3- to 5-year-old NH children (Russell *et al.*, 1998), even when non-verbal or low-verbal tasks are used (Levrez *et al.*, 2012; Schick *et al.*, 2007). Cochlear-implanted children of NH parents, using a range of communication modes, were found to pass the false-belief test between 7 and 9 years of age (Lundy, 2002; Macaulay and Ford, 2006; Moeller and Schick, 2006), or after 9 years of age (Peterson, 2004), demonstrating a 3- to 6-year delay compared to NH children (e.g., Moeller and Schick, 2006).

The few studies which have shown child CI users to develop theory of mind abilities on par with NH peers (Peters, Remmel, and Richards, 2009; Remmel and Peters, 2009; Ziv, Most, and Cohen, 2013) used groups of participants who, on average, had developed age-appropriate language skills, suggesting that language development is vital for children's understanding of others' mental states. This finding is supported by research showing that HI children's expressive language skills (Lundy, 2002), and lexical (Levrez *et al.*, 2012; Schick *et al.*, 2007; Sundqvist, Lyxell, Jönsson, and Heimann, 2014) and syntactic knowledge (Levrez *et al.*, 2012; Schick *et al.*, 2007) are correlated with their theory of mind abilities.

It is likely that the considerable delay exhibited by HI children is also due to their less effective early caregiver interactions compared to NH children. Indeed, there is some evidence that early implantation, and therefore earlier access to language, may improve some theory of mind outcomes for CI children – Sundqvist *et al.* (2014) investigated 16 prelingually deafened, oral 4- to 10-year-old cochlear-implanted children of NH parents, of whom half were implanted early (on average, at 1.4 years of age) and the rest were implanted late (at approximately 3.3 years of age). Although the two CI groups did not differ from each other on language level, the late-implanted group, but not the early-implanted group, differed from age-matched NH peers on a task assessing 'emotional' theory of mind, i.e. understanding of others' emotions, their causes and consequences. However, on a 'cognitive', standard false-belief task, both CI groups lagged behind NH peers. Nonetheless, the CI group demonstrated a wide range of variability in their theory of mind ability, with 40% of CI children performing within NH range.

The delayed theory of mind abilities of HI children may also be due to poorer exposure to mental state words (beliefs, desires and emotions) by family and friends – Moeller and Schick (2006) found that the theory of mind abilities of 4- to 10-year-old prelingually deafened severe to profound HI children of NH parents was predicted by the mother's use of mental state vocabulary in conversation with the HI child. Using qualitative methods, Silvestre, Ramspott, and Pareto (2007) also discovered that Spanish mainstreamed prelingually deafened moderate to profound 6- to 18-year-old HI children of NH parents, who mostly used oral communication, were exposed to very little conversation about mental states by teachers and peers.

These findings demonstrate that HI children who have delayed speech and language are likely to show deficits in theory of mind abilities even into middle and late childhood years – suggesting that it may be difficult for even older HI children to understand their listener's perspective in an interaction. However, most of the studies reviewed above were performed only on children with severe and profound hearing loss, and therefore it is unclear whether milder hearing losses also result in delays in theory of mind skills. Additionally, the false-belief task used by most of the studies is fairly artificial, and perhaps unnecessarily complex for HI children – Marschark, Green, Hindmarsh, and Walker (2000) have shown that theory of mind concepts were present in the storytelling of 7- to 15-year old severe to profound HI children of NH parents using total communication to an equivalent extent to NH children of the same age. Use of more naturalistic experiments (Ziv *et al.*, 2013), or tasks in which 'individual' belief is assessed, may demonstrate more advanced theory of mind skills in HI children. HI children's theory of mind abilities may also be related to their executive functioning (Ziv *et al.*, 2013) and working memory skills, which are known to be delayed in both preschool children with CIs (Beer, Kronenberger, Castellanos, Colson, Henning, and Pisoni, 2014) as well as in some older children with CIs (e.g., Beer, Pisoni, Kronenberger, and Geers, 2010; Fagan, Pisoni, Horn, and Dillon, 2007). As mentioned in section 1.3.3, these skills may be important for further theory of mind development in older children (Dumontheil *et al.*, 2010).

**Detecting and responding to feedback** The section above demonstrated that even older HI children may be delayed in their understanding of others' perspectives, which may make taking the listener's needs into account more difficult for them in an interaction. Due to their likely reduced experience in social interaction in general, as discussed in the previous section, it also seems probable that HI children will display deficits in inferring and adequately responding to a listener's needs from listener feedback. Research has indeed shown that HI children are delayed in their general pragmatic skills when compared to NH children (e.g., Dammeyer, 2012; Goberis, Beams, Dalpes, Abrisch, Baca, and Yoshinaga-Itano, 2012; Most, Shina-August, and Meilijson, 2010; Tye-Murray, 2003). On the other hand, as HI children are likely to encounter communication breakdown frequently (Caissie and Wilson, 1995; Tye-Murray, 2003), they may be more experienced than typical NH children in this aspect, and therefore may be able to respond more adaptively to listener feedback.

A few studies have demonstrated a delay in HI children's abilities to respond adaptively to listener feedback compared to NH children. In Jeanes, Nienhuys, and Rickards's (2000) study, 20 oral profoundly deaf HI children and adolescents between 8 and 17 years of age, who had NH parents and who were educated in mainstream schools, took part in a referential communication task with a friend of the same age and hearing status. Twenty NH peer pairs acted as controls. Although both groups of participants were found to give appropriate responses to clarification requests over approximately 90% of the time, the HI participants used significantly more repetitions in response to clarification requests compared to NH participants, who in turn used more adaptive responses that specifically added new information to the interaction. Similarly, Most (2002) investigated 16 11- to 18-year-old HI participants, all profoundly deafened and wearing digital hearing aids, who were judged to have age-appropriate expressive language skills. Half of the participants were rated as having high speech intelligibility, and half were rated as poorly intelligible. Ten age-matched NH participants acted as controls. As in Brinton *et al.* (1986), participants described pictures to an investigator, who at certain time points initiated a stacked three-part non-specific clarification request sequence. The group of HI participants with poor speech intelligibility were found to use repetition as a response strategy more often than both the high-intelligibility HI group and the NH group. Most (2002) suggested that this difference between the high- and low-intelligibility groups may have occurred because the participants with lower intelligibility tried to articulate more clearly in their repair response. The NH group, on the other hand, used certain types of adaptive responses to a greater extent than either of the HI groups, demonstrating that even these older HI children with age-appropriate language skills may not be adept at using adaptive strategies when faced with communication breakdown.

However, research on younger age groups have come to different conclusions. In a similar task protocol to that used by Most (2002), but investigating eight 4- to 7-year-old HI children who had profound losses and wore digital hearing aids and used total communication, Ciocci and Baran (1998) found that the HI participants used *more* adaptive responses (revision strategies) than age-matched

NH controls. The researchers reported that the HI children made as much effort as they could to be understood by the investigator, and particularly used gesture, pantomime and clearer speech articulation in their repair responses. Using a similar method, but without NH controls, Blaylock, Scudder, and Wynne (1995) showed that severe to profound 4- to 9-year-old HI children, most of whom used digital hearing aids and total communication, demonstrated a wide range of adaptive repair strategies, including adding gestures and orally or manually elaborating their original utterances after non-specific clarification requests. Additionally, more recent studies by both Sandgren, Ibertsson, Andersson, Hansson, and Sahlén (2010) and Most *et al.* (2010) indicate that HI children's repair behaviour may be similar to that of NH children's. Interestingly, both these studies included cochlear-implanted participants – in Sandgren *et al.*'s (2010) study, all 13 HI adolescents wore CIs, while in Most *et al.*'s (2010) research using a pragmatic skills rating scale, 11 of the 24 6- to 9-year-old children had CIs, and the rest had severe hearing loss and were digital hearing aid users. It may be, therefore, that with the greater access to language provided by modern advanced amplification devices such as CIs, HI children's skills at adaptively responding to a listener's feedback are no longer delayed.

On the other hand, HI children may not be as adept as NH children at inferring listener needs on referential communication tasks without listener feedback to guide them. Lloyd, Lieven, and Arnold (2005), using a similar referential communication task to Lloyd *et al.* (1998), investigated the descriptive referential speaking skills of 20 HI children between the ages of 7 and 12. The participants had severe and profound losses, used both oral and total communication, and were approximately 3.5 years delayed in their receptive language skills. Only a minority (3) had CIs. Compared to a group of younger 5- to 7-year-old NH children, HI children were similarly accurate in being able to provide the listener with unambiguous picture descriptions, suggesting a developmental lag in HI children on this skill. Elfenbein *et al.* (1994) also noted that the most frequent pragmatic error displayed by forty oral, mild to severe 5- to 18-year-old HI children when being interviewed by an investigator was in providing ambiguous information to the listener – even a younger NH control group did so less than the older HI participants.

There is also some evidence that HI children are delayed in monitoring their own comprehension. Goberis *et al.* (2012) noted that, on a pragmatic skills checklist filled in by parents of 126 3- to 7-year-old mild to profound HI children, most 7-year-old HI children were not reported to have mastered clarification requests, although the majority of NH control 4-year-olds had done so. Lloyd *et al.* (2005) also investigated the 'listening' skills of the HI children in their study. As in Lloyd *et al.* (1998), the investigator provided the participants with either ambiguous or unambiguous messages in a referential communication task. The number of times during which the HI children made clarification requests to ambiguous information was similar to that of 3.5-years younger NH control children, again suggesting a developmental delay in monitoring their own comprehension. As significant correlations to the HI children's receptive language performance was found, the investigators suggest that the participants' delayed comprehension monitoring skills were related to their delayed language skills. Alternatively, HI children may have learned to use 'learned helplessness' (Arnold, Palmer, and Lloyd, 1999), and tended to attribute any failures in understanding to themselves, rather than their interlocutor.

In summary, findings on HI children's abilities to detect and respond to feedback and to infer listener needs are mixed, with studies showing both similarities and differences between HI children and NH controls. Unfortunately, due to the distinct lack of studies investigating this aspect of HI children's pragmatic skills, especially in the last 10 years, it is very difficult to draw any distinct conclusions from the literature. However, it seems likely that HI children who exhibit language delays relative to NH peers will also have more difficulty in responding adaptively to listener feedback in both linguistic and acoustic aspects – if only because their delayed language and/or speech production skills may be less flexible in allowing them to adapt their original message to the listener. The next section will explore this possibility further.

**Flexibility in speech and language** As well as being able to model the listener, and being aware of the modifications which may be useful for the listener, the HI child also needs to have flexible mechanisms for changing their speech and language output to listener needs. For HI children, this may be particularly difficult due to their possible delays in speech and language production, which is characterised by (1) speech motor control problems (e.g., Niparko, 2009), leading to greater variability in speech voice characteristics compared to NH children, (2) differences in internal phonological structure and difficulty in the production of fine phonetic cues (e.g., Kosky and Boothroyd, 2003), which may impact their ability to produce distinctive phonetic constrasts, and (3) delayed expressive language skills, both in vocabulary and syntax (e.g., Wake *et al.*, 2004) (for more detail, see section 1.2.2).

As discussed in section 1.3.3, a flexible speech motor control system would presumably be required for a speaker to make global acoustic-phonetic adaptations, such as changing their mean F0, F0 range or vocal intensity. Compared to NH children, these may be more difficult for HI children to make, due to their possible problems with speech motor control – they have been found to exhibit higher mean F0 (e.g., Higgins *et al.*, 2003), greater F0 variability (e.g., Allen and Arndorfer, 2000) and, perhaps, more variable vocal intensity (Campisi *et al.*, 2005) than NH peers. However, as discussed in relation to NH children in section 1.3.3, this does not necessarily imply that HI children are not able to make such adaptations, but they may not be as adept at making them as NH peers are.

To enhance the segmental aspects of speech, HI children would need to have adequate internal articulatory targets to approximate, as well as knowledge of the particular phonetic cues which distinguish phonetic contrasts. Additionally, they would need to have good enough speech production skills to make the required fine phonetic detail adaptations. These kinds of phonetic adaptations may be much more difficult for HI children to make than global adaptations, due to HI children likely having less precise phonetic categories, both from a perception (e.g., Medina and Serniclaes, 2009) and production (e.g., Uchanski and Geers, 2003) point of view. It is likely that HI children with more profound losses will find segmental adaptation particularly difficult, due to their own severe losses obstructing important phonetic cues.

HI children's delays in their expressive language skills are also likely to make linguistic adaptations to a listener more difficult. For example, their limited vocabulary and syntax compared to NH peers may limit the amount of alternatives that they can use to adapt to their listener. The review on HI children's abilities to use adaptive repair in response to feedback showed that it is unclear whether HI children have adequate language skills to do this. It seems likely that HI children who are assessed as having a delay in speech and language compared to NH peers would find this aspect of adaptation especially difficult.

**Evidence for adaptation** From the evidence reviewed in this section so far, it seems likely that HI children will have difficulty in adapting to their listener's needs. However, hardly any research has been done to investigate this issue.

Very few studies have investigated whether HI children are able to adapt to another HI listener's needs. Severe to profound 2- to 5-year-olds using speech and sign were reported to try to communicate with both NH and HI peers as best they could, and to be able to respond to NH and HI peers according to the communication mode of the partner (Rodríguez and Lana, 1996). Lederberg *et al.* (1986) found similar results when investigating 4- to 6-year old profound HI children playing with NH and HI peers. Other studies (Arnold and Tremblay, 1979; Spencer *et al.*, 1994) found no differences between NH-directed and HI-directed communication modes used by HI children aged 2- to 5-years – HI children mostly used sign to communicate, regardless of the hearing status of the peer. These findings are not only old, and therefore reflect a very different environment for HI children in terms of amplification technology and education, but also limited in terms of the number of participants used in the studies (ranging from 4 in Spencer *et al.* (1994) to 14 in Lederberg *et al.* (1986)), the age group of the participants (all preschool aged), and the types of communication strategies investigated (mostly in terms of communication mode used by the participants).

There is some evidence, however, that HI adults are able to speak more clearly when instructed to do so - Ménard, Polak, Denny, Burton, Lane, Matthies, Marrone, Perkell, Tiede, and Vick (2007) asked NH adults and postlingually deafened cochlear implant recipients pre-implant, 1-month post-implant and 1-year post-implant to read sentences 'carefully without increasing loudness' to elicit a clear speaking style and 'at a conversational rate' to elicit a casual speaking style. The changes made between clear and casual speech were similar in NH adults and the CI users at all time points; both groups of adults produced larger vowel contrasts and longer vowel durations in clear than in casual speech. This suggests that at least hearing-impairments occurring in adulthood do not prevent people from adapting to a listener. However, adults with acquired hearing-impairments likely had fully-mature internal phoneme categories by the time of their hearing loss, and therefore this may make it easier for them to approximate targets and hyperarticulate in clear speech, than for HI children.

Other studies show that some of the difficulties faced by HI children may nonetheless not impact their ability to adapt to a listener's needs. Autistic children and adolescents aged between 6 and 16 years of age, who also exhibit delays in theory of mind development, are reported to be able to simplify their language when asked to speak to an imagined baby or non-native speaker, as opposed to a peer or adult (Volden *et al.*, 2006), although their simplification of language to these two groups was significantly less than that of controls'. When autistic participants were specifically asked to pay attention to their interlocutor's characteristics and their own language use, the autistic participants were able to further simplify their language to the listener. These results suggest that even a deficit in theory of mind abilities (as found in both autistic individuals and HI children) may not be a barrier for making adjustments to listener needs, but may lead to speakers making fewer spontaneous modifications to listeners than typical NH peers do. Additionally, Guralnick and Paul-Brown (1989) demonstrated that preschool children with mild developmental delays produced shorter MLUs to peers with severe to profound developmental delays, similarly to typical NH peers – suggesting that some delays in speech and language may not affect being able to take your listener's needs into account.

Even older speakers who have speech motor control problems due to Parkinson's disease or Multiple Sclerosis have been found to be able to change some of the same aspects of their speech acoustics as typical adults, if asked to 'speak clearly', such as decreasing articulation rate, increasing intensity, increasing F0 range and mean F0, and in some cases, making greater vowel formant movements (e.g. Goberman and Elmer, 2005; Tjaden, Lam, and Wilding, 2013; Tjaden, Sussman, and Wilding, 2014; Tjaden and Wilding, 2004; Whitfield and Goberman, 2014), suggesting that even speech motor control problems may not preclude individuals from being able to produce some characteristics of clear speech, at least if instructed to do so. However, the type of speech motor control problems exhibited by dysarthric speakers and HI children are likely to be quite different.

## 1.4 The current study

The current study is novel in assessing 'real' HI-directed speech by examining the speech adaptations made by older NH and HI children to a HI peer, compared to when they are talking to a NH peer. We use two communicative problem-solving tasks, the Grid task and the Diapix task, which are similar to tasks that NH and HI children may regularly face at school, to elicit NH- and HI-directed speech in a controlled but naturalistic setting. The 'Grid' task was developed to obtain multiple repetitions of three different kinds of speech contrasts in minimal pair keywords in spontaneous speech to enable a thorough investigation of segmental contrast enhancement in HI-directed speech, while the more complex spot-the-difference Diapix task was used to enable an analysis to be made of whether task difficulty may affect the amount of adaptation made to a HI interlocutor (discussed in greater detail in chapter 2). These tasks will enable the elicitation of several different types of speech measures – global acoustic-phonetic measures of speech, between- and within-speaker variability in several segmental contrasts, as well as a wide range of vocabulary and syntactic measures – to enable us to explore whether these aspects of NH and HI children's speech are enhanced in speech directed to a HI peer, compared to that elicited when speaking to a NH peer. As any communication difficulties may arise either due to a production problem by the speaker or due to a perception problem by the listener, the speech perception skills of the NH and HI children are also assessed. Additionally, the perceived clarity of the participants' speech is explored in each condition, and is related to their acoustic-phonetic enhancement strategies.

The next chapter (chapter 2) discusses the Grid and Diapix tasks and describes the development of the Grid task. The following chapter (chapter 3), details the participants and the methods used in the current study. Chapters 4 and 5 investigate whether the NH and HI participants differ in their speech perception skills and in their ability to complete the two communicative problem-solving tasks. Then, chapters 6, 7 and 8 explore the acoustic-phonetic and linguistic aspects of adaptations made by both NH and HI speakers. Chapter 9 investigates whether speakers' speech clarity varies according to the interlocutor's hearing status, and chapter 10 considers whether the variability in adaptations made is related to speaker and listener characteristics. Finally, in chapter 11, the findings of the study are summarised and discussed.

The following research questions are investigated in the current study:

**RQ1 What are the differences in the speech produced between NH and HI children in global acoustic-phonetic measures, segmental measures, and linguistic measures in spontaneous conversation? (chapters 6, 7 and 8)**

The literature review on the speech production skills of NH and HI children in section 1.2.2 demonstrated that HI children are likely to have difficulty in speech motor control, slower linguistic processing and differences in internal phonological structure, as well as delayed expressive language skills, compared to NH children. We would expect, therefore, the HI participants in this study to show different speech characteristics to NH children on each of these speech production measures.

**RQ2 Do NH and HI children adapt to the needs of an interlocutor with hearing loss? –**

**(A) Are adaptations made on the global acoustic-phonetic level, the segmental level and the linguistic level? (chapters 6, 7 and 8)**

As NH children's speech production skills are still developing, but starting to converge to adult levels by late childhood (see 1.2.2), we would expect that NH children will make similar modifications to those found in adults in the previous literature, but that these modifications may be made to a lesser extent compared to adults. It is of particular interest whether NH children enhance segmental contrasts, as there is conflicting evidence of segmental contrast enhancement in adults' speech adaptations (as discussed in section 1.3.2).

As HI children exhibit reduced speech motor control and greater variability in global acoustic-phonetic measures, which nonetheless does not necessarily preclude them from making adaptations on these measures, we would expect that HI children would make adaptations on this level, but to a lesser extent compared to

NH children. As linguistic adaptations require a certain level of linguistic competence, linguistic adaptations may be more difficult for HI children to make, at least for participants who are delayed in their language skills relative to NH children. Segmental adaptations are likely to be the most difficult for HI children to make, as, to enhance segmental distinctions, participants would need to have consistent articulatory targets for phonemes, acquired knowledge of the particular cues required to enhance phonetic contrasts, and also obtained adequate speech motor control.

**(B) Does the extent of speech adaptations made depend on task difficulty? (chapters 6, 7 and 8)**

As discussed in section 1.3.2, according to the Dual-Process Model (Bard and Aylett, 2005), it is likely that more difficult tasks will affect the amount of cognitive resources which can be allocated to adapting to the interlocutor. Therefore, it is hypothesised that fewer speech adaptations will be made in the more difficult Diapix task than in the relatively simpler Grid task.

**(C) Do the acoustic-phonetic adaptations made to listener needs make the speakers perceptively clearer in HI-directed than NH-directed conditions? Which acoustic-phonetic measures will predict speakers' perceived speech clarity? (chapter 9)**

We would expect that, as has been found for clear speech studies (section 1.3.2), HI-directed speech will be perceived as clearer than NH-directed speech, at least for the NH speakers. As discussed in section 1.3.1, increased speech clarity is likely to be associated with greater speech intensity, greater F0 variability, a decreased speech rate, an increased number of pauses, a bigger vowel space, and more discriminable phonetic categories.

**(D) Is individual variability between participants in their speech adaptations related to individual listener and speaker characteristics? (chapter 10) –**

*(i) Do NH and HI children make adaptations according to individual listener needs, such as the interlocutor's speech and language level?*

As the review of adult speakers' adaptations to the needs of listeners in section 1.3.2 demonstrated, adults are adept at adapting to particular listener needs. We would expect that NH children would also be able to make changes to their speech according to individual listener needs, although possibly to a lesser extent than adults. Therefore we would expect that NH children completing the tasks with HI children who are more severely affected by their hearing loss would also make more adaptations to that listener, compared to those conversing with HI children with milder hearing losses. HI children, on the other hand, may be unlikely to be able to make such fine-grained adaptations according to the specific needs of a listener.

*(ii) Does the extent of adaptations made by HI children depend on the HI speaker's characteristics, such as their speech and language level?*

As discussed in section 1.3.4, those HI children whose speech and language is more affected by their hearing loss may be less able to make adaptations to their HI interlocutor, especially on linguistic and segmental levels which require more advanced linguistic knowledge.

# Chapter 2

# Development of a communicative spontaneous speech task

## 2.1 Introduction

In this study, two different referential communication tasks are used to elicit speech – the DiapixUK task (Baker and Hazan, 2011) and the Grid task. Using these two tasks enables analyses to be made of NH and HI children's global acoustic-phonetic, segmental and linguistic measures of speech in both NH- and HI-directed conditions, and allows an additional assessment to be done on whether task difficulty affects the extent to which participants adapt to a HI interlocutor. This chapter describes the background, aims and implementation of the Grid task, which was developed in the current study specifically to elicit frequent repetitions of segmental contrasts in NH and HI children's peer-to-peer interactive communication. To our knowledge, no other referential communication task has been developed for HI children that allows the elicitation of all three types of measures in a communicative peer interaction paradigm.

### 2.1.1 Previously used methods of speech elicitation

Most previous studies of speech production in HI children use non-communicative tasks for speech elicitation. This is particularly the case if the experiment focuses on the production of phonemes – most commonly, either isolated sounds, syllables or words are elicited, typically through either imitation of the experimenter, through picture-naming, or through reading separate words in a list (Bernhardt *et al.*, 2003; Chuang *et al.*, 2012; Eriks-Brophy *et al.*, 2013; Horga and Liker, 2006; Liker *et al.*, 2007; Löfqvist *et al.*, 2010; Metz *et al.*, 1985; Monsen, 1974; Vandam *et al.*, 2011). Alternatively, isolated sentences are produced through reading or imitation (Chuang *et al.*, 2012; Harris *et al.*, 1985; Horga and Liker, 2006; Kosky and Boothroyd, 2003; Metz *et al.*, 1985; Neumeyer *et al.*, 2010; Uchanski and Geers, 2003). While the elicitation of specific segments using these tasks ensures a high 'event density' (i.e., the relative number of speech sounds of interest produced within the time frame is high) (Niebuhr and Michaud, 2015) and the context of the speech sounds is highly controlled, the speech elicited is not communicative in nature, and is therefore more likely to assess the performance rather than competence of speakers . Additionally, despite their elicitation method producing high event density, most of the studies cited above concentrate merely on assessing the % correct of phonemes produced (e.g., Bernhardt *et al.*, 2003; Eriks-Brophy *et al.*, 2013; Kosky and Boothroyd, 2003). Those that do analyse the acoustic-phonetic aspects of speech sounds typically elicit only a few tokens of each phoneme (e.g., Horga and Liker, 2006; Löfqvist *et al.*, 2010), therefore preventing a thorough analysis of both between- and within-category variability being conducted on the data. A few of these studies do include a greater number of token elicitations (e.g., Chuang *et al.*, 2012; Monsen, 1974; Neumeyer *et al.*, 2010), but nevertheless do not analyse within-category variability [1]. As discussed in chapter 1, clear speech studies are similarly often only concerned with between-category measures, rather than with within-category variability – despite recent evidence of the perceptual importance of within-category variability in talker intelligibility (e.g., Hazan *et al.*, 2013; Newman *et al.*, 2001).

The elicitation of spontaneous dialogic speech enables the analysis of more natural, expressive, and communicative competence-based speech (Niebuhr and Michaud, 2015) and, as discussed in chapter 1, is essential for assessing a speaker's adaptation to listener needs. A few previous studies on HI children's speech production elicit speech through conversation with either an examiner, clinician

[1]Although, for studies which assess phonetic variability in HI children to a limited extent, see (Metz *et al.*, 1985; Uchanski and Geers, 2003)

or a parent (Blamey *et al.*, 2001; Lenden and Flipsen, 2007; Nicholas and Geers, 2006; Serry and Blamey, 1999; Snow and Ertmer, 2009; Wiggin *et al.*, 2013), but these conversations are likely to be very different to peer interaction – and these studies mostly assess either the participants' speech intelligibility, the % correct of phonemes, or only the participants' expressive language skills using measures such as MLU. Indeed, in spontaneous conversational dialogues, researchers have much less control over the context of the speech elicited, potentially leading to a great amount of variability in measures due to factors other than those of interest (Ito and Speer, 2006). The event density is also usually low (Niebuhr and Michaud, 2015), which prevents the reliable within-category analysis of specific minimal pair contrasts in speech.

### 2.1.2 Referential communication tasks

Referential communication tasks can be used as an alternative to spontaneous conversation in the elicitation of speech. In a referential communication task, interlocutors are typically separated by a barrier, and they are given the task of identifying, describing or locating certain referents to each other. Thus, to enable the successful completion of the task, information must be exchanged between interlocutors, and both speaking and listening skills must be used in context by each participant, as in real communicative situations (Leinonen and Letts, 1997; Lloyd, 2003; Markman and Makin, 1998; Yule, 1997). Referential communication tasks enable the researcher to retain some control over the speech produced, as the context and referents produced in the talk will be constrained by the task. These tasks also typically enable a higher event density than spontaneous conversation, as the referents in the task can be controlled by the experimenter to include elements of interest. Critically for the current study, the researcher is aware of what each participant is trying to communicate, which is especially important when dealing with speakers who are less intelligible or who may be unable to produce certain phonemes accurately. The researcher also knows what the aim of the interaction is – trying to transmit information accurately to the interlocutor – thus enabling a measure of communicative success to be used to characterise the interactions. Although referential communication tasks are unlikely to elicit as natural speech as spontaneous conversation, the speech produced is still spontaneous and communicative. Children are likely to encounter referential communication tasks in their everyday lives, for example in the classroom or in games played outside of school (Lloyd, 2003).

As seen in chapter 1, referential communication tasks have been used previously to assess HI children's communication skills, but the measures examined have been limited to either the type of repair strategies or clarification requests used (e.g., Ibertsson, Hansson, Mäki Torkko, Willstedt Svensson, and Sahlén, 2009; Jeanes *et al.*, 2000; Most, 2002; Sandgren *et al.*, 2010), or the production and detection of ambiguity by participants (Arnold *et al.*, 1999; Lloyd *et al.*, 2005), with no known studies examining the acoustic-phonetic properties of speech produced by HI children in these tasks. In other research areas, referential communication tasks are often used to assess a speaker's ability to prosodically disambiguate utterances (Ito and Speer, 2006; Niebuhr and Michaud, 2015).

Many referential communication tasks used in studies with children involve picture-array tasks, in which one of the participants (the 'follower') is given an array of simple pictures differing in critical attributes, and the interlocutor (the 'leader') is given one of the pictures to describe to the other. The follower then has to select the correct picture based on the leader's description. This is the type of task used in, for example, the studies by Lloyd *et al.* (2005) and Lloyd *et al.* (1998) (see chapter 1). The Map Task (Anderson, Bader, Bard, Boyle, Doherty, Garrod, Isard, Kowtko, McAllister, Miller, Sotillo, Thompson, and Weinert, 1991; Brown, Anderson, Yule, and Shillcock, 1983) was created as a more complex alternative to picture-array tasks. The task leader is given a map with a route, and the task follower must draw the same route on their map based on the task leader's description. Both the task leader and follower's maps have 'landmarks' on them, which contain certain features of interest to the researcher – however, only some of these landmarks are shared by both interlocutors. The interlocutors need to navigate their route through the landmarks to achieve success in the task. However, in both the picture-array tasks and the Map Task, interlocutors are given particular pre-defined roles of leader and follower in the conversation, and therefore, unlike natural speech, the elicited speech is likely to be less balanced between interlocutors and, in the Map Task, may mostly include instructions and commands (Baker and Hazan, 2011).

One of the referential communication tasks used in the current study is the Diapix task, which was originally created by Van Engen, Baese-Berk, Baker, Choi, Kim, and Bradlow (2010), and which was further developed by Baker and Hazan (2011) to enable the elicitation of spontaneous speech in a context in which both interlocutors can contribute equally and collaboratively to the conversation. In the task, a pair of interlocutors is given a different version of a picture-scene, and their aim is to find 12 differences between their pictures without seeing each others' pictures. Because the picture-scenes are detailed, and contain small differences between pictures, the Diapix task elicits natural and fairly complex language from participants. The DiapixUK pictures contain a full set of 12 different picture-pairs which are of equal difficulty, and which therefore allow participants to complete several pictures both between and across conditions. The Diapix task has been used successfully in many recent studies, especially those investigating the global acoustic-phonetic properties of speech in adults (Hazan and Baker, 2011), children (Hazan *et al.*, submitted; Pettinato *et al.*, 2016), and second-language speakers (Granlund *et al.*, 2012; Wester, García Lecumberri, and Cooke, 2014). The Diapix task is used in the current study due to its elicitation of fairly sophisticated syntax and vocabulary, and due to its relative difficulty as a task – the interlocutors are not given specific strategies or roles within the interaction, but must negotiate these between themselves. However, it could not be used to elicit segmental contrasts in speech – although the task was designed to elicit minimal pair keywords, which were included as objects in the pictures, Baker and Hazan (2011) demonstrated that the task did not reliably elicit sufficient numbers of these keywords for segmental contrast analysis.

Therefore, because one of the main aims of the current study is the elicitation of several different types of measures from NH and HI children's peer interactions, including several repetitions of segmental contrasts to enable between- and within-category analyses to be done, an additional task was needed which would elicit frequent repetitions of minimal pair keywords in interaction. A few previous referential communication tasks have been developed for similar purposes. Sankowska *et al.* (2011) used a variant of the Map Task to elicit intrinsically long and short vowels in ADS, FDS and in noise. The target sounds were included as street names that participants needed to use to find the correct route on the map. In Garnier *et al.*'s (2010) river task, both interlocutors were given a sheet of paper with 17 river names, which included target syllables. The task leader had to tell the interlocutor to connect each river to two others to make a route – the task follower then drew the correct route according to the leader's directions. However, although the task elicited communicative speech from the task leader, it did not require much interaction. The elicited speech was also highly constrained, therefore preventing a linguistic analysis from being done on the data.

The referential communication task developed in the current study was inspired by the 'SAME/TRAP' task created by Hazan and Kim (2013). The aim of their task was to elicit consonant stop voicing (/b/-/p/) and place of articulation (/b/-/v/-/d/) contrasts, as well as four point vowels, in an interactive task. In their task, each participant was given a 4-by-4 grid, with a coloured object and a letter in each square. Some of the letters corresponded to the target consonants, and the object in the square was an item beginning with the same letter as that in the square. The colours of the letters, on the other hand, were designed to elicit each of the point vowels. Participants' squares were not identical, and the pair of participants were to identify the squares which were identical ('SAME'), and those in which the letter was the same but the object differed ('TRAP'), without directly naming the object in the square. Therefore participants produced sentences such as "Is your green V something you drive?" (Hazan and Kim, 2013, p.2). Although the task elicited fairly natural speech, and was successful in eliciting multiple repetitions of the target consonants and vowels, the nature of the target consonant elicitation prevents any minimal pair contrasts being used which occur elsewhere than in the syllable onset. The task may also be too complicated for HI children with language delays, as fairly sophisticated language skills are required to describe objects without naming them.

## 2.2 Development of the Grid task

### 2.2.1 Aims

Therefore, the Grid task was developed

1. to elicit multiple repetitions of several different types of phonetic contrasts produced within real words, which HI children may find difficult to produce and perceive. These may lead to potential miscommunications which therefore may require the HI child's interlocutor to enhance those phonetic contrasts in speech.

2. to enable an assessment of the influence of task difficulty on adaptation measures, by creating an interactionally simpler and easier task as an alternative to the Diapix task.

Thus, although one of the main aims of the Grid task was the elicitation of segmental contrasts, the task needed to be variable enough that both global acoustic-phonetic measures and linguistic measures could be taken from the speech in interaction, and compared to that produced in the Diapix task. The task also needed to be visually attractive and enjoyable enough for 9- to 15-year-old children and adolescents to be motivated to play the game several times in a session.

### 2.2.2 Implementation

**Segmental contrasts** The task was designed to elicit productions of three segmental contrasts: the bilabial voicing contrast /p/-/b/, the sibilant place distinction /s/-/ʃ/, and the high front vowel contrast /i/-/ɪ/. These contrasts were chosen as they have typically been found to be difficult for HI children to produce and perceive.

As discussed in sections 1.2.1 and 1.2.2, although temporal processing is often not affected in HI children (e.g., Halliday and Moore, 2010), the primarily temporal /p/-/b/ voicing contrast is likely to be difficult to perceive for HI children with more severe hearing losses, perhaps due to the cue being of fairly low amplitude, and visual cues being unavailable for the perception of this distinction (e.g., Kishon-Rabin *et al.*, 2002). CIs, however, likely assist children with profound losses in developing more accurate perception of voicing categories (Kishon-Rabin *et al.*, 2002). HI children may also produce less contrastive voicing categories than NH children (c.f., Lane *et al.*, 1994), although some CI users may be within NH range (Uchanski and Geers, 2003).

Most HI children have reduced spectral resolution, and therefore, due to the high frequency and lower amplitude of /s/, the spectral /s/-/ʃ/ contrast may be difficult for HI children with even mild losses to perceive, especially as HAs often do not sufficiently amplify higher frequencies (e.g., Stelmachowicz *et al.*, 2002, 2004) – accordingly, many HI children, including those with CIs, are unable to produce the distinction accurately (e.g., Elfenbein *et al.*, 1994; Kosky and Boothroyd, 2003; Uchanski and Geers, 2003).

The spectro-temporal /i/-/ɪ/ distinction has been less studied than consonant contrasts in HI populations, but it seems likely that HI children with milder losses do not have trouble perceiving or producing the contrast (e.g., Eisenberg, 2007). Due to their intact temporal processing, the secondary cue to the contrast, duration, is likely to be perceptible for HI children wearing HAs and CIs. Indeed, Monsen's (1974) study on severe-to-profound HI adolescents' production of the /i/-/ɪ/ distinction showed that HI children produced greater durational distinctions between the two vowels than their NH peers. However, the decreased spectral resolution of HI children with more severe losses, and those with CIs, is likely to affect their perception of fine-grained vowel contrasts (e.g., Giezen *et al.*, 2010) – and therefore it is likely that they are less able to produce spectral distinctions between these two vowels (see section 1.2.2).

HI children's productions of these three contrasts are also likely to be affected by the precise articulatory control needed to produce them, as well as their possibly differing phonological systems compared to NH peers, which may lead them to be more variable in producing these contrasts. Therefore, due to the potential difficulty that HI children may have in perceiving and producing the distinctions, encountering these contrasts will likely lead to frequent misunderstandings between interlocutors, thus necessitating the enhancement of the contrasts by both NH and HI children in HI-directed speech compared to NH-directed speech.

These distinctions also represent different kinds of contrasts (temporal, spectral and temporal- spectral) which may be enhanced differently in HI-directed speech. For example, HI children may have difficulty in enhancing spectral contrasts, but may be more able to make greater enhancements between temporal distinctions.

**Minimal pair keywords** Sixteen minimal pair keywords containing these contrasts were created, three per contrast (see Table 2.1). The keywords were chosen to be common vocabulary items which could be represented pictorially as concrete objects, and which both HI and NH children of 9 years of age would know. However, this was not possible with all the keywords – initial piloting with NH and HI 10-year-olds showed that some children were not familiar with the keywords 'cell' and 'shack'. Nonetheless, after a simple explanation of the word along with its picture, the children seemed to understand the words and readily used them in the task.

| **/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 2.1: The minimal pair keywords used in the Grid task.

Five versions of each of the keywords were hand-drawn, scanned on to a PC and coloured using GIMP (The GIMP Team, 2012) (see Appendix A.4 for all the pictures used in the task). The 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 types of objects but differing in details (i.e. bees with different numbers of stripes, different faces and different kinds of antennae and wings). A pack of five laminated cards was made for each of the 16 keywords and the packs were placed in a four-by-four tray, with each pack labelled with the keyword on top (see Figure 2.1b).

**Colour-number words** To enable the elicitation of several instances of vowels, for use in vowel space analysis, colour-number words reflecting as wide a range of vowels as possible were chosen for use in the task (see Table 2.2). The colours and numbers were combined in all possible ways, leading to 16 colour-numbers used in the task.

| **colour** | **number** | **vowel** |
|---|---|---|
| green | three | i |
| - | six | ɪ |
| red | - | ɛ |
| black | - | æ |
| - | four | ɔ |
| blue | two | u |

Table 2.2: The colour and number words used in the Grid task.

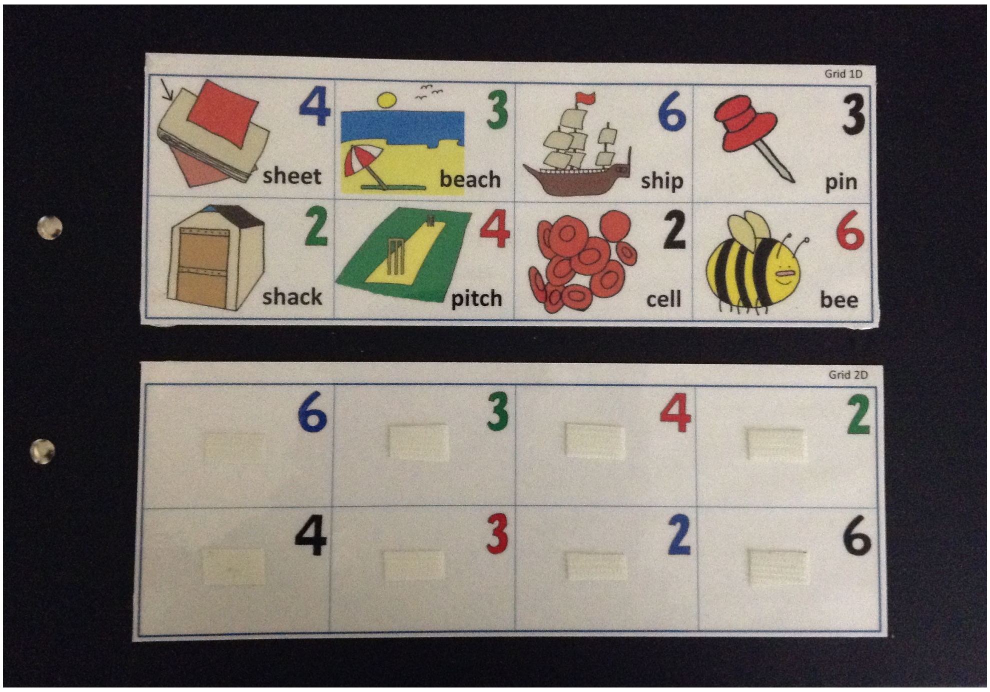
**Grid set-up** The 16 keywords and 16 colour-numbers were used to build eight pairs of two-by-four grids. The keywords and colour-number words were distributed between the grids so that each grid included keywords with several different target sounds. Several randomisation processes were applied to 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 randomisation procedures are described in more detail in Appendix A.2.

Additionally, each picture-grid was attached to an empty grid with only colour-numbers in its squares – it was placed underneath each picture-grid (see Figure 2.1). All the grids used in the current study are displayed in Appendix A.3.

**Task procedure** In the task, each participant is given the picture-grid, the empty grid and the tray containing the five different versions of the 16 keywords. The aim of the task is for each interlocutor, 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. Before the start of the task, the participants were shown a slideshow which instructed them to complete each of these three steps in order, to work from the top left hand corner horizontally, row by row, with the interlocutors taking it in turns to describe the squares on their picture-grid to each other (see Appendix B for the exact instructions given to participants). A pair of short practice grids was also created, containing the keywords 'lamp', 'castle', 'melon' and 'cat' which were not minimal pair words, and which therefore could be used to ensure that the participants understood the task. Prior to the start of the current study, extensive piloting of the task was done on adult native and non-native speakers, and on bilingual normal-hearing and hearing-impaired children to ensure that (1) the keywords were produced several times during the task, (2) the task was easy to understand, and (3) that the pictures and keywords were recognisable and familiar to children.

The task is thus similar to the picture-array tasks described in section 2.1.2 in that it involves one participant describing pictures and the other identifying the correct picture from an array. Although the Grid task therefore relies on similar 'instruction-giving' roles as the picture-array and Map tasks, the speech produced in the task does not only include instruction-giving, but also rich descriptions of pictures – especially as the differences between many of the pictures are not clear-cut, leading to the interlocutor having to ask questions and be involved in the interaction to find the correct picture. As each interlocutor described one square in their grid in turn, the contribution between interlocutors is also balanced. In the task, it is essential that the participants are able to produce and perceive the keywords correctly – therefore if either interlocutor has difficulty with one of the words, it is likely to lead to enhancement strategies being used.

Compared to the Diapix task, the Grid task is simpler – not only do the participants get given specific instructions on the exact steps to follow to complete the task, but each interlocutor also knows the specific role they play at each time, as either the 'describer' or the 'searcher'. The pictures used in the task are also more simple than those in the Diapix task, and therefore less advanced linguistic and vocabulary knowledge is likely needed for this task than for the Diapix task. In the current study, therefore, the use of both the Grid and Diapix tasks enables us to investigate any effects that task difficulty may additionally bring to the adaptations made by NH and HI children to an HI interlocutor.




(a) board


(b) tray

Figure 2.1: Example of a board and tray given to each participant in the Grid task.

# Chapter 3

# Method

## 3.1 Participants

The approach taken by the current study is somewhat different to that of the many highly-controlled studies on HI children's speech production, which typically assess the contribution of several factors affecting speech production outcomes. The focus of this study was to examine the communication skills of a group of HI older children which is broadly representative of HI populations in mainstream schools and hearing-impaired units in the UK. Thus, our aim was to be as ecologically valid as possible, both in terms of speech elicitation methods used and in terms of participants recruited. Therefore, our recruitment criteria were broad, and participants included varying levels of hearing loss, device use and language backgrounds. Due to the design of the study, we needed to recruit, from each school, a group of four children (2 HI, 2 NH) who were of similar ages and familiar with each other – this design imposed some restrictions on the schools, and thus we were also constrained by the availability of HI children within the schools. Children were only recruited from hearing-impaired units attached to mainstream schools, as they provided several HI child candidates per school for participating in the study.

Thirty-six participants in nine groups of four took part in the study. The groups were recorded in eight schools in Southern England. The mean age of the NH participants was 11.9 years (range: 9.0 - 14.3 years), and 11 of the 18 participants were female. The mean age of the HI participants was 12.0 years (range: 9.7 - 15.2 years), of whom 10 were female.

| **part.** | **gdr** | **age (y)** | **lang** | **SN** | **part.** | **gdr** | **age (y)** | **lang** | **SN** |
|---|---|---|---|---|---|---|---|---|---|
| NH1 | F | 12.8 | B | - | HI1 | F | 13.5 | M | - |
| NH2 | M | 12.0 | B | - | HI2 | M | 12.2 | M | - |
| NH3 | M | 11.2 | M | - | HI3 | M | 11.1 | M | - |
| NH4* | F | 11.0 | M | - | HI4* | F | 11.4 | B | - |
| NH5 | F | 14.3 | M | - | HI5 | F | 14.2 | M | - |
| NH6 | F | 14.4 | M | - | HI6 | F | 14.1 | M | - |
| NH7 | F | 12.8 | M | - | HI7 | F | 13.3 | SL | - |
| NH8 | F | 13.3 | M | - | HI8 | M | 12.7 | M | MA |
| NH9 | M | 9.8 | M | - | HI9 | F | 9.7 | P | MA |
| NH10 | M | 9.0 | M | - | HI10 | M | 9.9 | P | - |
| NH11 | F | 9.9 | M | - | HI11 | F | 10.3 | P | - |
| NH12 | M | 10.7 | M | MA | HI12 | M | 10.2 | P | - |
| NH13 | M | 14.3 | M | MA | HI13 | M | 15.2 | B | - |
| NH14 | F | 14.4 | B | - | HI14 | F | 13.6 | M | - |
| NH15 | M | 13.1 | M | - | HI15 | M | 12.7 | B | - |
| NH16 | F | 11.9 | M | - | HI16 | F | 11.8 | B | - |
| NH17 | F | 10.2 | M | - | HI17 | M | 9.9 | M | MA |
| NH18 | F | 9.7 | M | - | HI18 | F | 10.1 | M | - |
| **mean** | | **11.9** | | | **mean** | | **12.0** | | |

Table 3.1: Participant details. Part.=participant, gdr=gender, lang=language background (B: bilingual, M: monolingual, P: parents EAL, SL: BSL as first language), SN= additional special needs (MA: mild additional needs). The shaded areas between NH and HI participants indicate the group of four within which each child participated in the study. Participant numbers were paired so that children of opposite hearing statuses who shared a communication session were given the same participant number – therefore for example, NH1 was paired with both HI1 and NH2. *Participants NH4 and HI4 were excluded from the study due to equipment malfunction in one session.

Of the 18 NH participants, 15 were monolingual native Southern British English speakers, and three were bilingual, but all had received their entire schooling in English, and spoke English as their main language. Due to time constraints, only 8 out of the 18 NH children were given a pure tone audiometric hearing screening test on a calibrated laptop computer at octave frequencies between 250

Hz and 8k Hz in a quiet room in their school. Five of the eight children were found to have hearing thresholds within the normal range (25dB HL or better). One participant had slightly elevated thresholds between 30 and 40dB HL over several frequencies in one ear, but normal thresholds in the other. Another had an elevated threshold (30dB HL) only at 8k Hz bilaterally, and a further participant had thresholds of 30dB HL at 500Hz and 8k Hz in one ear. However, as the PTA was not conducted in audiological conditions, these slightly elevated thresholds should be treated with caution. The remaining 10 participants were presumed to have normal hearing based on parent and teacher reports. Two out of the eighteen NH participants had received some speech therapy during their school years, although one of them had been discharged over 3 years prior to testing. Two of the NH participants had mild additional needs, but the remaining participants were not reported to have any neurological or medical conditions. Further information on the NH participants can be found in Table 3.1.

According to a questionnaire administered to all participants' parents concerning their child's hearing and communication background (see Appendix C[1]), nine of the HI participants were monolingual Southern British English speakers, four HI participants' parents spoke English as an additional language (but only spoke to their children in English), four were bilingual, and one had British Sign Language (BSL) as her first language. However, as with the NH participants, all HI children had received their entire schooling in English (including some with sign support). The hearing loss level of the HI participants ranged from moderate to profound – 7 HI participants had one or two CIs, and the remaining 11 HI participants wore bilateral HAs. One HI participant had a sudden onset of hearing loss at age 2.5 years, and for one participant the age at onset of hearing loss was unknown, but the remaining participants' hearing-impairments were prelingual. One of the HI participants had progressive hearing loss. The aetiology of the participants' hearing loss was genetic (6), congenital illness (3), premature birth (1) and unknown (8) and, with the exception of one participant with mixed loss, all had sensorineural hearing loss. Three HI participants had mild additional needs, but the remaining 15 HI participants were not reported to have any ad-

[1]This questionnaire was based on a questionnaire used in Pimperton, Blythe, Kreppner, Mahon, Peacock, Stevenson, Terlektsi, Worsfold, Yuen, and Kennedy (2014).

| part. | dg(y) | HL | device | brand | CI(y) | CM | TA |
|---|---|---|---|---|---|---|---|
| HI1 | 1.0 | mod-to-sev | bilateral HA | N | - | O | A |
| HI2 | 5.5 | mod-to-sev | bilateral HA | N | - | O | D |
| HI3 | 1.5 | profound | bilateral CI | - | 3, 9 | O | A |
| HI4* | 0.5 | severe | bilateral HA | PH | - | O | D |
| HI5 | 1.0 | profound | unilateral CI | - | 4 | O | D |
| HI6 | 3.5 | moderate | bilateral HA | N | - | O | A |
| HI7 | 0.0 | mod-to-sev | bilateral HA | OT | - | TC,S | A |
| HI8 | 0.5 | sev-to-prof | CI and HA | - | 5 | O | D |
| HI9 | 0.0 | profound | bilateral CI | - | 3.5, 6 | O | D |
| HI10 | 2.5 | profound | CI and HA | - | 3.5 | O | D |
| HI11 | 0.0 | profound | bilateral HA | PH | - | N | D |
| HI12 | 5.0 | severe | bilateral HA | PH | - | N | S |
| HI13 | 1.5 | moderate | bilateral HA | PH | - | O | S |
| HI14 | 3.5 | profound | bilateral CI | - | 6, 13 | N | S |
| HI15 | 3.5 | severe | bilateral HA | PH | - | O,TC | D |
| HI16 | 0.0 | moderate | bilateral HA | N | - | O,TC | A |
| HI17 | 0.0 | moderate | bilateral HA | OT | - | O | D |
| HI18 | 0.0 | profound | bilateral CI | - | 4, 8.5 | O | D |
| **mean** | **1.6** | **severe** | | | | | |

Table 3.2: HI participant details. Diagnosis age (dg(y)) and implantation age (CI(y)) have been rounded to the nearest half year. Brand= for hearing aid users, device brand (PH=Phonak, OT=Oticon, N=no response to question on questionnaire, −=CI user) CM=communication mode with parents according to parent questionnaire (see Appendix C) (O=oral, TC=speech+sign, S=sign, N=no response to question on questionnaire). TA=teacher's assessment of child's language and communication skills (A=age-appropriate, D=delayed, S=somewhat delayed)

ditional neurological or medical conditions. The participants used mostly oral communication with their parents, although some also used total communication (speech and sign together). Participants' teachers were asked to assess the HI participants' language and communication skills in relation to peers. Only 5 of the 18 HI participants were assessed by their teachers as having age-appropriate language and communication skills. For further detail on the HI participants, see Tables 3.1 and 3.2[1].

[1] As most of the schools did not have up-to-date audiological records for the HI children who

The participants in each pair were familiar with each other prior to the recordings. According to a questionnaire administered to the participants during the study (see Appendix D), the extent to which the pairs knew each other varied widely but, on average, the participants considered each other friends and had known each other for approximately 3 years. As the NH participants all attended mainstream schools with hearing-impaired units, they were generally in regular contact with HI peers (NH mean: 4.2, on a scale of 1-not at all, 3-sometimes and 5-all the time) (see question 5 in Appendix D), and all had at least some HI friends (NH mean: 2.5, on a scale of 1-none, 3-some and 5-all) (see question 6 in Appendix D). Most NH participants usually used spoken English to communicate with HI peers, but some also used speech and sign together, BSL or gesture (question 7 in Appendix D). The HI participants were also frequently in contact with HI peers (HI mean: 3.7) and, on average, had both NH and HI friends (mean: 3.1). The HI participants used a variety of communication modes with HI peers, although most often, spoken English or speech and sign together were used. See Tables E.1 and E.2 in Appendix E for more detail on participants' familiarity with each other, their contact with HI peers, and the communication modes used with HI peers.

Each participant was given a certificate after completion of the study. Additionally, each school received a donation of 15 pounds per participant towards a charity of their choice for taking part. The study was approved by the UCL Ethics board.

## 3.2 Procedure

Each participant took part in the study for approximately 3.5 hours in three sessions; two of the sessions involved communication tasks done with a friend, and one session consisted of speech production and perception tasks which were completed alone (see Table 3.3). All sessions were recorded during school hours at the children's schools in a quiet room.

participated in the study, we were unable to collect their PTA data reliably, and it is therefore not provided here.

| Example session order | time |
|---|---|
| 1. Communication session 1 with NH friend | 1-1.5h |
| 2. Communication session 2 with HI friend | 1-1.5h |
| 3. Individual session | 45min |

Table 3.3: An example session order for participants in the study.

### 3.2.1 Communication sessions

#### 3.2.1.1 Testing Procedure

Within each group of four from each school, each participant completed one communication session with a NH friend, and the other communication session with a HI friend. Therefore, in each group of four, there were two HI-NH pairs (c and d in Figure 3.1), one HI-HI pair (b in Figure 3.1) and one NH-NH pair (a in Figure 3.1). This design enabled the elicitation of both NH-directed and HI-directed speech from both the NH and HI participants in a communicative context. In each group of four, 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 (HI-NH pairs), and the order of 'same' and 'different' pairs was counterbalanced between groups.

In the communication tasks, each participant wore an Audiotechnica AT8531 lapel microphone which was connected to a Scarlett 2i2 USB audio interface which fed into a laptop running Audacity with a sample 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 (see Figure 3.2). The video recordings were not examined in the current study, but have been used in five student projects (Chu, 2015; Dunn, 2014; Ebrahim, 2015; Harris, 2014; Ní Almhain, 2014) to analyse the interactions from a Conversation Analysis point of view (Sacks *et al.*, 1974). Future analyses are also planned to be made on the participants' use of eye gaze and gesture in the tasks.

In the two communication sessions, the pair of participants completed two sets of grids from the Grid task, followed by two Diapix pictures, and finally a further

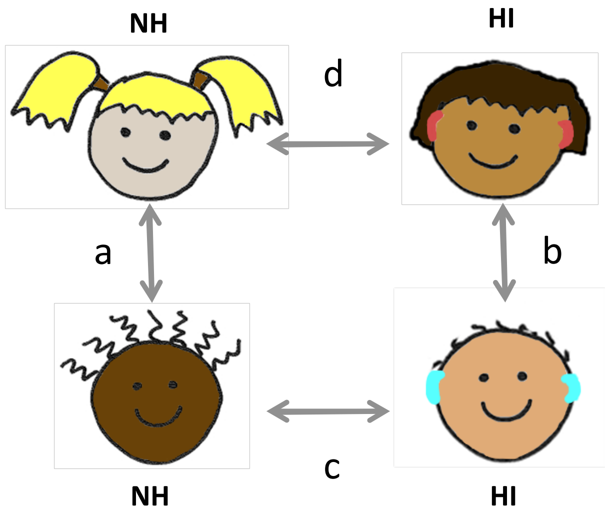



Figure 3.1: Participant pairings for the communication sessions in each group of four.

two sets of grids together (see Table 3.4). Due to time constraints, and some pairs taking fairly long to complete each set of grids, it was not always possible to complete all tasks as planned – however, all pairs completed at least two sets of grids, and one set of Diapix pictures together. At the end of the session, the participants were asked 'feedback questions' by the experimenter about the difficulty of each of the tasks, as well as the strategies they used if their friend could not understand them. Although some participants reported regularly using both speech and sign with their HI friend (see Appendix E.2), they were asked not to use sign language in the communication sessions. This instruction was given to enable the analysis to be done only on their spoken language skills. However, there were several instances in the recordings where participants did not adhere to this rule.

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 (see chapter 2 for greater detail on the Grid task). Before the start of the task, the participants were shown a slide show on the laptop of two characters playing the

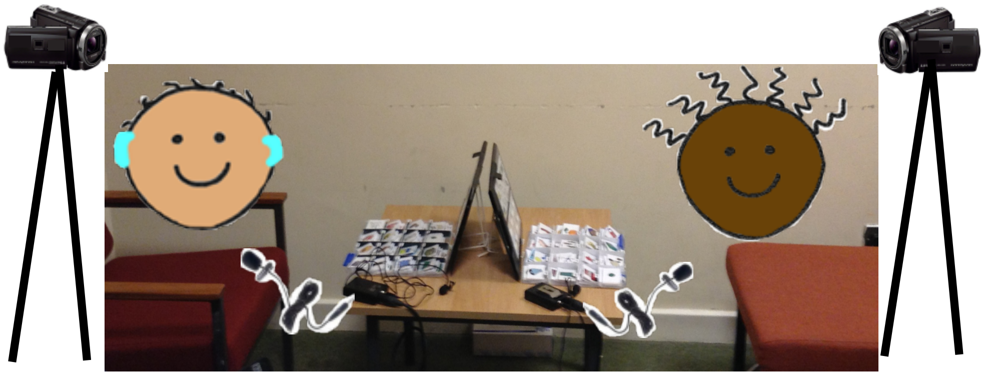


Figure 3.2: Experimental set-up in the communication sessions.

game, illustrating how the Grid task is completed (see chapter 2 and Appendix B for further detail). The participants' familiarity with the keywords was also checked, and they completed a few squares of a practice grid together to ensure that they understood the task. The order of the grids given to the participants was randomised (see Appendix A.2 for further detail) 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, per participant per condition, 32 minutes of Grid task conversation was recorded, and 3.2 grids were completed.

For the Diapix task, the same set-up as in the Grid task was used; each participant's Diapix picture was attached to their board, and the participants faced each other so that they could not see each others' pictures. Participants were asked to start the description of their pictures from the top left- hand corner and work clockwise; they were also asked to contribute equally to the conversation, and to circle any differences that they found (for further detail on the background of the Diapix task, see chapter 2). Before the start of the task, the participants were shown an example picture, and the different types of differences found in the example picture were discussed. Then, to ensure that the participants had understood the task, they were given a training picture and asked to find a few differences. The participants were given a 10-minute time-limit for finding the 12 differences[1]. Only four different Diapix pictures were used in the experiment –

[1]However, due to time constraints, for several pairs of participants, less than 10 minutes could be spent on each Diapix picture.

two each from the 'Farm' and 'Beach' picture-scenes from the full set of pictures (Baker and Hazan, 2011) were included, as they contained less writing than the 'Street' scenes, and were therefore thought to be more appropriate for HI children. If two Diapix pictures were done in one condition, participants completed one Farm and one Beach picture scene together. Any written words in the four chosen pictures which were deemed to be difficult for HI children to read were changed to simpler ones. The order of the picture-scenes and the particular pictures presented to participants was counterbalanced between pairs and groups, and no same picture was completed by a participant twice. After each picture was completed, the experimenter and the participants looked at the two pictures together and reviewed the differences which the pair had found. On average, per participant per condition, 13.5 minutes of Diapix task conversation was recorded, and 1.8 Diapix pictures were completed.

| **Communication sessions** |
|---|
| 1. Grid task x2 |
| 2. Diapix task x2 |
| 3. Grid task x2 |
| 4. Feedback questionnaire |

Table 3.4: Procedure for sessions 1 & 2 (communication sessions).

While the participants were engaged in the Grid and Diapix tasks, the experimenter was seated in the same room, and monitored the recording on the laptop. As it was desired that the conversation between the two participants would be as natural as possible, the experimenter attempted to seat herself as far from the participant pair as possible, and only took part in the conversation if explicitly asked questions by participants. For a few pairs, a teacher was also seated in the same room, but was concentrated on other tasks and did not participate in the recorded conversation.

The same procedure was used in both communication sessions, except that in the second communication session the participants were not given introductions to each task or asked to complete practice pictures. Each communication session lasted between 1 and 1.5 hours.

#### 3.2.1.2 Preliminary file processing

In the Grid and Diapix tasks, each participant's speech was saved on to a separate channel. Due to equipment malfunction, the audio recordings for the communication session between NH4 and HI4 were unusable, and therefore both participants were excluded from further analyses. In total, excluding the data from participants NH4 and HI4, 51.6 hours (341 files) of single-channel recordings were made. Of those, 36.3 hours (218 files) were Grid recordings, and 15.3 hours (123 files) were Diapix recordings.

The dual-channel recordings were transcribed orthographically by the author using Praat (Boersma and Weenink, 2015). Each speaker's channel was given its own transcription tier, and the transcriber specified the locations of the boundaries of utterances in each speaker's tier[1]. The transcription criteria followed the general guidelines of those used in Hazan and Baker (2011), Hazan *et al.* (submitted) and Pettinato *et al.* (2016). A description of the protocol used can be found in Appendix F. 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. In the Grid task, particular attention was paid to ensuring that the transcribed keyword was the one intended to be produced by the speaker – this was done by checking the speaking context, i.e. by ensuring that the keyword matched the speaker and/or interlocutor's grid, by examining the pictures chosen by the speaker in their finished grid (see section 5.3.1.2), and by viewing the video recording, if necessary. If, even after checking the context, the transcriber was still unsure of which word the speaker intended to produce, the keyword was transcribed as unintelligible. Any speech overlap between participants was also tagged. Within-speaker pauses over 500ms in length were marked as 'SIL' to enable later pausing analyses to be done. For this measure, 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), and thus is long enough not to include stop closure durations. As well as speaker tiers, additional

[1]This was done to ensure that any speech from the interlocutor's channel would not interfere with the accuracy of later alignment.

tiers were created for transcribing external events (such as the school bell ringing) or instances in which participants mispronounced words. Additionally, in the Diapix task, the points during the interaction at which participants found differences between their pictures were labelled on a point tier.

The transcribed utterances from each speaker's tier were extracted using a Praat script and converted to txt 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. The individual utterances were then reinserted into the speaker's original long single-channel wav file. Then, the alignment of each single-channel recording was checked on the word-level using Praat – approximately 65% of the checking was done by the original transcriber, and the rest was checked by two researchers with extensive experience in phonetics. Those word intervals for which a change was made during checking were extracted using a Praat script, and realigned on the phoneme level. The new phoneme-level alignment was then reinserted into the word-level checked TextGrid. Each single-channel wav file, and its corresponding TextGrid file with phoneme- and word-level tiers was then used for analysis. Further file processing relevant to each analysis is described in each following chapter.

### 3.2.2 Individual session

#### 3.2.2.1 Testing procedure

In the individual session, each participant completed a picture-naming task, as well as three speech perception tests (see Table 3.5). They sat facing a laptop in a quiet room, with the researcher sitting beside them. The same audio recording set-up as in the communication session was used. A Dell A215 loudspeaker was positioned 70 centimetres from the participant, at eye-level directly above the laptop screen. The three perception tests (VCV, BKB and WiNiCS tasks) were presented via the loudspeaker at 70dB SPL.

The order of the VCV and BKB tests was counterbalanced between participants to avoid effects of speaker familiarity; the WiNiCS test was always done after the first speech perception test. At the end of the session, participants were given a short questionnaire on their familiarity with their interlocutors from sessions 1 and 2 and on their experience in communicating with HI peers. Time permitting, NH participants were given a hearing screening test at the end of the session. Session 3 lasted approximately 40 to 50 minutes per participant.

| **Individual session** |
|---|
| 1. Picture-naming task (1/2) |
| 2. VCV perception task* |
| 3. Speech-in-noise (WiNiCS) task |
| 4. BKB perception task* |
| 5. Picture-naming task (1/2) |
| 6. Questionnaire |
| 7. Hearing screening (NH only) |

Table 3.5: Procedure for session 3 (individual session). *The order of the VCV and BKB tests was counterbalanced between participants.

### A. Speech production

**Picture-naming task** A speech production task was used to elicit segmental contrasts in minimal pair keywords. Participants were asked to name pictures on a screen using the carrier sentence ‘I can see a [keyword]’. The task was the same as that used in Romeo *et al.* (2013).

The task was designed to elicit the word-initial /p/-/b/ contrast and the /s/-/ʃ/ distinction. Four keywords were included for each sound; therefore there were sixteen pictures representing different keywords used in the task (see Table 3.6). Altogether, the task elicits 64 keywords and 32 ‘distractor’ keywords.

Participants were familiarised with the pictures used in the task and completed a practice session on the computer before starting the task. Half of the picture-naming task was done at the beginning of the session, with the remaining part completed at the end of the session (see Table 3.5). DMDX (Forster and Forster, 2003) was used to present the pictures and to record the sentences with a sample rate of 22,050Hz. The 96 pictures were pseudo-randomised so that the same

| /p/ | /b/ | /s/ | /sh/ |
|---|---|---|---|
| pea | bee | cell | shell |
| pin | bin | seat | sheet |
| pill | bill | sack | shack |
| peach | beach | C | sheep |

Table 3.6: The keywords used in the picture-naming task.

keyword did not occur twice in a row. Although the task was planned to be used to explore within- and between-speaker variability in the NH and HI children's speech, as the Grid task was found to provide a sufficient number of elicitations of each target phoneme to enable within-speaker variability to be analysed, the picture-naming task itself is not further analysed in the current study.

### B. Speech perception

**Audiovisual consonant (VCV) identification task** An audiovisual consonant perception task was used to investigate the HI children's ability to perceive consonant contrasts without any lexical cues. A set of recordings by a female Standard Southern British English (SSBE) speaker made at UCL were used. They consisted of consonants presented in a vowel-consonant-vowel (VCV) context. Sixteen consonants (/b,d,f,g,j,k,l,m,n,p,s,ʃ,t,v,w,z/) were presented once in three different vocalic contexts (/ɑ,i,u/). Participants were presented with the video and an array of 16 consonants on the screen[1], and they were instructed to both repeat what they had heard and to point at the consonant on the screen that they perceived. The researcher then clicked on the consonant which the participant indicated as correct. The participants completed one list, with the consonants and vocalic contexts randomised for each participant. They were presented with 48 VCVs altogether. At the beginning of the task, participants were familiarised with the consonant array, and the researcher read through the consonants together with the participant. A few practice VCV stimuli were also given to the participants. The participants were audio-recorded, but the correctness

[1]The consonants /j/ and /ʃ/ were represented as 'y' and 'sh', respectively.

of their response was judged on the basis of selecting the consonant array. Due to time constraints, three NH participants (NH3, NH4 and NH8) were unable to complete the task.

**Sentence (BKB) perception task in quiet** The participants were given a sentence perception task to explore their ability to understand simple sentences. Two lists from the BKB (Bench, Kowal, and Bamford, 1979) test recorded at UCL by the same female SSBE speaker as in the VCV test were used. One list contains 16 sentences with three or four keywords each. One list was presented in the audiovisual (AV) mode and the other in audio-only (A) mode on a computer screen in quiet, and the order of the A and AV tests, and the sentence lists were counterbalanced between participants. The sentences within each list were randomised. The participants in the first group (HI1, HI2, NH1 and NH2) were only presented with audiovisual BKB sentences from one list. The participants were instructed to repeat the sentence they had heard, and their speech was audio-recorded for later coding. Before starting the task, participants were presented with a few practice sentences by the experimenter. Due to time restrictions, only 13 out of the 18 NH participants completed the BKB sentence test. However, all HI participants except HI14 were given the test.

**Speech-in-noise task (WiNiCS)** The Words in Noise in Connected Speech (WiNiCS) task (Messaoud-Galusi, Hazan, and Rosen, 2011) was used to measure the participants' ability at perceiving speech in noise with no visual cues. The same procedure as in Messaoud-Galusi *et al.* (2011) was used to enable comparisons to be made between the current study's participants and the 51 normal-hearing children in their study. In the task, the participants were presented with sentences such as 'Show the dog where the [colour] [number] is'. The screen displays an array of six buttons, all of which have the same digit on them in different colours. The participants were instructed to click on the button with the digit of the correct colour. A three-up / one-down adaptive procedure was used; the first sentence was presented at 20 dB SNR, and the SNR level was varied to track 79.4% correct in the test. For the first group of participants, the test ended either after eight reversals or a total of 30 trials. As one of the hearing-impaired participants had fewer reversals than expected during the first 30 trials, which made the participant's results unusable, from the second group of participants onwards, the maximum number of trials was extended to 40. Participants were shown a screenshot of the task screen set-up before starting the task.

The file processing and analyses for the speech perception tasks are presented in the next chapter, chapter 4.

# Chapter 4

# Perception results

## 4.1 Introduction

As discussed in chapter 1, due to deprived early auditory experiences and the poorer quality of auditory input received by HI children even through HAs and CIs, the majority of HI children display deficits in the perception of speech compared to NH peers (e.g., Moeller *et al.*, 2007c) – although the extent to which speech perception and receptive language is delayed is likely to vary greatly between HI children (see section 1.2.1). Specifically, previous studies have found that HI children with milder losses perform better on speech perception and receptive language tests than those with more severe losses (e.g., Eisenberg, 2007). In this chapter, the results of the three speech perception tests conducted on the participants in this study are analysed to examine the extent to which the speech perception skills of the HI participants in the current study differ from their NH peers'.

## 4.2 File Processing

### 4.2.1 Sentence (BKB) perception task

The researcher listened to the BKB recordings from each participant and scored their responses according to the number of keywords correct in each condition. Minor grammatical errors, such as incorrect tense production, were ignored when determining keyword scores. Unintelligible words were deemed as keyword errors. A second researcher, who is experienced in listening to HI children's speech, checked the scoring of any sentences of which the first scorer was uncertain. Lastly, the % correct keywords per condition per participant was calculated.

### 4.2.2 Audiovisual consonant (VCV) identification task

For the audiovisual VCV test, the % of correctly identified consonants was calculated for each participant.

### 4.2.3 Speech-in-noise task (WiNiCS)

The adaptive procedure in the WiNiCS test automatically calculated the SNR level at which the listener achieved 79.4% words correct. This was calculated for each participant.

## 4.3 Results

### 4.3.1 Sentence (BKB) perception task

All NH participants scored between 96 and 100% correct in both the A and AV conditions, and therefore their performance on the BKB test was not analysed further. Figure 4.1 displays the results of the BKB test for all HI participants. Only five HI participants scored below 90% in any condition – HI5, HI9, HI10, HI11 and HI18, all of whom had profound hearing loss. With the possible exception of HI9, each of these five participants also seemed to receive substantial benefit from visual information in the test – the audiovisual benefit (calculated as the score from the A condition subtracted from the score from the AV condition) ranged from 6 percentage points for HI9 to 30 percentage points for HI11 (mean: 19 percentage points). All other HI participants performed at or near ceiling, and therefore the contribution of visual information to their speech comprehension could not be analysed.

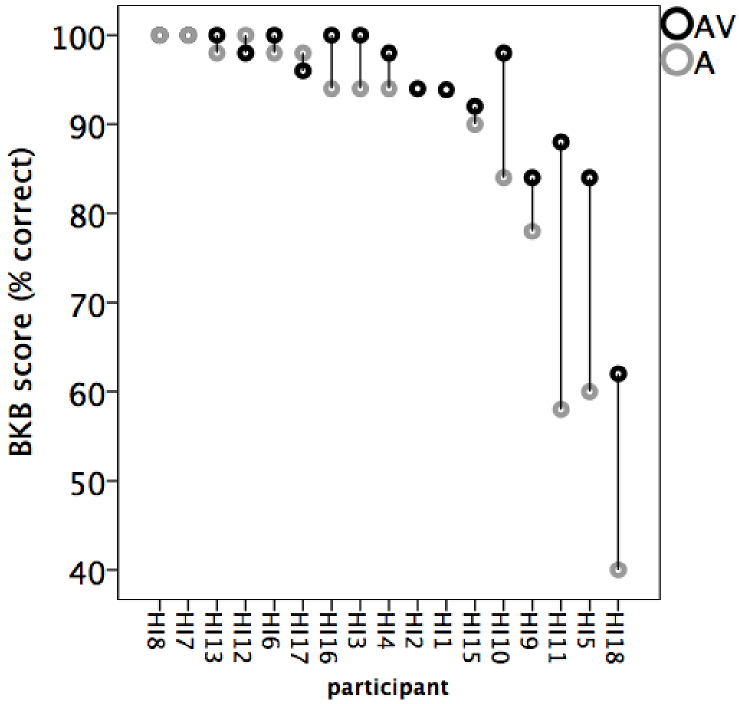
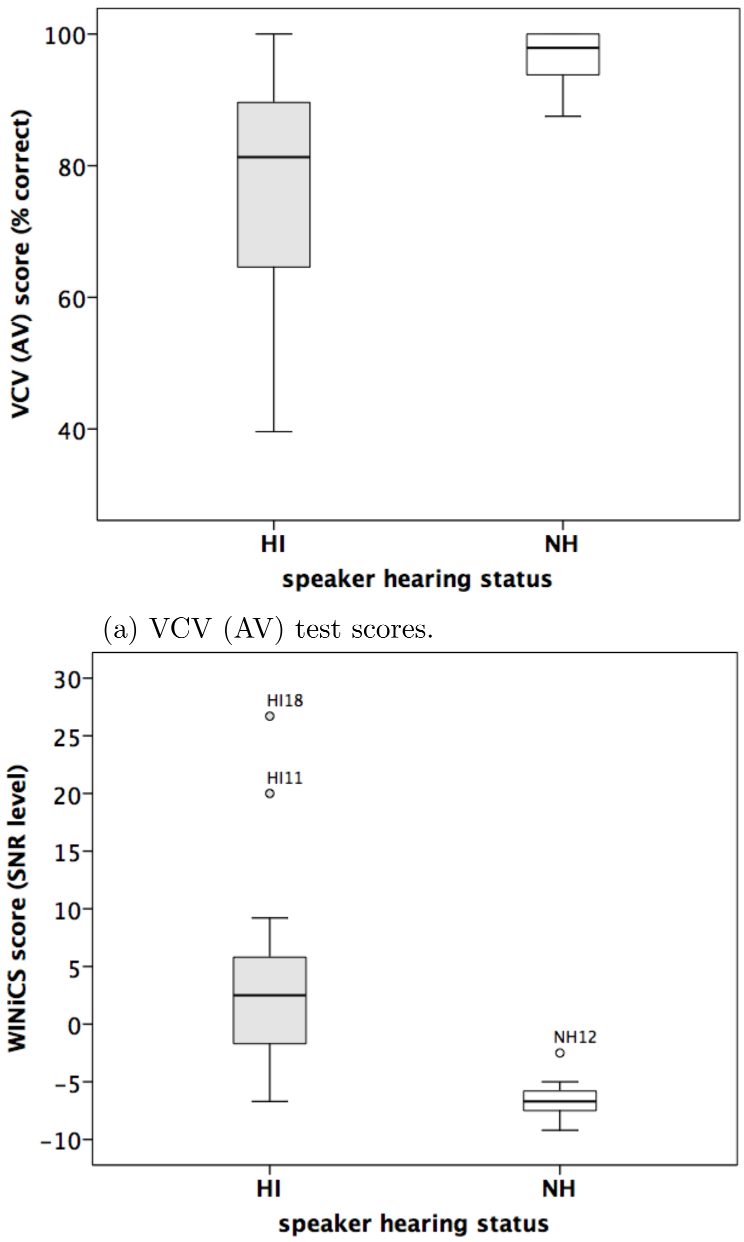



Figure 4.1: The results of the BKB sentence perception test for each HI participant in the AV (black) and A (grey) conditions. HI1 and HI2 only completed the test in the AV condition.

### 4.3.2 Audiovisual consonant (VCV) identification task

An independent samples t-test between NH and HI groups' scores for the VCV perception test demonstrated that, as expected, NH participants (mean: 96.3%) had higher accuracy scores than HI participants (mean: 77.4%) ($t(31)=-4.0$, $p<0.001$) (see Figure 4.2a). The range of scores obtained by HI participants was nonetheless large – ranging from 100% (HI13) to 40% (HI9). Seven of the HI participants (HI1, HI4, HI6, HI7, HI12, HI13 and HI17) scored within NH range (88% or above) – of these participants, all but HI4 and HI12 had either moderate or moderate-to-severe hearing loss.

When specifically examining the perception scores of the two consonant contrasts which also occur in the Grid task, HI participants were 84.3% accurate in identifying /p/ and /b/ on average (SD: 19.4%, range: 33.3-100%) and 72.2% correct in /s/ and /ʃ/ identification, on average (SD: 19.0%, range: 33.3-100%).




(a) VCV (AV) test scores.

(b) WiNiCS test scores.

Figure 4.2: The results of the VCV (AV) and WiNiCS perception tests for NH and HI participants.

Thus, while some HI participants were accurate in perceiving these consonants, most HI participants had at least some difficulty with perceiving these sounds. Therefore it is likely that HI participants will find it difficult to perceive at least some of the minimal pair keywords correctly in the Grid task.

### 4.3.3 Speech-in-noise task (WiNiCS)

In the WiNiCS test, HI participants required a higher SNR level (mean: 3.5dB SNR) than NH participants (mean: -6.6dB SNR) to achieve 79.4% words correct ($t(33) = 4.7$, $p<0.0001$) (see Figure 4.2b). The scores for NH participants did not significantly differ from those of the 51 6- to 13-year-old NH children tested in Messaoud-Galusi *et al.* (2011) ($t(67)=1.1$, $p=0.272$, n.s.). Again, a wide range of variability in the HI participants' scores was evident. Only two HI participants obtained results within one standard deviation from the NH mean – HI7 and HI13, who had moderate-to-severe and moderate hearing loss, respectively.

### 4.3.4 Correlations between measures

A Pearson's correlation demonstrated that the speech perception scores for the HI participants obtained in the WiNiCS and the VCV tests were strongly correlated with each other ($n=16$; $p<0.001$; $r=-0.771$; $R^2=0.594$) – the higher the participant's VCV score, the lower the SNR required to obtain 79.4% correct on the WiNiCS test. As shown in Figure 4.3, HI participants with severe-to-profound and profound hearing losses generally performed worse than HI participants with milder hearing loss levels. Notably, all of the five HI participants who achieved low scores in the BKB sentence test were also the poorest performers in both the VCV and WiNiCS tests.

## 4.4 Discussion

In summary, in this chapter it was confirmed that the HI participants in this study have significantly greater difficulty in both consonant and speech-in-noise perception than their NH peers. As expected from previous studies (Eisenberg,

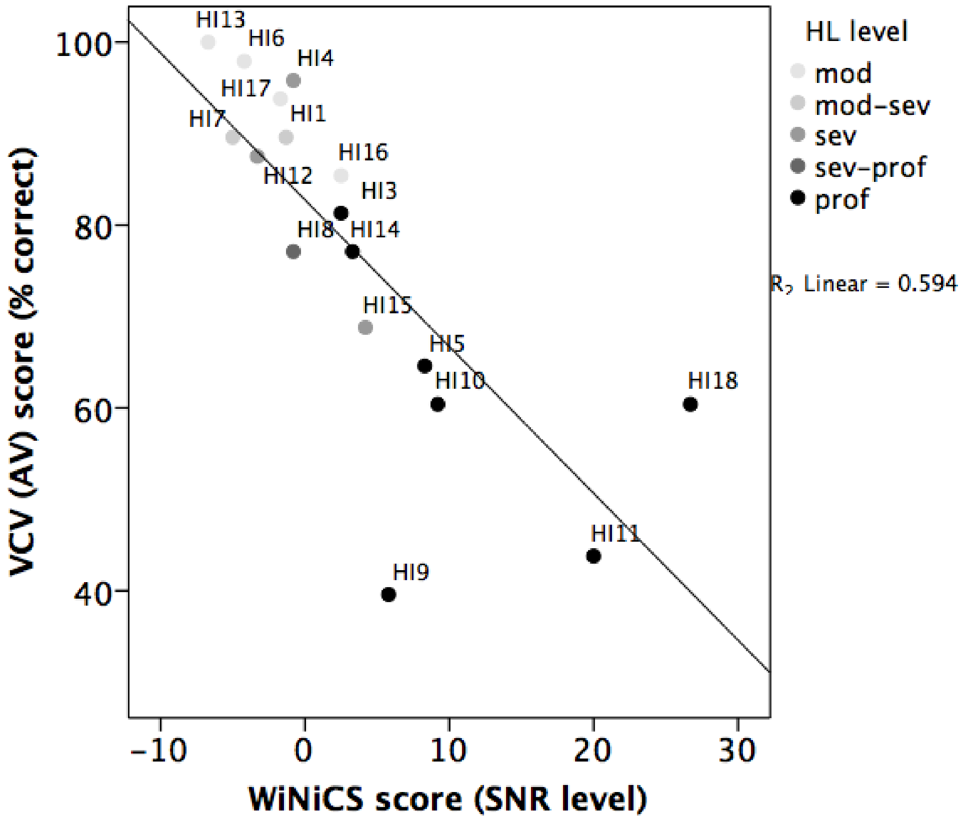



Figure 4.3: The correlation between VCV and WiNiCS scores for HI listeners.

2007; Hennies *et al.*, 2012), there was a great deal of variability between HI participants' performance on the tests, which seemed to be generally related to the hearing loss level of the participants – those with more severe losses performed worse than those with milder losses. In particular, five HI participants with profound losses consistently obtained the lowest scores out of all HI participants on all three tasks. Some HI participants even achieved similar scores to their NH peers on both the VCV and WiNiCS tests – these tended to be the participants with only moderate or moderate-to-severe hearing loss levels (as in e.g., Borg *et al.*, 2007). However, the level of hearing loss did not account for all the variability in the data – for example, three of the HI participants with profound or severe-to-profound hearing loss (all wearing CIs) performed similarly to HI participants with milder losses.

Due to the ceiling effect in the BKB test for most participants, the contribution of visual information in word perception could not be assessed for most participants. However, for the five HI participants who achieved lower scores on the test, visual information significantly contributed to their word comprehension.

Based on the results from this chapter, we would expect that, due to the HI participants' speech perception deficits, participants would find the communication tasks more difficult when performed with a HI interlocutor. Therefore, it is likely that in those conditions speech modifications would have to be made to adapt to the HI listener's needs.

# Chapter 5

# Task results

## 5.1 Introduction

As discussed in chapter 2, the Grid task was developed in the current study to elicit multiple repetitions of segmental contrasts in NH and HI children's peer interactions. Additionally, the Grid task was developed as an easier and less complex alternative to the Diapix task, to enable an analysis to be done on the effect of between-task difficulty in children's speech adaptations. The first part of the current chapter assesses whether the Grid task was successful in these aims.

The second objective of this chapter is to examine whether the speaking condition (interlocutor hearing status) influenced the difficulty participants had in completing the Grid and Diapix tasks. Task transaction time – the time taken to complete a task – is used in this study to measure within-task difficulty. The measure has been used successfully in previous studies using the Diapix task methodology to discriminate between the communicative efficiency of native and non-native interlocutor pairs (Van Engen *et al.*, 2010), and to demonstrate that interacting through a communication barrier is more effortful for speakers than communicating through a normal auditory channel (Hazan and Baker, 2011; Hazan *et al.*, submitted). Based on the perception results in chapter 4 showing that the HI participants in the current study display deficits in their speech perception skills relative to the NH participants, it was hypothesised that participants would find the communication tasks more difficult to complete in pairs involving HI interlocutors. If this is the case, it is likely that speakers will need to modify their speech and language to the needs of their HI interlocutors in these conditions.

## 5.2 File Processing

### 5.2.1 Task evaluation

**Total number of keywords and phonemes elicited in the Grid task** The word-level checked TextGrid files were used to find all tokens of all keywords in the Grid task. However, keywords which were spoken in overlap with the interlocutor, words spoken while laughing, and words which were only partially spoken were excluded, as an acoustic-phonetic analysis of these tokens would not be possible. Plurals of keywords were included, as were keywords for which noise did not occur at the target segment. The target phonemes in the keywords were segmented manually using Praat (see chapter 7 for greater detail). Only the keywords in which the target segments were usable were included in the total keyword count. Similarly, only the target segments which were analysable were included in the total target phoneme count analysed in this chapter.

**Participants' evaluation of between-task difficulty** As described in chapter 3, a feedback questionnaire was given to each participant pair at the end of each communication session. For each task, the experimenter asked the participants: "What did you think of this game? Was it easy, difficult, or just right?". Responses were scored according to whether participants said a task was easy (1), just right (2) or difficult (3). If a participant mentioned two possible categories ("easy and in-between") then their response was scored as being between the two categories (1.5 or 2.5).

**Grid correctness** The videos recorded in each communication session were viewed, and the number of correct and incorrect pictures in each participant's finished grid was counted. The number of incorrectly completed squares in the grids and the total number of grid squares completed by each pair were then calculated to obtain a measure of the % correct picture-squares per pair. The type of error (keyword error or incorrect picture-version error) was also noted. Only very few errors in the location of the grid picture were observed, and were not treated as errors in this analysis.

**Proportion of time spent speaking in each task** The total duration of all words spoken by each speaker in each file was calculated using a Praat script. The total speech duration was then divided by the duration of the entire file to obtain the proportion of speech per file by each interlocutor. The mean time spent speaking per participant per task was then calculated to enable an analysis of speech elicitation efficiency for each task to be made.

### 5.2.2 Task transaction time

**Grid task** 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 (see section 5.2.1 above) to obtain the mean time taken to find one correct picture in the grid task. The mean over the two to four grids completed by each pair was calculated to obtain a mean transaction time score per participant pair.

**Diapix task** As mentioned in chapter 3, while transcribing the Diapix files, the transcriber marked on a separate tier on the TextGrid file the points at which the pair of participants found each difference in each Diapix picture. A Praat script was used to obtain the time at the 8th marked difference for each file. Then, for each Diapix picture per participant pair, the mean time taken to find one difference out of the first eight differences found was taken as a measure of task difficulty. If less than eight differences were found, the measure was the mean time taken to find each difference. If two Diapix pictures were completed by the pair, the mean time over the two pictures was calculated, to obtain one measure of task difficulty per participant pair. The pair HI15-NH15 was excluded from this measure as they did not find any differences between their Diapix pictures in either of the two picture-sets they completed.

## 5.3 Results

### 5.3.1 Task evaluation

#### 5.3.1.1 Number of target phonemes elicited

**Keyword elicitation** One of the main aims of the Grid task was to elicit, in spontaneous communicative speech, multiple repetitions of keywords with certain target sounds. Figure 5.1 displays the mean number of keyword elicitations per speaker per condition. Altogether, 4397 keywords, which could be used for measuring the target speech segments, were elicited in the Grid task – approximately 20.2 tokens per completed grid per speaker per condition. A mean of 4.0 tokens (SD: 0.6) of each keyword were elicited per speaker per condition.

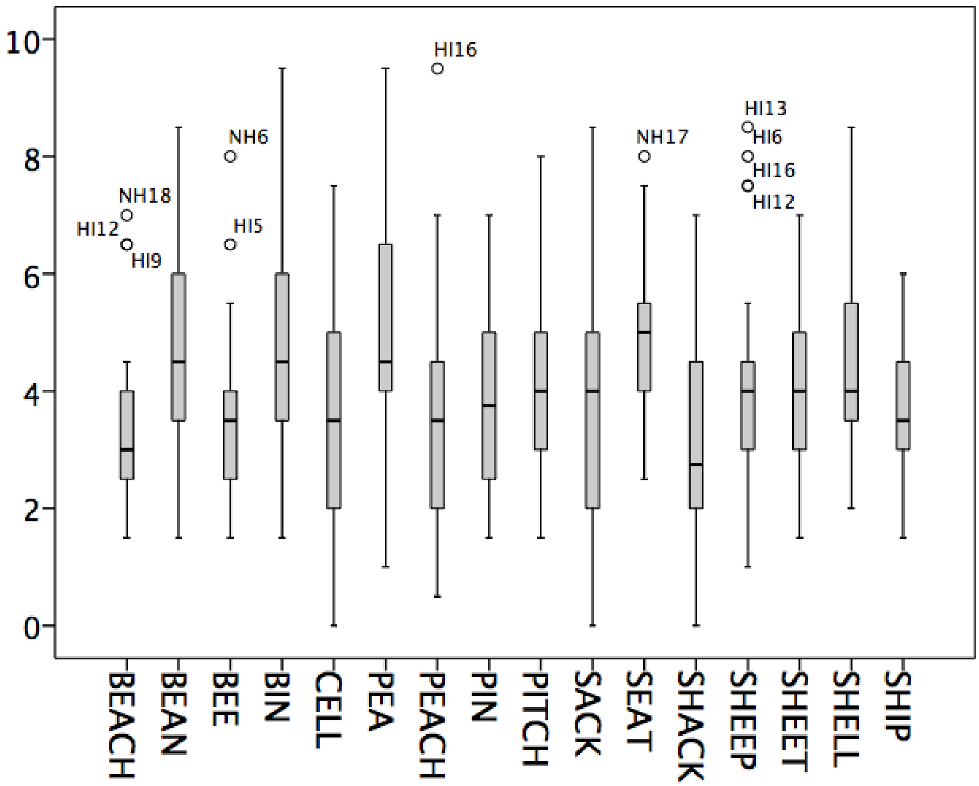
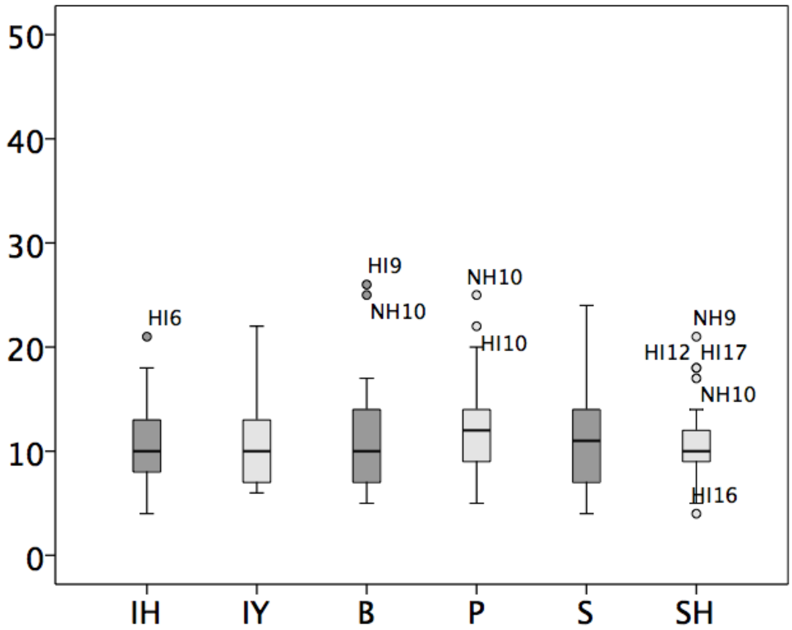



Figure 5.1: The mean number of keyword elicitations per speaker per condition.

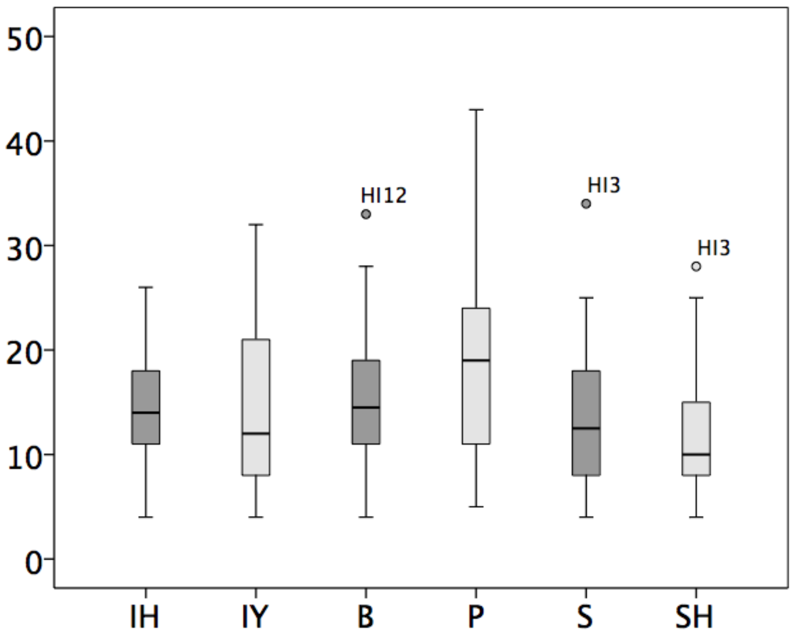
**Target phoneme elicitation** As Figures 5.2a and 5.2b show, the Grid task generally elicited several repetitions of each segment. In the NH-directed condition, a mean of 11.2 (SD: 2.9) instances of each target sound per speaker per condition were elicited, while in the HI-directed condition, there were a mean of 14.8 (SD: 5.3) occurrences of each segment per speaker per condition. The number of repetitions obtained per speaker per condition is considerably lower than that in Romeo *et al.* (2013), in which a picture-naming task was used to elicit approximately 30 instances of each target phoneme per speaker. However the number of target phonemes elicited in this study is generally greater than that of previous spontaneous speech studies. For example, Garnier *et al.* (2010) analysed approximately five repetitions of each of /b/, /p/ and /m/ phonemes in their river task. In Sankowska *et al.*'s (2011) study, approximately 9.6 and 9.4 tokens of each vowel type (long or short) (or, approximately 3 tokens per vowel) were elicited in ADS and Lombard speech conditions, and 17.8 tokens were elicited per vowel type (or approximately 6 per vowel) in the FDS condition. The number of repetitions obtained in the Grid task is also substantially higher than in the two previous studies which have analysed variability in HI children's speech – in Uchanski and Geers (2003), 4 to 8 tokens of each phoneme were elicited, and in Metz *et al.* (1985), only 4 repetitions of each segment were obtained. Based on these previous studies, it seems that the number of segment repetitions elicited in the Grid task is sufficient for analysing contrast enhancement and variability in the current study.

A paired t-test demonstrated that a greater number of target sounds were elicited in the HI-directed condition than in the NH-directed condition ($t(33)=3.87$, $p<0.01$). The more frequent repetition of target sounds in the HI-directed condition suggests that HI children found the target contrasts more difficult to produce and perceive, and this may therefore lead their interlocutors to enhance these contrasts in HI-directed speech. Further analysis of the segmental contrasts is conducted in chapter 7.

**Summary** In summary, the Grid task was successful in its first aim of eliciting multiple repetitions of segmental contrasts which HI children may find difficult to produce and perceive. The next section explores whether the Grid task was




(a) NH-directed condition.




(b) HI-directed condition.

Figure 5.2: Number of target sounds elicited in each condition. ‘IH’ = /ɪ/, and ‘IY’= /i/.

successful in its second main aim of being an easier and less complex task for NH and HI children than the Diapix task. It also explores whether the Grid task is as efficient at eliciting speech as the Diapix task.

#### 5.3.1.2 Between-task difficulty

**Examples of elicited speech** Tables 5.1 and 5.2 display examples of the type of interaction elicited from both the Grid and Diapix tasks. As pointed out by Ní Almhain (2014) and Ebrahim (2015), the Grid task involves interlocutors working through distinct, predetermined steps in building each others' grids – first the correct keyword is established, followed by a description of the correct picture version, and finally the location of the picture is determined. In the Diapix task, however, neither speaker takes up a specific role, and interlocutors together contribute to the conversation to find the differences between their pictures. The differences between the two tasks imply that the Diapix task may be more difficult to complete than the Grid task.

**Between-task difficulty** In the feedback questionnaire, participants rated the Diapix task as being between 'just right' and 'difficult' (mean: 2.4), and the Grid task as being between 'easy' and 'just right' (mean: 1.6). If participants gave a reason for rating the tasks as difficult, they often mentioned that, in the Diapix task, it was difficult to identify a difference if only one of the pair had a particular object in their picture, and that sometimes it was difficult to locate the object that their friend was describing. For the Grid task, a few participants mentioned having difficulty with some of the minimal pairs in the task, and others commented on it being difficult to describe some of the Grid task pictures.

These responses imply that the Diapix task is indeed seen as a more difficult and complex task by the participants, perhaps partly due to the lack of shared referents in the task. Indeed, as shown by figure 5.3, participants found it difficult to identify all the differences in their pictures, with NH-NH pairs finding approximately 73% of all differences (mean 8.8 out of 12), HI-NH pairs finding approximately 58% of all differences (mean 6.9 out of 12) and HI-HI pairs finding only about 51% of all differences (mean 6.1 out of 12) in their pictures[1].

[1] Due to time constraints, however, not all pairs had the same amount of time to find the

| speaker | speech |
|---|---|
| NH9 | okay SIL it's a bee SIL it's got a smiley face SIL s- <SIM_and> |
| NH10 | <SIM_yeah> |
| NH9 | it's sort of like that |
| NH10 | is it a bean or a bee |
| NH9 | bee B E E SIL <LS> it's got a smiley face |
| NH10 | is it on a flower |
| NH9 | no SIL it's sort of like stretching |
| NH10 | has it got like a p- pink nose |
| NH9 | no got no nose |
| NH10 | no nose |
| NH9 | and his ears are curly |
| NH10 | hmm hopefully I got it SIL what number is it |
| NH9 | black no SIL red four |
| NH10 | four got it |

Table 5.1: An example interaction from the Grid task of pair NH9-NH10 completing Grid 3D-4D. 'SIL' denotes a within-speaker pause, words transcribed as '<SIM>' are spoken simultaneously by interlocutors, and <LS> is a lip-smack.

| speaker | speech |
|---|---|
| NH6 | there's two err older people sitting on SIL deck chairs SIL outside the food shack |
| NH5 | the woman's on a turquoisy blue and the man's on an orange |
| NH6 | yeah and there's a little white radio next to <SIM_them> |
| NH5 | <SIM_yeah> |
| NH6 | the woman's wearing red sandals |
| NH5 | yeah and the man's wearing |
| NH6 | red socks with brown <SIM_sandals> |
| NH5 | <SIM_no> |
| NH6 | oh |

Table 5.2: An example interaction from the Diapix task of pair NH5-NH6 completing a Diapix beach scene picture.

differences.

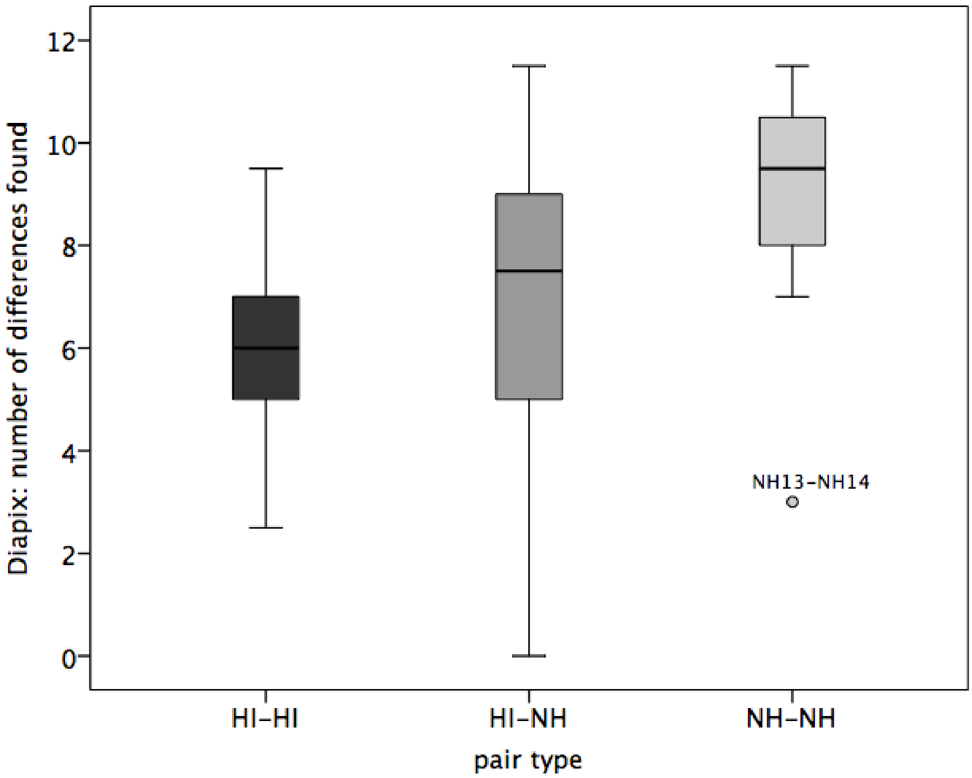



Figure 5.3: Mean number of differences found in the Diapix task (n: HI-HI=9; HI-NH=17; NH-NH=9).

In the Grid task, however, participants were able to identify most of the picture-squares correctly (see Figure 5.4). On average, NH-NH pairs found approximately 97%, HI-NH pairs identified about 90% and HI-HI pairs found approximately 87% of picture-squares correctly.

Of the errors which were made, the majority (approximately 65%) were ones in which the incorrect picture version was chosen for a certain keyword. The remaining errors were those in which an incorrect keyword was chosen. These kinds of errors were very rare for NH-NH pairs, for whom only 1% (5) of all completed picture-squares contained an incorrect keyword. Of those five errors, most were confusions between the keywords 'bean' and 'pea'. For HI-NH and HI-HI pairs, keyword errors accounted for approximately 4% and 5% of all completed picture-squares, respectively (total: 43 keyword errors). Table 5.3 displays the keyword errors made in terms of the incorrect phonetic contrast in the keyword

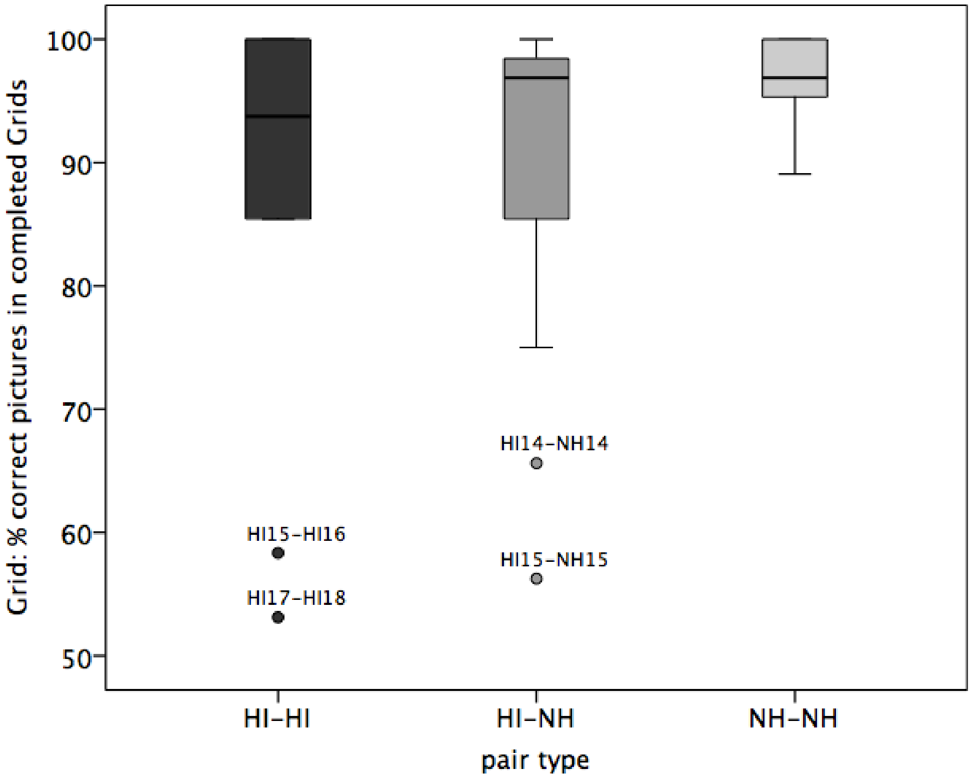



Figure 5.4: Percent correct picture-squares in the completed grids in the Grid task (n: HI-HI=9; HI-NH=17; NH-NH=9).

for HI-NH and HI-HI pairs, as well as the most common keyword confusion for that phonetic contrast. The most common errors were minimal pairs involving the contrasts /i/-/ɪ/ and /s/-/ʃ/. Most of the keyword objects for the contrasts which were confused, with the possible exception of 'bean' and 'pea', are very different in type, and therefore the description of the correct object by the describer must have led to great confusion by the interlocutor. Therefore, for these errors to have occurred in the completed Grids, a major miscommunication between interlocutors must therefore taken place without it being adequately repaired. The above analysis therefore suggests that the keywords chosen for the Grid task were ones which the NH and HI participants had trouble producing and perceiving, and which therefore may have led to many opportunities of miscommunication and repair.

| contrast | % errors | example |
|---|---|---|
| /i/-/ɪ/ | 34.6 | peach-pitch |
| /s/-/ʃ/ | 21.2 | sack-shack |
| no coda - final /n/ | 19.2 | pea-bean |
| /b/-/p/ | 13.5 | bean-pea |
| unrelated | 7.7 | cell-shack |
| final /p/ - final /t/ | 3.8 | ship-sheet |

Table 5.3: Types of Grid task keyword errors made by HI-NH and HI-HI pairs (out of 43 errors made) according to phonetic contrast, along with an example of the most common keyword error made within that category.

**Proportion of time spent speaking** To investigate the efficiency of each task in eliciting speech, a paired t-test was conducted on the proportion of time each speaker spent speaking in the Grid and Diapix tasks. Speakers spent more time speaking in the Diapix task (mean: 28.5%) than in the Grid task (mean: 23.1%) ($t(33)=7.69$, $p<0.0001$). Therefore it is likely that for general speech elicitation purposes, the Diapix task is more efficient than the Grid task.

**Summary** Altogether, the Grid task was indeed considered a less complex task by the participants, and the participants were more successful in completing it than the Diapix task. The types of errors elicited in the final finished grids also implied that the target contrasts selected for inclusion in the Grid task were ones that the participants found difficult to perceive and produce. Finally, although the Grid task was found to fulfil both its main aims, the analysis on the proportion of time spent speaking in each task demonstrated that the Diapix task was, nonetheless, more efficient in eliciting speech from participants than the Grid task.

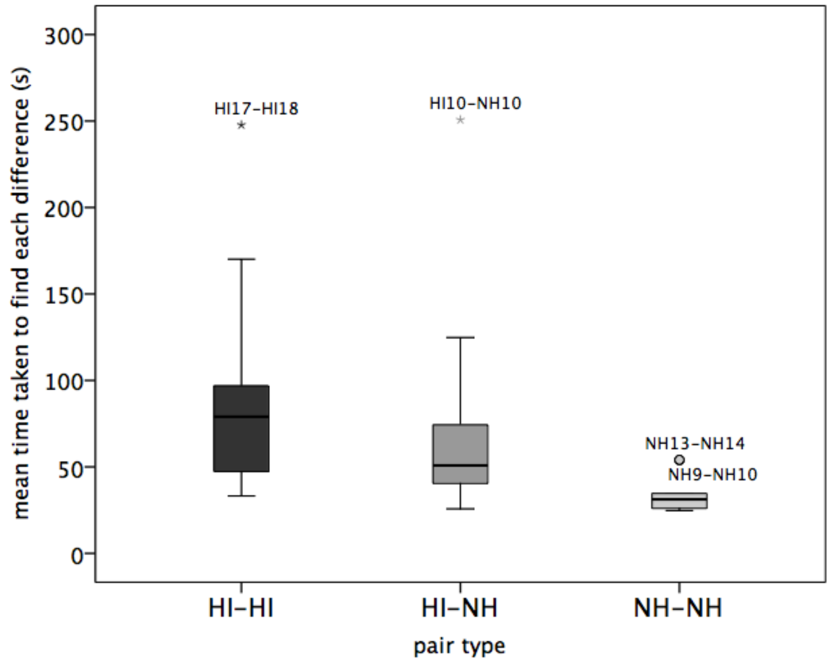
### 5.3.2 Speaking condition and transaction time

To explore whether conditions involving a HI interlocutor were more difficult for participant pairs than those involving a NH interlocutor, the transaction times for each of the three participant pair types were analysed in both the Diapix and Grid tasks.

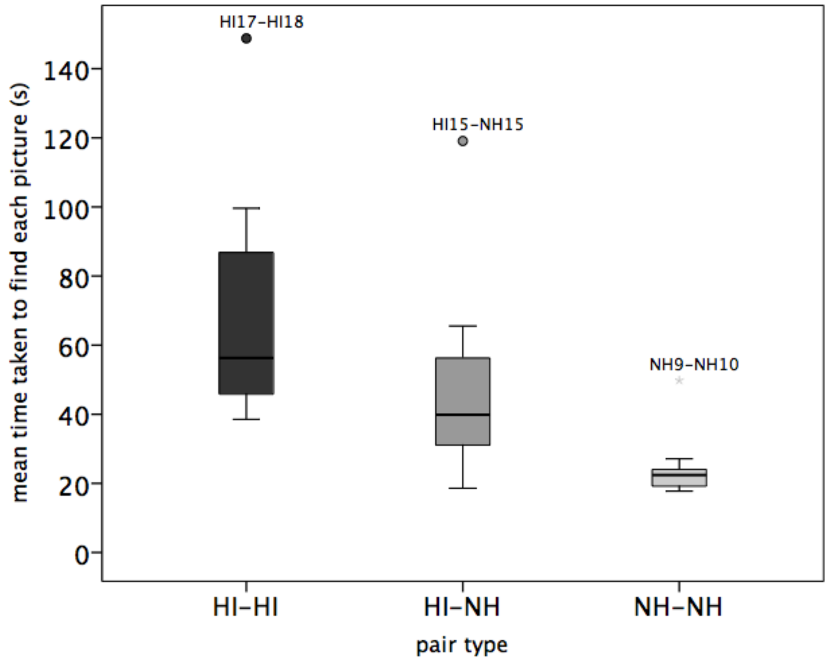
**Diapix transaction time** As a Shapiro-Wilk test showed that Diapix transaction time was non-normally distributed, Wilcoxon rank sum tests were used to investigate whether the transaction times of each pair type differed from each other. The NH-NH pairs took less time to find each difference (mean: 34.4s) than HI-NH pairs (mean: 68.2s) (W=112, p=0.02, effect size r=-0.45). However, there was no difference in transaction times between HI-NH pairs and HI-HI pairs (mean: 97.6s) (W=93, p=0.25) (see Figure 5.5a). In the Diapix task, therefore, the difficulty of the task increased for NH participants when completing the task with a HI friend compared to a NH friend, but task difficulty did not differ between HI-NH and HI-HI pairs.

**Grid transaction time** Grid transaction times were also non-normally distributed. Wilcoxon rank sum tests revealed that in the Grid task, NH-NH pairs took less time to find each grid picture correctly (mean: 25.0s) than did HI-NH pairs (mean: 45.5s) (W=128, p=0.004, effect size r=-0.56) who, in turn, found each picture more quickly than did HI-HI pairs (mean: 70.6s) (W=116, p=0.03, effect size r=-0.42) (see Figure 5.5b). Therefore, unlike in the Diapix task, each increase in the number of HI interlocutors participating in the task led to more time being spent on completing the Grid task correctly.

**Transaction time and speech perception skills** It is possible that the reason for HI-HI participants having greater difficulty than HI-NH pairs in the Grid task is that in HI-HI pairs, both interlocutors may be impaired in their perception of the phoneme contrasts which are crucial for Grid task completion. To explore this possibility, a Pearson's correlation was run on Grid transaction time and mean pair VCV score. The two were negatively correlated with each other (n=35; p<0.001, r=-0.614, $R^2$=0.376) – the better the pair's combined VCV perception score, the more quickly they were able to complete each square in the Grid task (see Figure 5.6). This implies that the ability to perceive the phoneme contrasts in the Grid task is an essential component of being able to rapidly complete the task.




(a) Diapix task (n: HI-HI=9; HI-NH=16; NH-NH=9).




(b) Grid task (n: HI-HI=9; HI-NH=17; NH-NH=9).

Figure 5.5: Boxplots showing mean task transaction times for each pair in the Diapix (a) and Grid (b) tasks.

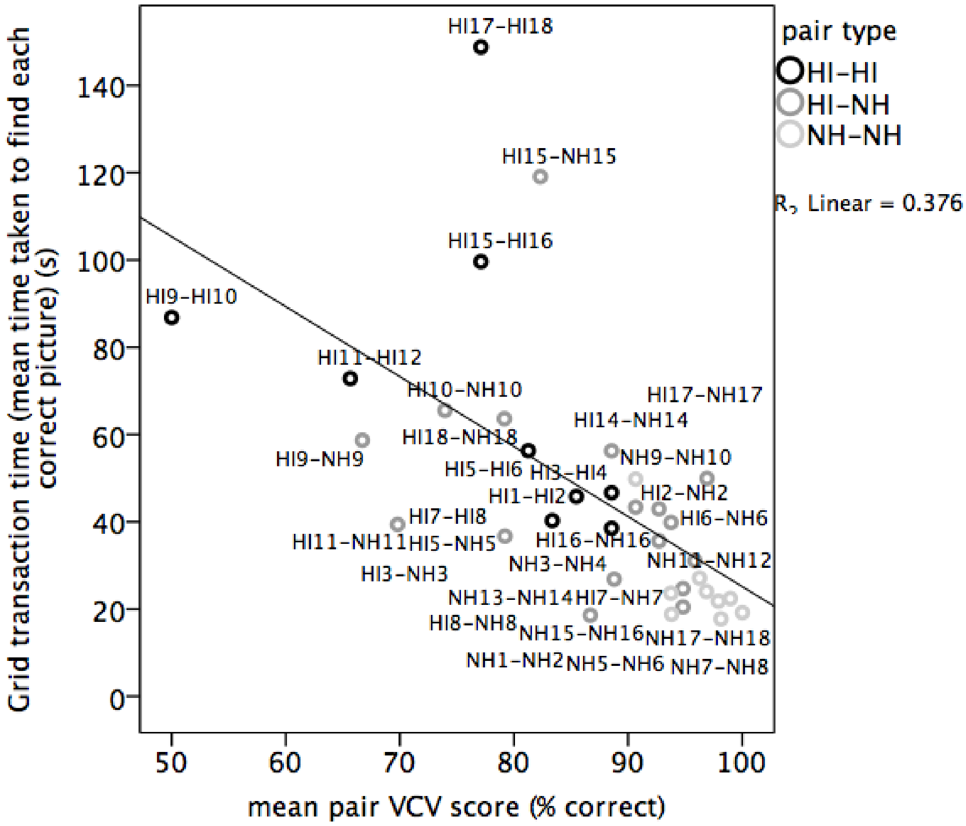



Figure 5.6: Scatterplot of the mean time taken to find each picture in the Grid task and the mean VCV score for each pair.

**Correlation between transaction times** To investigate whether, for each pair, difficulty in one task was related to difficulty in the other task, a correlation was run between Diapix and Grid transaction time measures per pair. The transaction times of the two tasks were highly correlated with each other (Spearman's rho=0.86, S=914, p<0.0001) (see Figure 5.7), suggesting that, for each pair, interactional difficulty was, nonetheless, more pair-specific than task-specific.

**Summary** Overall, it was found that the greater the number of HI interlocutors involved in the interaction, the more difficult it was to complete the communication tasks. This was particularly the case in the Grid task, likely due to the importance in the task of being able to perceive phoneme contrasts accurately, which, as shown in chapter 4, the HI participants perform more poorly on com-

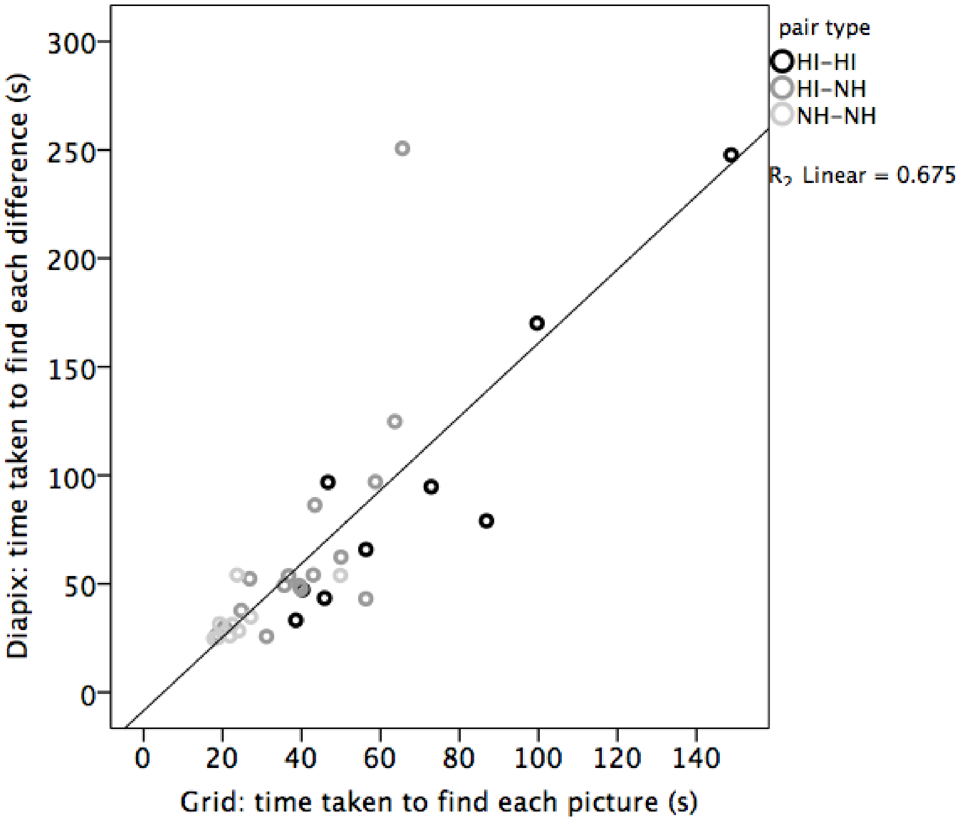



Figure 5.7: Mean transaction time in the Diapix and Grid tasks for each pair (n=34).

pared to the NH participants.

## 5.4 Discussion

In summary, the Grid task was found to be successful in eliciting frequent repetitions of minimal pair keywords and target phonemes, and the number of tokens elicited was found to be sufficient for an analysis of the segmental variability of the participants' speech to be made. Additionally, it was confirmed that the Grid and Diapix tasks in this study are complementary in enabling an analysis to be made of children's speech adaptations in both an easier and a more complex task. An analysis of the proportion of time spent speaking in each of the tasks demonstrated, nonetheless, that the Diapix task is more efficient in eliciting spontaneous speech from speakers than the Grid task – this may be due to both interlocutors in the Diapix task being able to contribute to the conversation without predefined roles. Additionally, in the Grid task, time is likely spent on several task-specific actions, such as finding the correct stack of pictures in the tray, and looking at the picture alternatives – and these parts of the interaction may be conducted in silence.

Interestingly, the Grid task was found to be better able to discriminate between HI-NH and HI-HI pairs in terms of task transaction time than the Diapix task – likely due to the importance of accurate phoneme contrast perception in the Grid task. Indeed, participants' keyword errors which occurred in even the final completed grids demonstrated that some participants in the task encountered major miscommunications, mainly in two target contrasts, which were left unrepaired. Nevertheless, a highly significant correlation between the Grid and Diapix tasks showed that the performance of pairs was more pair-specific than task-specific.

Overall, the greater difficulty that pairs have in conditions involving a HI interlocutor is likely to lead to speakers having to make changes to their speech and language to adapt to the needs of their interlocutor in these conditions. These adaptations are explored in the following chapters.

# Chapter 6

# Global Acoustic-Phonetic Adaptations

## 6.1 Introduction

In this chapter, measures of F0 median and range, speech intensity, articulation rate and pausing are taken to explore whether NH and HI children enhance the global acoustic-phonetic aspects of their speech when talking with a HI interlocutor. As discussed in section 1.3.1, due to the reduced acoustic input received by HI children, an increased speech intensity in speech directed to HI children may enhance important speech cues. An increase in intensity may be accompanied by an increase in F0 median (Titze, 1989), which however may not be beneficial to HI listeners as such. Increasing F0 range in HI-directed speech may make parts of the speech signal more salient to HI listeners by highlighting important words – although it may be less beneficial to users of CIs which only transmit F0 information weakly (e.g., Kuo *et al.*, 2008). A decreased articulation rate may allow the listener more time to process the signal, and an increased number of pauses is likely to assist the HI listener in parsing utterances (Cooke *et al.*, 2014a).

It is unclear, however, whether NH children are able to adapt the global acoustic-phonetic characteristics of their speech to a HI peer. As discussed in section 1.2.2, NH children's global speech production differs from adults' until at least the early teenage years (e.g., Hazan *et al.*, submitted; Lee *et al.*, 1999), and displays greater within-speaker variability probably due to immature speech motor control and inexperience in speech production (e.g., Koenig *et al.*, 2008; Lee *et al.*, 1999), but this may not, nonetheless, prevent them from making adaptations on the global acoustic-phonetic level (Hazan *et al.*, submitted). On the other hand, in section 1.3.4, it was demonstrated that especially younger HI children may be delayed in taking their listener's perspective and in detecting and responding to listener feedback. Section 1.2.2 showed that HI children with more severe hearing losses may have problems with speech motor control (e.g., Niparko, 2009), and may produce greater variability in speech (e.g., Allen and Arndorfer, 2000) than NH peers. Therefore it seems likely that HI children will not be as adept at adapting the global acoustic-phonetic aspects of their speech to the needs of a HI listener as NH peers, although no previous studies have investigated the issue directly.

This chapter therefore investigates NH and HI children's abilities at adapting the global aspects of their speech to a HI listener, as well as whether the difficulty of the task affects the extent to which adaptations are made to HI listeners. According to the Dual-Process model (Bard *et al.*, 2000), tasks which require greater cognitive load are likely to lead to speakers making fewer adaptations to their interlocutor. Therefore, it is expected that both NH and HI speakers will make fewer adaptations to their HI listener in the more difficult Diapix task than in the simpler Grid task.

## 6.2 File processing

Measures were calculated for each file separately to obtain one value per file per participant per condition, for each task individually. For all measures, any part of the speech signal containing the interlocutor's speech was not analysed. Any partially spoken words, simultaneously spoken words, unintelligible words, words spoken in noise and silences were excluded from analysis. For all measures except articulation rate[1], any outliers over or below 3 standard deviations from the mean of each participant group per condition were excluded from analysis.

[1]For this measure, outliers were not excluded as all words had been manually checked.

### 6.2.1 F0 measures

The fundamental frequency (F0) in all files was measured using a Praat script. The script calculated the median F0 and 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), the median value was used to ensure that erroneous F0 calculations did not influence the measure to a great extent, while semitones were used as an attempt to normalise for speaker and gender across participants[1]. The accuracy of the automatic F0 estimation was visually checked for a subset of the files, and no major problems were detected, except for two male participants' (NH13, NH2) F0 values, which were excluded as outliers. F0 analyses were not separated by gender, as previous studies show that significant gender differences in F0 appear after approximately age 12 years (Hollien, Green, and Massey, 1994) and in this study, only two male speakers (one NH, one HI) who were over age 13 were included in the analyses of F0 measures. Therefore no major gender differences in F0 values between genders were anticipated.

### 6.2.2 Intensity

For each file, speech intensity was measured using a Praat script. 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 3k Hz. The mean energy per file between those frequencies was then calculated.

### 6.2.3 Articulation rate

To analyse articulation rate, 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 software (Rinker, 2013) on R (R Core Team, 2014). The total

[1]Conversion to semitones involves a logarithmic transformation of the scale, which more accurately reflects human perception of pitch.

number of syllables was divided by the total duration of words to determine the number of syllables produced per second in each file. Although the calculation of the number of syllables in the files correspond to the transcribed orthographic syllables, rather than actual spoken syllables, this approach was taken as orthographic syllables match the intended, 'target' production of syllables within a word.

### 6.2.4 Pausing

As discussed in chapter 3, during transcription, any within-speaker pauses over 500ms in length which were not interrupted by the interlocutor were tagged as 'SIL'. To prevent the inclusion of pauses which occurred due to task-related factors[1], in the current analysis, the maximum length of a SIL pause was determined as 4 seconds. The number of SIL pauses in each file was calculated using a Praat script. This measure was normalised to file length by dividing the number of SILs by the number of total words in the file (excluding simultaneously spoken words). Only the recordings from the Diapix task were used for this measure, as the distinction between within-speaker pauses and silence due to task demands (such as looking for the correct picture-cards) was often unclear in the Grid task.

## 6.3 Results

### 6.3.1 Statistical approach

Linear mixed effects models were used for most of the statistical analyses reported in the following chapters in this thesis. These models are useful because, unlike traditional models such as the ANOVA, linear mixed effects models are able to model not only the effects which influence the mean and which are of interest to the analyst (fixed effects) but also those that introduce variance into the data but are not variables of interest (random effects). This leads to more powerful statistical models (Crawley, 2007). Linear mixed effects models are also able to deal with data which violate traditional models' assumptions on independence of

[1]Such as the circling of a difference during the Diapix task.

all data points, enabling the inclusion of several data points from each participant in a repeated-measures design, as in the current study. Additionally, linear mixed effects models are able to deal with missing data without deleting existing data points, which traditional models do (Field, Miles, and Field, 2012).

For each global acoustic-phonetic measure in this chapter, the lmer function in the lme4 package for R (R Core Team, 2014) was used to exclude ineffective random factors according to the Akaike Information Criterion (AIC). Then, the same linear mixed effects model approach as in Pettinato *et al.* (2016) was used - the lme function in the nlme package was used to choose the best-fitting model with a bottom-up hierarchical approach, in which each predictor is added one-by-one to the baseline model (Field *et al.*, 2012). The fixed and random factors included in each analysis are detailed in the sections below. Because of their robustness, t-tests with Bonferroni correction were used for all post-hoc tests in which the data was normally or near-normally distributed. For very non-normally distributed data, Wilcoxon rank-sum tests and Wilcoxon signed-rank tests, with Bonferroni correction, were used. The Pearson correlation coefficient was used to calculate the effect size for significant findings[1]. For normally distributed data, the effect size r was calculated using the rcontrast function provided in Field *et al.* (2012) in R. For non-normally distributed data, the same approach as in Field *et al.* (2012) was used to calculate r from the wilcox.test function in R.

To examine whether the NH and HI participants enhanced the acoustic-phonetic and linguistic characteristics of their speech when talking to a HI interlocutor, their speech production in the NH-directed and HI-directed conditions was compared. In addition, to explore the influence of the difficulty of the task used to elicit speech, we investigated whether there were any differences in the speech enhancements made in the Grid and Diapix tasks.

Therefore, for the following analyses, speaker hearing status ('spHstatus', [HI, NH]), listener hearing status ('directed', [HI-directed, NH-directed]) and type of task ('task', [Grid, Diapix]) were included as fixed factors in the model[2]. Unless

[1] Frequently used interpretations for r as an effect size are as follows: $0.1<r<0.3$: small effect; $0.3<r<0.5$: medium effect; $r>0.5$: large effect; the same numbers apply for the negative. However, these interpretations are to be used with caution (Field *et al.*, 2012).

[2] With the exception of the pausing measure, for which only the Diapix task was used (see section 6.2.4), and therefore only the fixed factors of speaker hearing status and listener hearing

otherwise specified, speaker, task number and age as a continuous variable were treated as random factors[1]. A significant interaction between speaker hearing status and listener hearing status would imply that the NH and HI participants use different strategies when speaking with a HI interlocutor. A main effect of task or any interactions involving task would indicate that the type of spontaneous speech task had an effect on the speech produced.

### 6.3.2 F0 median (semitones)

The final model included speaker hearing status, listener hearing status and task as fixed factors, and age as the random factor. None of the interactions were significant. There was no significant main effect of task or of speaker hearing status, indicating that median F0 values were similar across HI and NH participants, and across tasks. There was a significant main effect of listener hearing status ($\chi 2(4) = 110.17$, $p<0.0001$, r=-0.44) - participants increased their F0 median when talking to a HI interlocutor (mean: 94.57 st) compared to when speaking to a NH interlocutor (mean: 93.76 st). Therefore the strategy of increasing F0 median in HI-directed speech was used similarly by both NH and HI participants, with no effect of the type of task used to elicit the speech (see Table 6.1).

### 6.3.3 F0 range (semitones)

In the final model, speaker hearing status, listener hearing status and task were included as fixed factors, and speaker and task number were used as random factors. There were no significant interactions, although the interaction between speaker hearing status and task was near-significant ($\chi 2(9) = 3.37$; p=0.066, n.s.). There was no main effect of task. A significant main effect of speaker hearing status ($\chi 2(6) = 11.86$, $p<0.001$, r=-0.40) was found, with HI participants using a wider F0 range (mean: 3.83 st) than NH participants (mean: 2.98 st). There was also a main effect of listener hearing status ($\chi 2(5) = 52.22$, $p<0.0001$, r=-0.34),

status were used.

[1]Note that, due to the relatively small groups analysed in this study, as well as intragroup variability relating to, for example, hearing loss severity and device use, age and gender effects were not analysed in this study. However, both age and gender were relatively well matched between speaker groups.

| measure | task | NH-NHD | NH-HID | HI-NHD | HI-HID |
|---|---|---|---|---|---|
| F0 median (st) | Grid | 93.8 (2.6) | 94.9 (1.9) | 93.7 (5.1) | 94.2 (4.7) |
| | Diapix | 93.8 (2.1) | 94.8 (2.5) | 93.7 (5.3) | 94.4 (5.2) |
| F0 range (st) | Grid | 2.7 (0.7) | 3.3 (0.8) | 3.6 (1.1) | 4.3 (1.2) |
| | Diapix | 2.7 (0.8) | 3.3 (0.9) | 3.3 (0.9) | 4.1 (1.0) |
| ME 1-3k Hz (dB) | Grid | 62.2 (2.1) | 64.0 (2.4) | 64.1 (3.0) | 64.9 (3.1) |
| | Diapix | 61.6 (2.3) | 63.6 (2.5) | 63.6 (3.3) | 65.0 (3.3) |
| artic. rate (syll/s) | Grid | 3.6 (0.5) | 3.5 (0.6) | 3.3 (0.7) | 3.1 (0.5) |
| | Diapix | 4.0 (0.5) | 3.7 (0.5) | 3.4 (0.6) | 3.4 (0.6) |
| prop. of pauses | Grid | - | - | - | - |
| | Diapix | 0.052 (0.025) | 0.048 (0.021) | 0.072 (0.040) | 0.058 (0.028) |

Table 6.1: Means and standard deviations (in brackets) for each global acoustic-phonetic measure. NH-NHD: NH speaker in NH-directed condition; NH-HID: NH speaker in HI-directed condition; HI-NHD: HI speaker in HI-directed condition; HI-HID: HI speaker in HI-directed condition.

which indicated that participants increased their F0 range in HI-directed speech (mean: 3.71 st) compared to NH-directed speech (mean: 3.09 st) (see Figure 6.1 and Table 6.1). These results suggest that the same strategy of increasing F0 range when interacting with a HI interlocutor was used by both groups of participants, regardless of the task.

### 6.3.4 Intensity (ME 1-3k Hz)

The final model included speaker hearing status, listener hearing status and task as fixed factors, and age and task number as random factors. A significant interaction of speaker hearing status and listener hearing status was found ($\chi 2(10) = 8.54$, $p<0.01$) . Paired t-tests revealed that both NH participants (mean NHD: 61.98dB; mean HID: 63.84dB; $t(82)=10.77$, $p<0.0001$, $r=0.77$) and HI participants (mean NHD: 63.94dB; mean HID: 64.91dB; $t(69)=3.04$, $p<0.01$, $r=0.34$) increase the intensity of their speech when talking with a HI friend. Independent t-tests showed that in the NH-directed condition, NH participants have a lower speech intensity than HI participants (mean NH: 61.98dB; mean HI: 63.94dB) ($t(145.73)=-4.83$, $p<0.0001$, $r=0.37$). However, in the HI-directed condition, the speech intensities of NH and HI participants are more similar to each other (mean NH: 63.84dB; mean HI: 64.91dB; $t(140.05)=-2.41$, $p=0.017$, approaching signif-

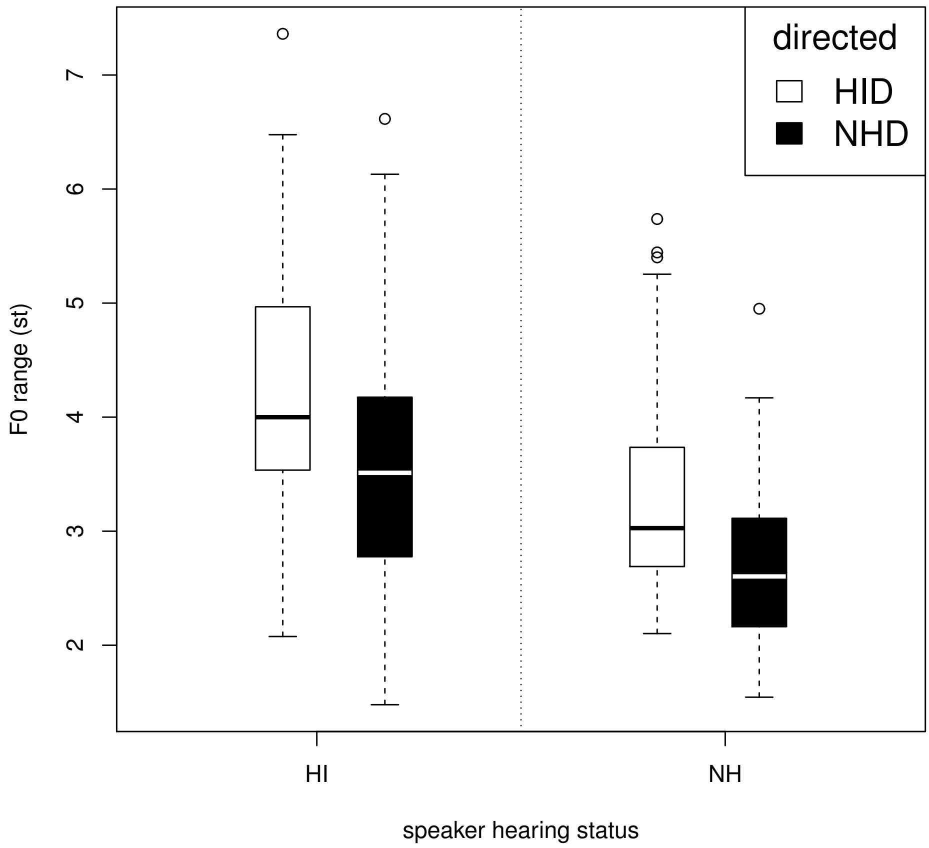



Figure 6.1: F0 range (semitones).

icance[1], r=0.20). These results, as well as the larger effect size found for the NH than HI participants when comparing the NH-directed and HI-directed conditions, together suggest that NH speakers increase the intensity of their speech when talking to a HI interlocutor more than do HI speakers, perhaps because HI speakers already talk more loudly than NH speakers in NH-directed speech. Fig. 6.2 confirms this interpretation. No other interactions were significant. An additional main effect of task ($\chi 2(7) = 7.16$, $p<0.01$) suggests that participants spoke slightly more loudly in the Grid task (mean: 63.67dB) than in the Diapix task (mean: 63.43dB). These findings therefore indicate that both groups increase the

[1]Bonferroni correction set the significance level to 0.0125

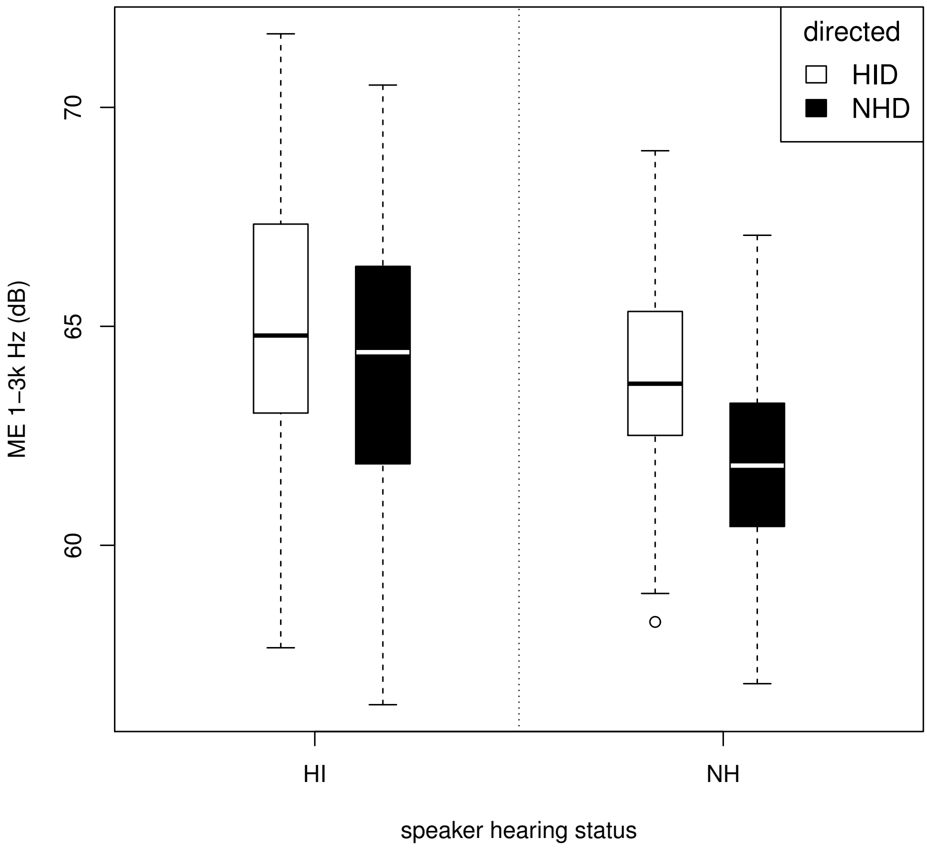



Figure 6.2: Mean energy between 1-3k Hz.

intensity of their speech when talking with a HI interlocutor, but NH speakers do so more than HI speakers.

### 6.3.5 Articulation rate

The fixed factors of speaker hearing status, listener hearing status and task, and the random factor of age were included in the final model. Table 6.1 displays the means and standard deviations of each condition. The interaction between speaker hearing status, listener hearing status and task was just significant ($\chi2(10) = 3.69$, p=0.055). However, after splitting the data per task and investigating the effects of speaker hearing status and listener hearing status separately, it was found that there was no interaction of factors in either the Diapix ($\chi 2(6)$ = 2.88, p=0.09; n.s.) or the Grid ($\chi 2(6)$ = 1.33, p=0.24; n.s.) task. Therefore the next highest interaction, speaker hearing status and task, was investigated using post-hoc tests.

T-tests revealed that both in the Grid task (t(199.56)=4.06, p<0.0001, r=0.28) and the Diapix task (t(112.86)=4.33, p<0.0001, r=0.38), NH participants spoke faster (Grid mean: 3.5 sylls/s; Diapix mean: 3.8 sylls/s) than did HI participants (Grid mean: 3.2 sylls/s; Diapix mean: 3.4 sylls/s). Additionally, both NH and HI participants spoke faster in the Diapix task than in the Grid task (NH: t(62)=-5.31, p<0.0001, r=0.56) (HI: t(59)=-3.09, p<0.0001, r=0.37); the effect sizes demonstrate that the speech rate difference between the two tasks was greater for the NH participants than the HI participants (see Figure 6.3).

There were no significant interactions involving listener hearing status, but a main effect of listener hearing status was found ($\chi 2(4)$ = 18.08, p<0.0001, r=0.10), with participants speaking more slowly to a HI interlocutor (mean: 3.41 sylls/s) than to a NH interlocutor (mean: 3.52 sylls/s). However, the relatively modest effect size demonstrates that this is a fairly small effect (see Figure 6.4).

These results suggest, therefore, that HI participants speak more slowly than NH participants in both tasks. Despite this, the strategy of slightly decreasing speech rate in response to a HI interlocutor is used by both NH and HI participants. The Diapix task seems to enable speakers to speak faster than the Grid task, perhaps due to the interactions in the Grid task being frequently interrupted by task-based requirements, such as looking for the correct picture set in the Grid tray. This effect is more evident in NH participants' than in HI participants' speech.

### 6.3.6 Number of pauses

The fixed factors of speaker hearing status and listener hearing status, and the random factor of speaker were included in the final model. The interaction was not significant. The main effect of speaker hearing status was only near-significant ($\chi 2(5)$ = 3.5, p=0.0613). However, a significant main effect of listener hearing

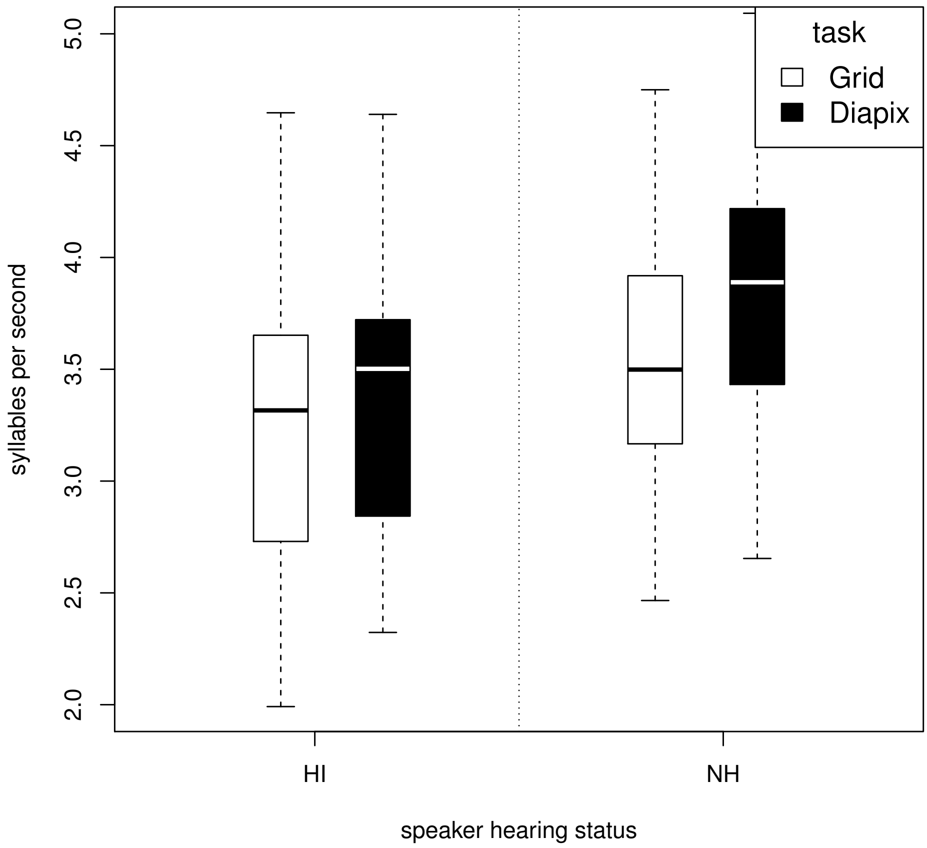



Figure 6.3: Syllable rate for HI and NH participants in each task.

status ($\chi2(4) = 6.05$, $p<0.05$, r=-0.17) was found, with proportionately more within-speaker pauses used in NH-directed (mean: 0.062) than in HI-directed (mean: 0.053) speech. These results suggest that both groups use more pauses in NH-directed than HI-directed speech, contrary to expectations that speakers would increase the number of pauses to benefit HI listeners. However, this finding may be due to the durational limits of within-speaker pauses used in this study (see section 6.2.4) which may discard important shorter or longer pauses used by speakers.

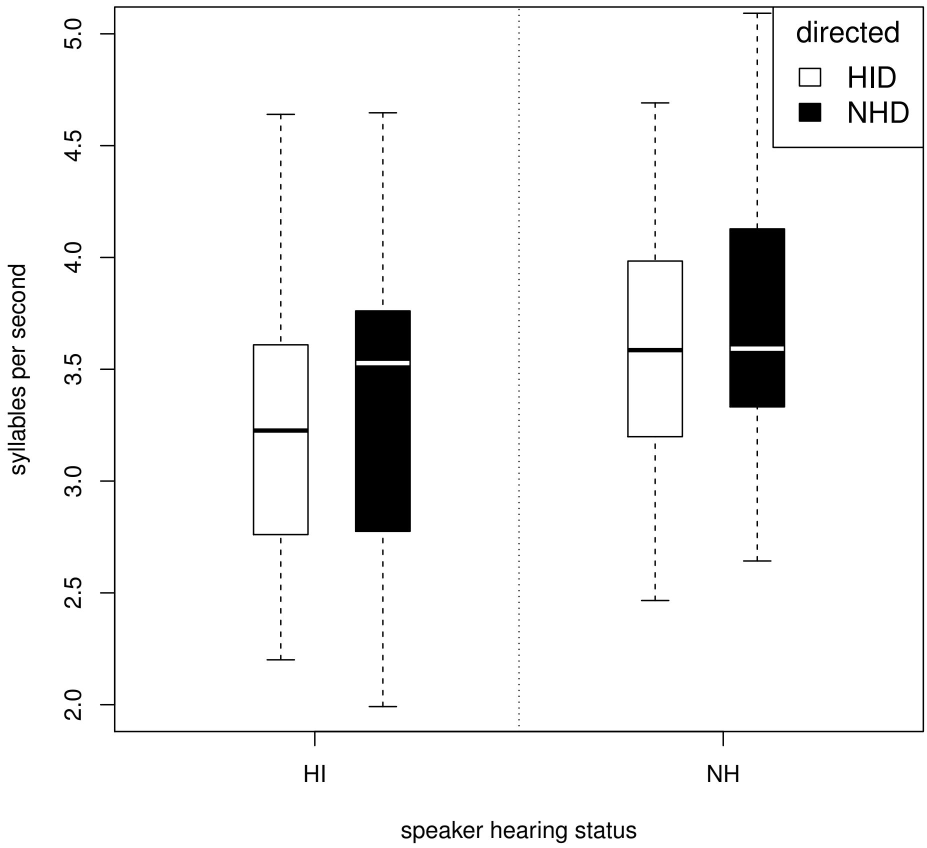



Figure 6.4: Syllable rate for HI and NH participants in NH- and HI-directed conditions

### 6.3.7 Summary

HI participants were found to use a wider F0 range and to speak more loudly and more slowly than NH participants. Both groups of speakers, however, increased their F0 median and F0 range, and slightly decreased their articulation rate when talking with a HI interlocutor. Although both groups increased the intensity of their speech in response to a HI interlocutor, NH participants were found to use this strategy to a greater extent than HI participants. Contrary to expectations, both groups were also found to decrease the number of pauses in their speech when talking with a HI interlocutor than with a NH interlocutor.

The type of task used to elicit speech did not affect these acoustic-phonetic measures to a great extent, with the exception of articulation rate: both groups of participants were found to talk more slowly in the Grid task than in the Diapix task, although this effect was greater in NH than in HI participants. Additionally, speakers were found to talk slightly more loudly in the Grid task than in the Diapix task. No interactions between listener hearing status and task were found, implying that task difficulty did not affect the extent to which the global acoustic-phonetic measures were enhanced in HI-directed speech.

## 6.4 Discussion

### 6.4.1 NH children

**Global acoustic-phonetic adaptations** The findings of this chapter are novel in examining the global acoustic-phonetic strategies used by older NH children when speaking with their HI peers in spontaneous interaction. The results demonstrate that NH children are able to spontaneously adapt to the needs of their HI friend on the global acoustic-phonetic level – they increase their F0 median and F0 range, slightly decrease their articulation rate, and increase the intensity of their speech.

Previous literature indicates that NH children's global speech production is likely to differ from adults' until at least the early teenage years (e.g., Hazan *et al.*, submitted; 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). However, NH older children's less flexible speech production systems do not appear to prevent them from making adaptations to their listener on the global acoustic-phonetic level. These findings are similar to those of Hazan *et al.* (submitted), who found that in the vocoder condition compared to a 'no-barrier' condition, both completed in audio-only conditions, 9- to 14-year-olds completing the Diapix task with peers increased their F0 median and speech intensity, and decreased their articulation rate similarly to adults. These changes were made despite some age groups' global values in the no-barrier condition significantly differing from adults'.

Due to the lack of previous studies examining speech directed to a real HI interlocutor, few direct comparisons can be made to previous literature. However, all the strategies found to be made by NH children in this study are ones that adult speakers in clear speech studies have also been found to use when asked to 'speak clearly as if to a hearing-impaired person' (Smiljanić and Bradlow, 2009). Compared to clear speech studies, however, there are apparent differences in the extent of the modifications made. This is most evident in articulation rate – in adult clear speech studies, speakers decrease their articulation rate on average between 26% to 48% (Picheny *et al.*, 1986; Smiljanić and Bradlow, 2005), but the NH speakers in this study did so by only approximately 5%. Although the extent of the increase in F0 range seems to be similar in instructed clear speech studies (e.g., Smiljanić and Bradlow, 2005; Van Engen *et al.*, 2012) and the current study, in at least some clear speech studies, the amount of intensity change seems to be smaller than in the current study (e.g., under 0.5% in Van Engen *et al.*, 2012, compared to 3.6% in the current study). These discrepancies between the studies may reflect developmental differences between NH children and adults, and may also at least partly be due to the interactive nature of the current study (Hazan and Baker, 2011) – leading speakers to make greater changes to their speech during interlocutor miscommunication, but fewer changes are made when there are no misunderstandings. This 'tension' between speaker effort and communicative efficiency is predicted by Lindblom's H&H model (Lindblom, 1990), and is not reflected in instructed read clear speech studies which do not include real interlocutors.

Although the current study differs significantly from instructed adult clear speech studies both methodologically and in terms of participant age, the finding that speech rate decreases only little in NH children's HI-directed speech is somewhat surprising. It may be that a substantial decrease in articulation rate is not a useful strategy when talking to a HI peer. In two out of the five sources which give instructions to HI persons' communication partners mentioned in section 1.3.1, communication partners are asked not to speak too slowly as it may affect the HI person's ability to lip read (Action on Hearing Loss, 2012; NDCS, 2012b). Indeed, decreasing articulation rate to a great extent may distort the timing of lip patterns – while previous studies show that a 'normal' speaking rate is more favourable for speechreading than faster speaking rates (Massaro, Cohen, and Gesi, 1993), there is some evidence that exaggerated articulation may negatively affect consonant perception, but increase the visual recognition of sentences (Franks, 1979) (cited in Mohammed, 2007). On the other hand, Berger (1972) (cited in Mohammed, 2007) reported that hearing-impaired adults favoured a slightly slower speech rate compared to normal when speechreading. Even without visual speech information, it is possible that a slow articulation rate would not enhance intelligibility for HI listeners – there is evidence that a decreased speech rate does not increase speech intelligibility even to NH listeners in noise (e.g., Cooke, Mayo, and Villegas, 2014b).

Alternatively, a speaker may decrease their speech rate by adding pauses to speech to enhance speech parsing and linguistic processing by the HI listener. Indeed, some studies (e.g., Picheny *et al.*, 1986) have found that speakers increase the number of pauses in clear speech. However, in the current study, a decrease in the number of pauses used in HI-directed compared to NH-directed speech was observed. It is possible that the lower limit in marking a within-speaker pause in this study (500 ms) was not short enough to detect all meaningful pauses in speech – a study by Heldner and Edlund (2010) showed that in the adult Map Task Corpus (Anderson *et al.*, 1991), 56% of all within-speaker pauses were under 500 ms in length – therefore it is likely that the analysis of pausing in the current study was unreliable and did not capture all pausing behaviour by the speakers. On the other hand, it is possible that the decrease in the number of pauses found in HI-directed speech is a consequence of more careful speech production planning by speakers, leading to fewer hesitation-related pauses being produced when talking with a HI friend.

In summary, although previous clear speech and simulated CI studies have shown NH speakers to use similar global acoustic-phonetic strategies to those used by NH children to a real HI listener in the current study, the extent of modifications, especially to articulation rate, seem to differ. Notably, the NH children in the current study frequently interact with their HI peers, and are therefore more likely to use global acoustic-phonetic strategies which enhance communication with their HI friends.

### 6.4.2 HI children

**Global speech characteristics** Overall, compared to NH children, HI children were found to have a greater F0 range, greater speech intensity in the 1 to 3k Hz range, and a slower articulation rate. The second column of Table 6.2 displays the mean % difference between all HI and NH children, and the third column shows the mean % difference between the seven CI users and all NH children, in the NH-directed condition for these global acoustic measures.

The differences between the HI and NH groups are similar to those reported in the previous literature. Although previous studies have investigated F0 production and speech intensity in HI children, they have mostly examined participants' productions of a sustained /ɑ/ for frequency or amplitude variation, rather than F0 range or mean intensity in spontaneous speech, with findings generally showing greater variation in both F0 and amplitude for HI children with profound hearing losses both with and without CIs (e.g., Campisi *et al.*, 2005; Campisi, Low, Papsin, Mount, and Harrison, 2006; Hocevar-Boltezar *et al.*, 2006). Despite methodological differences, the current study thus supports the findings of previous studies showing a wider F0 range and greater speech intensity in HI children's speech, and extends it to spontaneous speech, and imply that HI children are likely to have reduced speech motor control for these voice characteristics.

| **measure** | **all HI %diff** | **CI %diff** |
|---|---|---|
| F0 range | 30.5 (32.4) | 33.6 (41.5) |
| intensity | 3.5 (4.7) | 4.7 (6.3) |
| artic. rate | -9.0 (17.7) | -14.3 (18.3) |

Table 6.2: Measures for which HI children differed significantly from NH children in the NH-directed (NHD) condition (in the Grid and Diapix tasks together). 'all HI %diff'= % mean difference between all HI participants (n=17) and the mean of all NH participants (n=17) in each measure in the NHD condition. 'CI % diff'= % mean difference between HI participants wearing CIs (n=7) and the mean of all NH participants (n=17) in the NHD condition. Standard deviations are presented in parentheses.

For articulation rate, previous studies on sentence reading tasks have shown a mean difference of 10%[1] between 92 8- to 9-year-old HI children who had worn CIs for at least 4 years and predominantly used an oral communication mode, and 24 same-aged NH controls (Uchanski and Geers, 2003) – thus similar to the 14% articulation rate difference for spontaneous speech in CI users compared to NH controls in the current study. The reduced articulation rate of HI children implies that they may be slower at processing and planning speech than their NH peers. Previous studies showing a reduced speech rate for HI children have suggested that much of the difference to NH controls comes from the increased number and length of pausing (Burkholder and Pisoni, 2003; Chuang *et al.*, 2012), which did not show a significant difference between HI and NH participants in the current study. However, as discussed in the previous section, it is likely that the 500 ms lower limit for transcribing within-speaker pauses was not sufficient for observing all pausing behaviour in this study.

**Global acoustic-phonetic adaptations** Surprisingly, despite the above findings of HI children's reduced speech motor control and generally slower speech production planning, when talking to a HI friend, HI children modified the global acoustic-phonetic characteristics of their speech similarly to their NH peers – they increased their median F0 and F0 range and slightly decreased their articulation rate. They also increased the intensity of their speech, although this strategy was used to a lesser extent by HI children than by NH speakers – perhaps due to HI children having a greater speech intensity than NH children already in the NH-directed condition, and therefore being unable to increase their intensity further. This result is contrary to the hypothesis that HI children may display deficits in making global speech modifications to a great extent, due not only to their speech motor control deficits and their reduced auditory feedback, but also the delays found in HI children in taking a listener's perspective and in detecting and responding to feedback (see section 1.3.4).

No known previous studies have explored whether HI children are able to make global acoustic-phonetic adaptations to a listener's needs. The only previous study examining a clinical population's abilities in global acoustic-phonetic lis-

[1] Calculated from Uchanski and Geers's (2003) Table 3 on the duration of the final word in read sentences.

tener adaptation are instructed clear speech studies on adult dysarthric (Parkinson's disease/Multiple Sclerosis) patients (Goberman and Elmer, 2005; Tjaden and Wilding, 2004). In these studies, when asked to 'speak clearly', 'speak loudly' or 'speak slowly' when reading a paragraph, dysarthric patients were able to increase their mean F0 and F0 range, increase the intensity of their speech, and decrease their articulation rate, compared to their 'habitual' speech. These studies show, therefore, that even a serious deficit in speech motor control, such as dysarthria, does not prevent a speaker from being able to make global acoustic-phonetic adaptations to their speech. However, it is unclear whether dysarthric patients would have been able to make these changes without instruction, or in spontaneous conversation – the current study is thus novel in suggesting that speakers from a clinical population with known deficits in speech production are able to spontaneously adapt to the needs of their listener in interaction.

The result here is similar to that in Hazan *et al.* (submitted) in showing that an immature global speech production system does not prevent a speaker from making global adaptations. Similarly, Weppelman *et al.* (2003) demonstrated that even 4-year-olds were able to speak more slowly to an infant than to an adult (although they did not make F0 changes to their speech), and Syrett and Kawahara (2014) showed that 3- to 5-year-olds increased their F0 range and their speech intensity in vowels when teaching new words to a puppet compared to telling an adult about those words.

It is possible that the reason HI children and even younger preschool children are able to make these global acoustic-phonetic adaptations is that global adaptations require little linguistic experience from the speaker. Global acoustic-phonetic adaptations, unlike segmental and linguistic adaptations, do not necessarily require language-specific knowledge for a speaker to be able to make them – there is evidence that fairly proficient non-native speakers of English are able to make similar global modifications to speech as native speakers, both when talking through a vocoder in spontaneous speech (Granlund *et al.*, 2012) as well as when asked to 'speak clearly' in read clear speech (Smiljanić and Bradlow, 2011). Global adaptations may be somewhat automatic modifications made to speech when a speaker encounters a listener in an adverse listening condition, and may thus require both less linguistic experience and less maturity in the speech production system than other types of adaptations.

### 6.4.3 Task difficulty

According to the Dual-Process model (Bard *et al.*, 2000), task difficulty affects the amount of adaptation that a speaker can make to their listener, as more complex tasks require cognitive resources which then cannot be allocated to listener adaptation. Although it was found that speakers' articulation rate was greater in the Diapix than in the Grid task, and the Grid task was spoken with slightly greater speech intensity than the Diapix task – likely both due to the Grid task eliciting shorter and less complex utterances than the Diapix task – speakers were found to make adaptations to a similar extent in both the easier Grid task and the more difficult Diapix task, contrary to the hypothesis. However, it is possible that the two tasks did not differ in difficulty enough for speakers to be unable to allocate enough resources to listener adaptation in the Diapix task. Alternatively, as hypothesised above, global speech modifications may be more automatic processes, and therefore may not require many cognitive resources for modifications to be made. It is possible that segmental or linguistic adaptations, investigated in the following chapters, require greater effort and cognitive processing, as they involve the consultation of language-specific linguistic knowledge.

### 6.4.4 Summary

Altogether, this chapter demonstrated that both NH and HI children are remarkably adept at making global acoustic-phonetic adaptations to their HI listener. However, as discussed above, it is possible that these adaptations are easier for speakers to make than adaptations requiring more linguistic experience and knowledge – thus segmental and linguistic adaptations are investigated in the following chapters 7 and 8. Although the speakers were found to be able to make global adaptations, it is not clear whether these adaptations increased their intelligibility – this aspect is further analysed in chapter 9. Additionally, the individual variability in strategy use is investigated in chapter 10.

# Chapter 7

# Segmental Acoustic-Phonetic Adaptations

## 7.1 Introduction

Some evidence from clear speech studies suggests that, when asked to speak clearly 'as if to a hearing-impaired person', speakers increase the distance between vowel categories by expanding their vowel space, and by increasing between-category distance for vowel or consonant contrasts (e.g., Bradlow, 2002; Ferguson and Kewley-Port, 2002; Kang and Guion, 2008). Enhancing phonetic contrasts by approximating phonetic targets more closely may indeed be helpful to listeners – there is evidence that speakers who have more separable and internally more consistent phonetic categories, i.e. greater distances between phonemes, less within-category dispersion and less overlap between categories, are more intelligible to listeners than those whose categories are less discriminable (Hazan *et al.*, 2013; Newman *et al.*, 2001). However, clear speech studies, as well as speech production studies examining HI children's speech, typically only report on mean measures per category without examining discriminability and variability within categories.

It is unclear whether either NH or HI children are able to make segmental adaptations to a listener's needs. NH children's speech production development has been found to continue until at least early adolescence (e.g., Lee *et al.*, 1999), and they demonstrate greater between-category distance but more variable within-category dispersion in phonetic contrasts compared to adults (Romeo *et al.*, 2013). HI children, on the other hand, exhibit deficits in the production of fine phonetic detail in speech (Pratt and Tye-Murray, 2008) – particularly in producing less distinct phonetic categories and greater variability than NH peers (e.g., Kosky and Boothroyd, 2003; Uchanski and Geers, 2003) – and they may not receive adequate auditory feedback from their own speech to be able to enhance fine-grained segmental distinctions.

This chapter therefore aims to explore whether both NH and HI children enhance the discriminability of phonetic contrasts when in spontaneous task-based conversation with a HI friend compared to a NH friend. Vowel space measures are taken from both the Diapix and Grid tasks to enable a comparison to be made regarding task difficulty. The Grid task is used to elicit spontaneous production of phonetic contrasts which are likely to be difficult for the HI children to both produce and perceive. Clarifying these contrasts in the task had a specific communicative role, as listeners had to perceive the keywords accurately to be able to complete the task. The interlocutor may therefore need to enhance these contrasts in HI-directed speech, but not in NH-directed speech. As well as measuring vowel space area and category means for several different types of phonetic contrasts, we also measure category dispersion and overlap to obtain a more comprehensive assessment of segmental contrast enhancement by children in a realistic speaking situation. If speakers attempt to approximate phonetic targets when speaking clearly, we are likely to find greater between-category distance, less within-category dispersion and less overlap between categories in HI-directed speech compared to NH-directed speech.

## 7.2 File Processing

As in section 6.2, for all measures, the parts of the signal containing the interlocutor's speech were excluded from analysis. Unintelligible words, words spoken in noise or while laughing, words spoken in overlap, partially spoken words and silences were not analysed in any of the measures.

### 7.2.1 Vowel space measures

For the vowel space measures, files from both the Diapix and Grid tasks were used to enable analyses to be done on the task effect. In both tasks, the vowel midpoint F1 and F2 of /i/, /æ/ and /ɔ/ in content words were measured using a Praat script. These vowels were chosen due to being at the extremes of the vowel space for Southern British English, and due to their high frequency of occurrence in the speech elicited here. 14567 vowels in total were extracted from the files. For /æ/ and /ɔ/, the default formant track settings on Praat[1] 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. Then, for each of the three vowels, F1 and F2 values above or below two standard deviations from the mean per person per condition were excluded from analysis. As an attempt to normalise for differences between speakers, an R script was used to transform the values to equivalent rectangular bandwidth (ERB) values. ERB is traditionally classified as a 'vowel-intrinsic' normalisation method (c.f., Adank, Smits, and van Hout, 2004), and as such, is more equivalent to 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.

In total, 13269 /i/, /æ/ and /ɔ/ vowels were used to calculate the vowel space area and F1 and F2 ranges in the tasks. On average, in the Diapix task, 15.9 /æ/, 14.8 /ɔ/ and 29.0 /i/ vowels per speaker per condition were used. In the Grid task, a mean of 30.2 /æ/, 19.7 /ɔ/ and 85.7 /i/ vowels per speaker per condition were used.

#### 7.2.1.1 Vowel space area

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

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

method, by taking the square root of s*((s-a)*(s-b)*(s-c)). One vowel space area measure was obtained per speaker per condition per task.

#### 7.2.1.2 F1 range

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/ from the mean F1 of /æ/, as /i/ is the vowel with the lowest F1 values and /æ/ the vowel with the highest F1 values in this vowel set. One F1 range measure per speaker per condition per task was obtained.

#### 7.2.1.3 F2 range

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 per task 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.

### 7.2.2 Phoneme category distinctions

The aligned and word-level checked TextGrid files were used to find all instances of the keywords in the Grid task (see section 3.2.1.2). Plurals of keywords were included in the analysis, as were keywords for which noise did not occur at the target segment. The target sounds (/p/-/b/, /s/-/ʃ/ and /i/-/ɪ/) in each of the keywords were then manually segmented using Praat. Altogether, 5200 segments extracted from the keywords in the Grid task were included in the analyses. It is acknowledged that in spontaneous speech each of these segments may be influenced by several different factors, but the large amount of data analysed in this study as well as the relative control over the context of elicitation in the Grid task is likely to counteract at least part of this variability.

#### 7.2.2.1 Stop voicing contrast

The voice-onset-time of /p/ and /b/ in the keywords 'pin', 'bin', 'peach', 'beach', 'pea' and 'bee', and the letters 'P' and 'B' was determined as the interval between the first peak of the stop burst and the zero-crossing of the onset of the first glottal cycle for the following vowel. Tokens were excluded if the burst could not be identified, or if the stop was produced as a fricative. Segments were labelled as prevoiced only if voicing continued up to the release of the burst. In total, the VOT of 1953 /p/ and /b/ tokens were calculated using a Praat script. On average, 15.8 /b/ tokens and 18.5 /p/ tokens per speaker were analysed in the HI-directed condition. For the NH-directed condition, an average of 10.9 /b/ tokens and 12.2 /p/ tokens were analysed per speaker.

#### 7.2.2.2 Fricative spectral contrast

The keywords containing the /s/-/ʃ/ contrast ('cell', 'shell', 'seat', 'sheet', 'sack' and 'shack') were manually segmented by determining the part of the fricative interval in which there was no mixed excitation (as in Romeo *et al.*, 2013). Tokens of /s/ and /ʃ/ which were erroneously produced as consonant clusters, voiced fricatives or stops were excluded from analysis[1]. A Praat script then band-pass filtered the file between 300 Hz and 20000 Hz, and determined the centre of gravity for each fricative for the mid 50% of each fricative using DFT spectra[2]. The first spectral moment, corresponding to the spectral centre of gravity for each fricative, was used in this analysis, as it has been found to be the primary cue for category distinctions between /s/ and /ʃ/ in English (e.g., Jongman, Wayland, and Wong, 2000). Altogether, 1625 fricative tokens were used in the analysis. On average, 13.3 /s/ tokens and 12.3 /ʃ/ tokens per speaker were used in the analysis of HI-directed speech. For NH-directed speech, an average of 11.3 /s/ tokens and 10.9 /ʃ/ tokens were analysed per speaker.

[1]Erroneous productions of fricatives were mostly found for certain HI participants, in particular HI7, HI10, HI11 and HI14.

[2]This was done using the 'Get centre of gravity' command in Praat. The measure has been shown to be highly correlated with measures using multitaper spectra (Romeo *et al.*, 2013).

#### 7.2.2.3 Tense-lax distinction

The vowel intervals for the keywords containing the /i/-/ɪ/ distinction ('bean', 'bin', 'peach', 'pitch', 'sheep' and 'ship') were manually segmented by finding the zero crossing of the first glottal cycle of the vowel and the zero crossing at the end of the last full glottal cycle of the vowel. If the boundary between the vowel and a following nasal was unclear, the vowel was excluded from analysis. A Praat script was then used to to determine the F1 and F2 of the midpoint of the vowel. Visual inspection of the vowel formants showed that F2 values varied greatly between participants, especially for the vowel /ɪ/ in male speakers of different ages. Therefore, the Praat script was run on the dataset twice with two different formant tracking settings (females - F2: 2600Hz and 3000Hz; males - F2: 2400Hz and 2800Hz), and the output of the two scripts was compared. If a difference in F2 between the two scripts was found, the vowel was manually checked and corrected. This method led to the checking of 165 /i/ and /ɪ/ vowels (9.7% of all tokens). Vowels with F1 or F2 values over or under two standard deviations from the mean per vowel per speaker per condition were excluded, leaving a total of 1622 vowels to be analysed. The F1 and F2 values of the remaining vowels were normalised to ERB, and, to obtain one spectral measure per vowel, the Euclidean distance between F1 (ERB) and F2 (ERB) was taken for each vowel[1]. The duration of each vowel was measured using a Praat script. Altogether, 1622 vowels were analysed, with a mean of 13.8 /i/ tokens and 13.4 /ɪ/ tokens per speaker in the HI-directed condition, and a mean of 10.3 /i/ tokens and 10.2 /ɪ/ tokens per speaker in the NH-directed condition.

#### 7.2.2.4 Category distinction measures

For each of the target segmental contrasts in sections 7.2.2.1, 7.2.2.2 and 7.2.2.3, an additional four measures were calculated to investigate the category distinctiveness between the /p/-/b/, /s/-/ʃ/ and /i/-/ɪ/phonemes per speaker per condition. These measures have been used in previous studies to quantify the discriminability of phonetic categories (e.g. Romeo *et al.*, 2013).

[1]This was calculated as $\sqrt{(F1 - F2)^2}$.

**Between-category distance** For each phoneme per speaker per condition, the mean value over all tokens was first calculated. Then, the measure of between-category distance was taken as the difference in the mean of phoneme 1 and the mean of phoneme 2.

**Within-category dispersion** The standard deviation of the measure for each phoneme over all tokens for each speaker per condition was determined. Within-category dispersion was calculated as the mean of the standard deviations of the measure for phoneme 1 and phoneme 2.

**Category overlap** Category overlap was measured by subtracting the individual token with the lowest value of phoneme 1 from the highest token value of phoneme 2. A negative value indicates category overlap, while a positive value shows the distance between the extreme values of the two phonemes.

**Category discriminability** Overall category discriminability was calculated as: $(mean.phon1-mean.phon2)*\sqrt{2}/\sqrt{((sd.phon1^2)+(sd.phon2^2))}$ where phon1 = phoneme 1 and phon2 = phoneme 2, mean = mean of all tokens, and sd = standard deviation of all tokens, and it was calculated for each segmental contrast per speaker per condition. This measure therefore takes into account both category distance as well as category dispersion, and is equivalent to sensitivity from signal detection theory (Romeo *et al.*, 2013).

# 7.3 Results

## 7.3.1 Vowel space enhancement

### 7.3.1.1 Statistical Approach

The same statistical approach using linear mixed effects models as in section 6.3.1 was taken. Speaker hearing status ('spHstatus', [HI, NH]), listener hearing status ('directed', [HI-directed, NH-directed]) and task ('task', [Grid, Diapix]) were included as fixed factors in the model, as in chapter 6, and speaker and age as a continuous variable were treated as random factors. An interaction between speaker hearing status and listener hearing status would suggest that NH and HI speakers are using different strategies for enhancing their vowel spaces when speaking with their HI friend compared to speaking with their NH friend. Significant task effects would imply that the type of task influenced the vowel spaces produced by speakers.

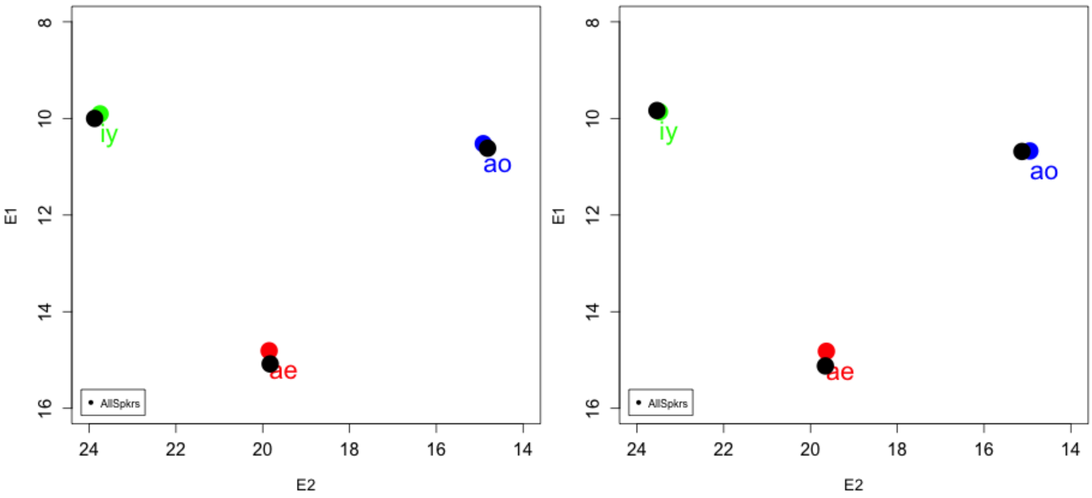
| measure | task | NH-NHD | NH-HID | HI-NHD | HI-HID |
|---|---|---|---|---|---|
| Vowel space area (ERB$^2$) | Grid | 20.91 (2.93 ) | 22.35 (3.77) | 20.06 (4.19) | 20.83 (5.44) |
| | Diapix | 19.68 (3.67) | 20.95 (3.25) | 18.82 (4.70) | 20.05 (5.07) |
| F1 range (ERB) | Grid | 4.95 (0.45) | 5.15 (0.62) | 5.02 (0.72) | 5.32 (0.95) |
| | Diapix | 4.78 (0.68) | 4.96 (0.68) | 4.85 (1.0) | 5.13 (0.76) |
| F2 range (ERB) | Grid | 8.96 (0.63) | 9.13 (0.60) | 8.59 (0.93) | 8.39 (1.18) |
| | Diapix | 8.73 (0.79) | 8.97 (0.51) | 8.40 (0.84) | 8.40 (1.14) |

Table 7.1: Means and standard deviations (in brackets) for each vowel space measure. NH-NHD: NH speaker in NH-directed condition; NH-HID: NH speaker in HI-directed condition; HI-NHD: HI speaker in HI-directed condition; HI-HID: HI speaker in HI-directed condition.

#### 7.3.1.2 Vowel space area (ERB$^2$)

For the vowel space area measure, the fixed factors of speaker hearing status, listener hearing status and task, and the random factor of age were included in the final model. Neither the three-way interaction, nor any of the two-way interactions were significant (all $p>0.66$). The effect of speaker hearing status was not significant ($p=0.49$). However, the effect of listener hearing status was significant ($\chi2(4) = 7.44$, $p<0.01$, $r=0.30$); speakers' vowel spaces were slightly larger when talking with a HI interlocutor (mean: 21.04) than when talking with a NH interlocutor (mean: 19.87) (see Figure 7.1). Surprisingly, therefore, the HI and NH participants did not differ in the size of their vowel spaces, and both groups made their vowel spaces larger when talking with a HI friend.

There was also a significant effect of task ($\chi2(6) = 7.91$, $p<0.01$, $r=0.38$), with larger vowel spaces found in the Grid task (mean: 21.04) than in the Diapix task (mean: 19.87).




(a) NH speakers.

(b) HI speakers.

Figure 7.1: Mean vowels /i/, /æ/ and /ɔ/ in the NH-directed (colour) and HI-directed (black) conditions (E1= F1 (ERB); E2= F2 (ERB)).

#### 7.3.1.3 F1 range (ERB)

The fixed factors of speaker hearing status, listener hearing status and task, and the random factor of age, were included in the final model for F1 range. None of the interactions were significant (all $p>0.50$). The effect of speaker hearing status was also not significant ($p=0.54$). However, there was a significant effect of listener hearing status ($\chi 2(4) = 10.27$, $p<0.01$, $r=0.36$), with greater F1 ranges found in the HI-directed condition (mean: 5.14) compared to the NH-directed condition (mean: 4.90). In terms of F1 range, then, there were no differences between the NH and HI speakers, and both groups increased their F1 range in response to a HI interlocutor (see Figure 7.2).

The effect of task was also significant ($\chi 2(6) = 6.03$, $p<0.05$, $r=0.34$) - F1 range was slightly larger in the Grid task (mean: 5.11) than in the Diapix task (mean: 4.93).

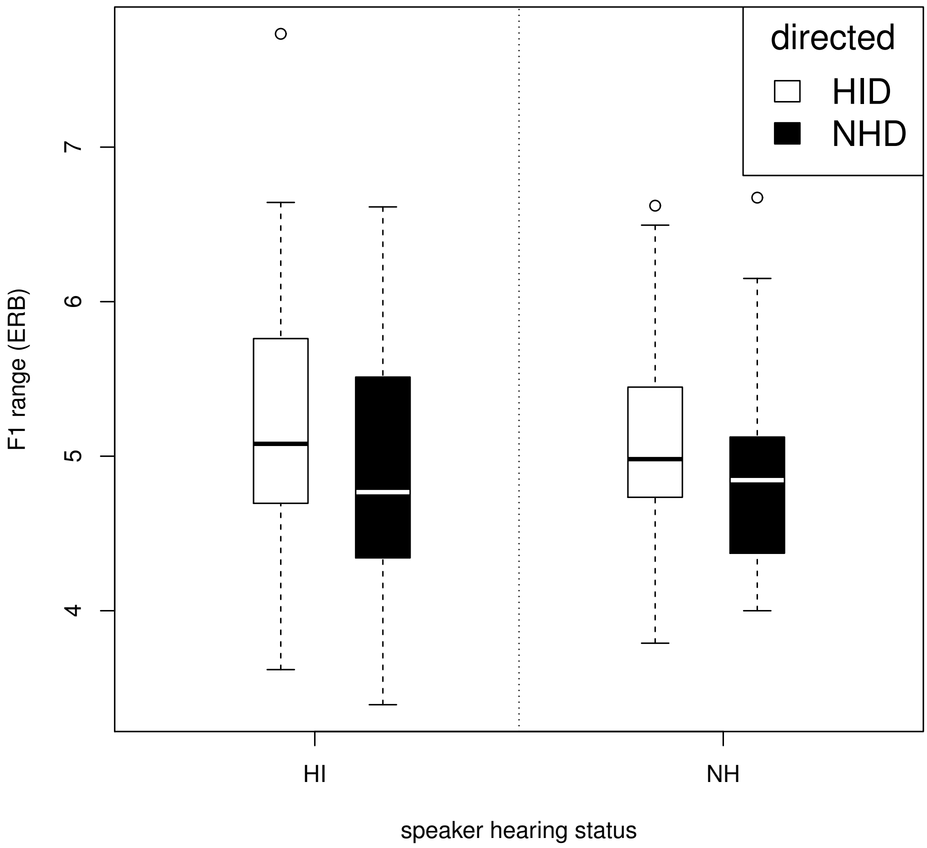



Figure 7.2: F1 range for HI and NH speakers in the two speaking conditions.

#### 7.3.1.4 F2 range (ERB)

The final model for F2 range included the fixed factors of speaker hearing status, listener hearing status and task, and the random factor of age. None of the interactions were significant (all $p>0.12$). The effect of speaker hearing status was significant ($\chi2(5) = 4.09$, $p<0.05$, $r=0.31$); NH speakers were found to have a slightly wider F2 range (mean: 8.95) than HI speakers (mean: 8.45) (see Figure 7.3). The effects of listener hearing status ($p=0.65$) and task ($p=0.18$) were not significant. Therefore, although HI participants were found to have a smaller F2 range than NH participants, neither group of speakers was found to change this aspect of their speech when talking with a HI interlocutor.

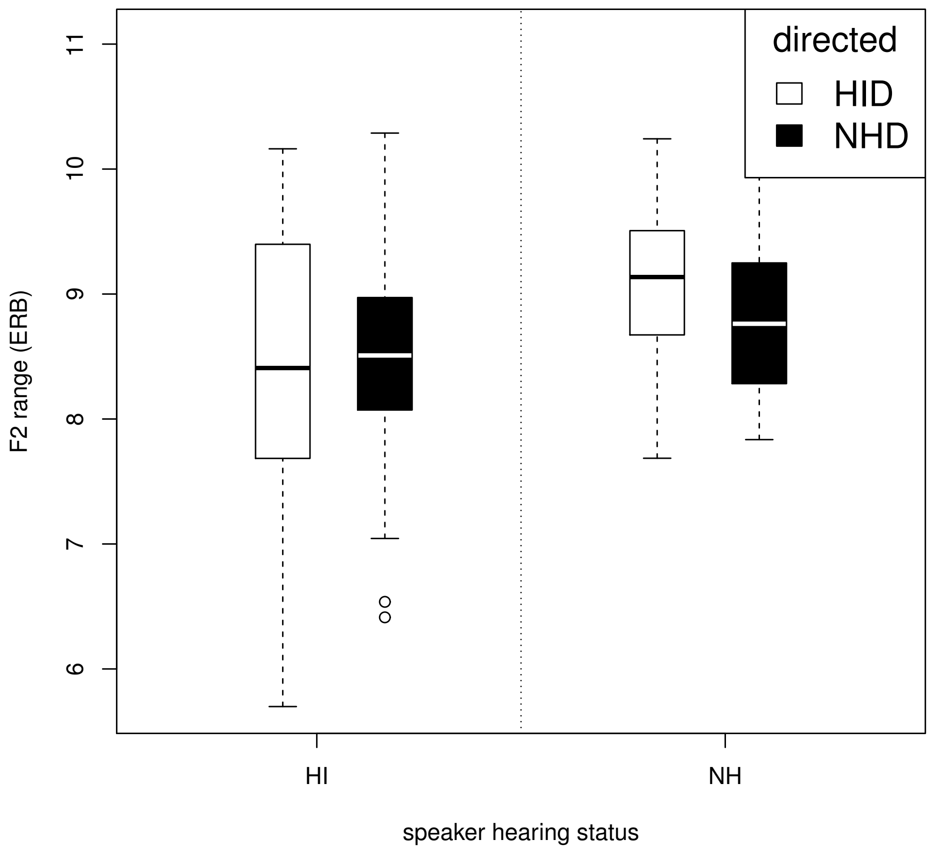



Figure 7.3: F2 range for HI and NH speakers in the two speaking conditions.

#### 7.3.1.5 Summary

In summary, HI and NH speakers were found to differ only in terms of the F2 range of their vowels. However, both groups of speakers were found to slightly increase their vowel space area and F1 range in HI-directed speech. As can also be seen from Figure 7.1, the main strategy for increasing vowel space area by both groups of speakers seemed to be by decreasing F1 for the vowel /æ/. However, F2 range did not change as a function of the listener's hearing status. The type of task also had an effect on vowel space measures - the Grid task was found to elicit a greater vowel space area and a greater F1 range from speakers than the Diapix task – perhaps due to the slower articulation rate in the Grid task giving speakers more time to reach vowel targets.

## 7.3.2 Enhancement of phonetic contrasts

### 7.3.2.1 Statistical approach

As in section 6.3.1, a linear mixed effects model approach was taken for statistical analysis of the three phonetic contrasts elicited from the Grid task. To assess the overall difference between conditions for each contrast in NH and HI speakers, overall statistical analyses of the /p/-/b/ VOT contrast, the /s/-/ʃ/ spectral contrast and the spectral and durational /i/-/ɪ/ contrast were performed. Speaker hearing status (‘spHstatus’, [HI, NH]), listener hearing status (‘directed’, [HI-directed, NH-directed]) and phoneme (’phoneme’, [phoneme1, phoneme2]) were included as fixed factors in each model, and speaker, task number, age and keyword were used as possible random factors in the model. Any significant effects involving phoneme would suggest that in this measure, the contrasting phonemes significantly differ from each other. The overall means and standard deviations can be found in Table 7.2.

To investigate whether speakers make phonetic contrasts more distinct in their speech when talking with a HI interlocutor, the three contrasts were also explored in terms of their category distinctiveness in HI-directed and NH-directed conditions. For these measures, speaker hearing status and listener hearing status were included as fixed factors in the model, and speaker and age were taken as possible random factors.

For both of these analyses, an interaction between speaker hearing status and listener hearing status would imply that HI and NH speakers are using different strategies for enhancing phonetic contrasts in HI-directed speech.

### 7.3.2.2 Voice Onset Time

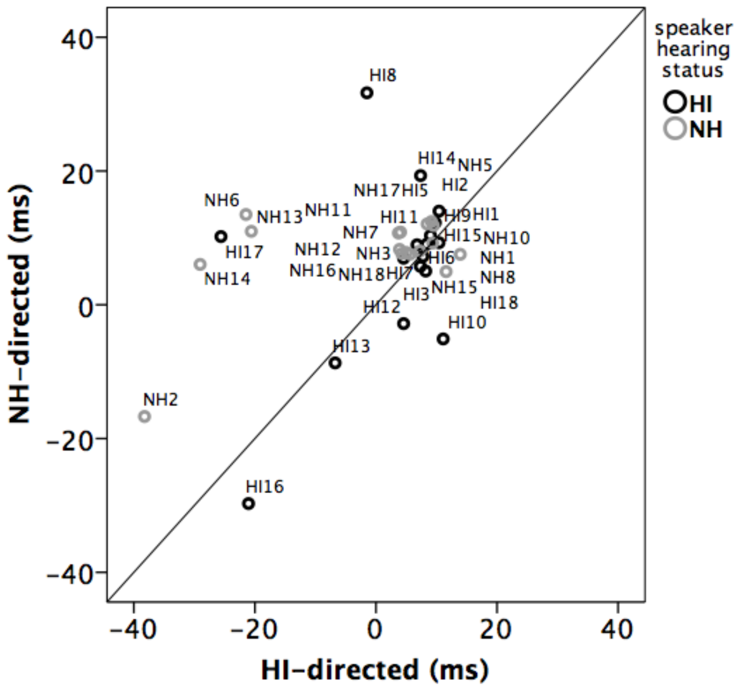
**Overall** For the VOT measure in the /p/-/b/ contrast, the final model included speaker hearing status, listener hearing status and phoneme as fixed factors,

| contrast | phoneme | NH-NHD | NH-HID | HI-NHD | HI-HID |
|---|---|---|---|---|---|
| VOT (ms) | /p/ | 78.3 (32.7) | 77.4 (31.0) | 77.5 (44.5) | 80.6 (42.6) |
| | /b/ | 8.1 (1.7) | 0.0 (28.8) | 5.0 (28.6) | 3.3 (26.0) |
| fricative spectral (Hz) | /s/ | 7762.0 (1783.6 ) | 7986.6 (1382.3) | 6536.3 (1676.6) | 6652.0 (1449.5) |
| | /ʃ/ | 4531.1 (804.0) | 4724.5 (762.1) | 4525.2 (1066.3) | 4688.5 (1028.7) |
| tense/lax Euclidean (ERB) | /i/ | 14.4 (1.0) | 14.2 (1.2) | 13.8 (1.4) | 13.9 (1.5) |
| | /ɪ/ | 10.9 (1.2) | 11.4 (1.3) | 12.0 (1.5) | 12.2 (1.7) |
| tense/lax durational (ms) | /i/ | 196.4 (82.4) | 217.8 (104.8) | 209.5 (97.3) | 253.3 (157.7) |
| | /ɪ/ | 134.0 (47.2) | 138.8 (42.6) | 165.5 (73.3) | 160.3 (73.9) |

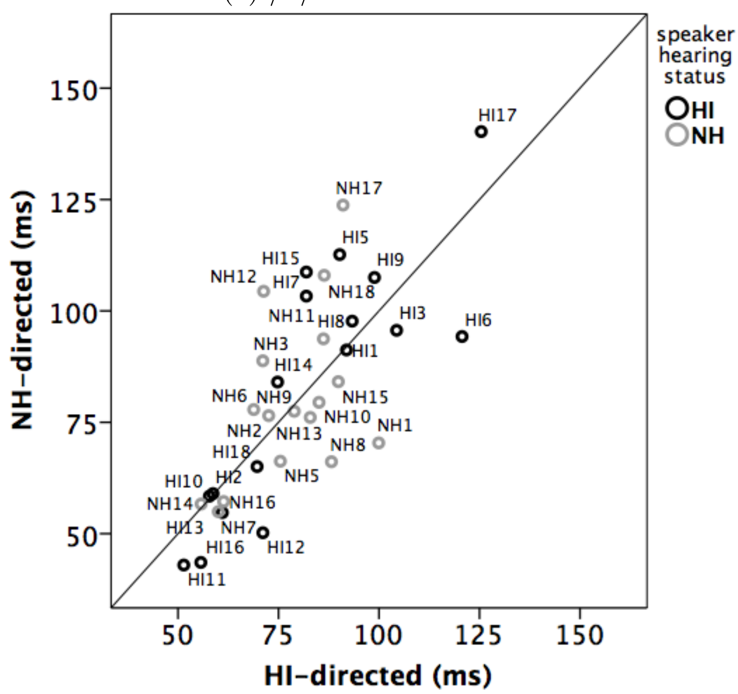
Table 7.2: Means and standard deviations for the phonetic contrasts in NH- and HI-directed conditions for both groups of speakers. NH-NHD: NH speaker in NH-directed condition; NH-HID: NH speaker in HI-directed condition; HI-NHD: HI speaker in HI-directed condition; HI-HID: HI speaker in HI-directed condition.

and speaker and keyword as random factors. None of the three-way or two-way interactions were significant ($p>0.10$), nor were the main effects of speaker hearing status ($p=0.89$) or listener hearing status ($p=0.16$). Only the effect of phoneme was significant ($\chi 2(7) = 404.95$, $p<0.0001$, $r=-0.59$) - as expected, the VOT of /p/ was longer (mean: 78.7ms) than that of /b/ (mean: 3.5ms) (for overall means, see Table 7.2). These results suggest that, despite the HI children having a slower articulation rate compared to NH children, there were no differences between the two groups in their production of the /p/-/b/ VOT contrast. They also imply that speakers did not enhance the VOT distinction in HI-directed speech - this will be explored further in the analyses below.

Figure 7.4 displays the mean VOT values per speaker for each phoneme for NH- and HI-directed conditions. Speakers close to the diagonal line have similar VOT measures in both conditions, while those further from the line produce different mean VOT values per condition. For /b/, only a subset of six speakers (NH2, NH6, NH13, NH14, HI8 and HI17) are located considerably above the diagonal line, therefore indicating lower VOT values for /b/ in their HI-directed speech than in their NH-directed speech. Five of the six speakers do this by pre-voicing. For /p/, most speakers are close to the diagonal line, with approximately equal numbers of speakers on either side of the line. The distribution of VOT values for /p/ and /b/ for NH and HI speakers over both conditions can be found in Figure 7.5.




(a) /b/




(b) /p/

Figure 7.4: Scatterplot of VOT means for /b/ (a) and /p/ (b) per speaker in NH- and HI-directed conditions.

**Between-category distance** The final model for the between-category distance measure for /p/-/b/included speaker hearing status and listener hearing status as fixed factors, and speaker as the random factor. Neither the interaction nor the main effects of speaker hearing status and listener hearing status were significant (all $p>0.21$). This finding suggests that NH and HI speakers have similar distances between their /p/ and /b/ phonetic categories, and that neither group enhances the distance between these two categories in HI-directed speech compared to NH-directed speech.

**Within-category dispersion** For within-category dispersion, the fixed factors of speaker hearing status and listener hearing status, and the random factor of age were included in the final model. None of the effects were significant (all $p>0.14$), implying that the dispersion of /p/ and /b/ categories did not differ between NH and HI participants, and that neither group changed their within-category dispersion when talking with a HI friend.

**Category overlap** The final model for the category overlap measure included speaker hearing status and listener hearing status as fixed factors, and speaker as a random factor. None of the effects were significant (all $p>0.078$), suggesting that the NH and HI participants did not differ in their /p/-/b/ category overlap, nor did either group decrease their category overlap in HI-directed speech, compared to NH-directed speech.

**Category discriminability** For the /p/-/b/ discriminability measure, the final model included speaker hearing status and listener hearing status as fixed factors, and speaker as a random factor. Neither the interaction nor the main effects were significant (all $p>0.30$), implying that there were no differences in category discriminability between NH and HI speakers, and that neither group enhanced the discriminability of their /p/ and /b/ categories when talking with a HI interlocutor.

**Summary** Altogether, it was found that there were no differences in the production of /p/ and /b/ between NH and HI speakers, and neither group enhanced

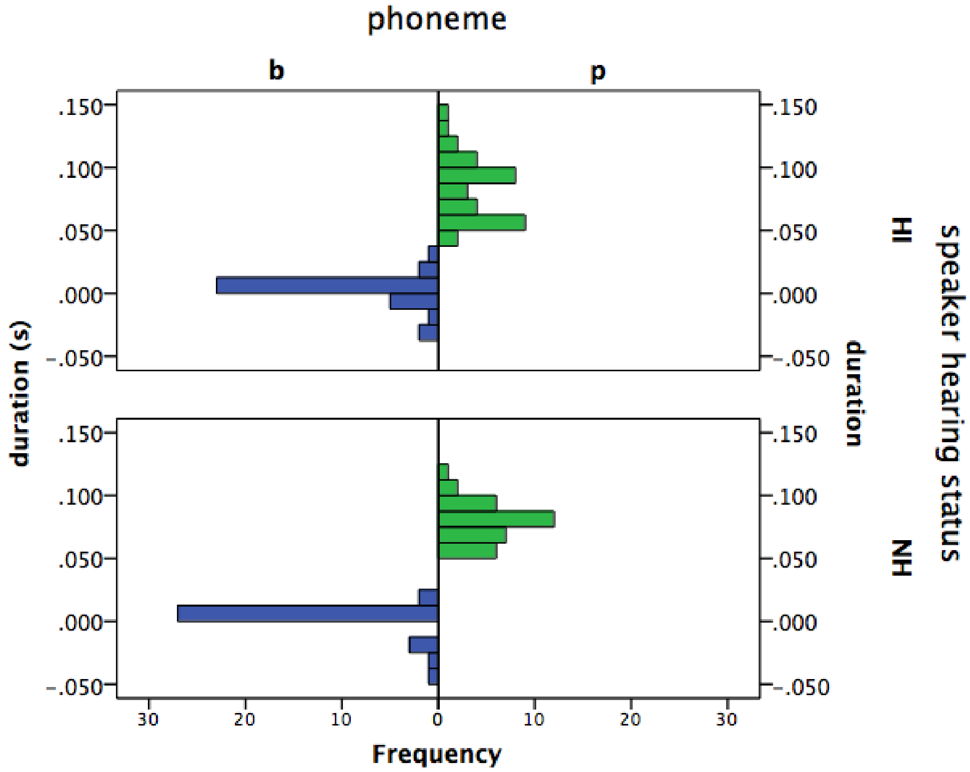



Figure 7.5: The distribution of mean VOT per speaker for /p/ and /b/ in HI speakers (upper) and NH speakers (lower) over both conditions.

the VOT category distinction in response to a HI interlocutor. However, inspection of mean VOT values per speaker in Figure 7.4a suggests that a small subset of speakers prevoiced /b/ in HI-directed speech, and therefore perhaps use this strategy in an attempt to increase the distinctiveness of their /b/ category compared to the /p/ category.

#### 7.3.2.3 Fricative spectral contrast

**Overall** For the /s/-/ʃ/ spectral centre of gravity measure, the final model included speaker hearing status, listener hearing status and phoneme as fixed factors, and age and keyword as random factors. The three-way interaction was not significant, and neither was the interaction between speaker hearing status and listener hearing status or the interaction between listener hearing status and phoneme (all $p>0.44$). The interaction of speaker hearing status and phoneme was significant ($\chi2(9) = 53.85$, p<0.0001, r=-0.30) (see Figure 7.6). A Wilcoxon rank sum post-hoc test showed that NH participants' centre of gravity for /s/ was significantly higher (mean: 7887.7Hz) than HI participants' (mean: 6596.4) (W=128955, p<0.0001, r=-0.30). However, the NH and HI participants' /ʃ/ productions did not significantly differ from one another (mean NH: 4632.3Hz; mean HI: 4612.9Hz) (W=81778, p=0.21). Nevertheless, for both speaker groups, /s/ and /ʃ/ were significantly different from each other (NHs: V=82115, p<0.0001, r=-0.42; HIs: V=68081, p<0.0001, r=-0.36) . In addition, the effect of listener hearing status was significant ($\chi2(5) = 5.67$, p<0.05, r=-0.05), with slightly higher centre of gravity values for fricatives produced in the HI-directed condition (mean: 6099.4Hz) than those produced in the NH-directed condition (mean: 5867.8Hz) - however, the effect size for the difference is very small.

Overall, therefore, HI speakers were found to produce /s/ with a lower centre of gravity than NH speakers. The two groups did not differ in their production of /ʃ/, however. Both groups were found to slightly increase the fricative centre of gravity for both phonemes in HI-directed speech (see Table 7.2). Below we explore in more detail whether speakers enhance the category distinctions between /s/ and /ʃ/ in HI-directed speech. Figure 7.7 also displays the distribution of the mean centre of gravity per speaker for each phoneme.

**Between-category distance** The final model for the between-category distance measure included speaker hearing status and listener hearing status as fixed factors, and age as a random factor. The interaction was not significant (p=0.61). The effect of speaker hearing status was significant ($\chi2(5) = 24.15$, p<0.0001, r=0.69), with greater between-category distances for NH speakers (mean: 3352.6Hz) than for HI speakers (mean: 1802.3Hz). The effect of listener hearing status was not significant (p=0.89), implying that category distances were not enhanced by either group when talking with a HI friend.

**Within-category dispersion** For the within-category dispersion measure, the final model included speaker hearing status and listener hearing status as fixed factors, and age as a random factor. The interaction was not significant. The effect of speaker hearing status was just-significant ($\chi2(5) = 3.57$, p=0.059, r=-

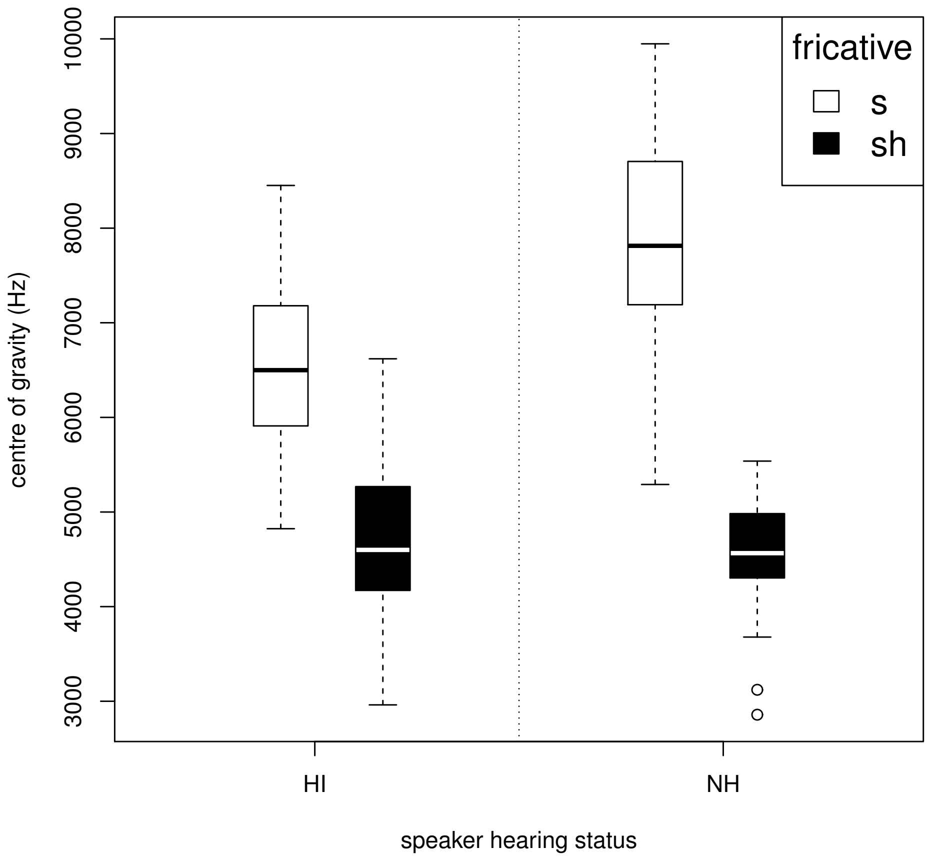



Figure 7.6: Mean centre of gravity for /s/ and /ʃ/ in NH and HI participants.

0.26) - NH speakers had smaller within-category dispersion for /s/ and /ʃ/ (mean: 804.9Hz) than did HI speakers (mean: 987.1Hz). The effect of listener hearing status was not significant (p=0.28), suggesting that neither group changed their within-category dispersion in response to a HI interlocutor.

**Overlap** For the category overlap measure, speaker hearing status and listener hearing status were included as fixed factors, and age was chosen as a random factor in the final model. The interaction was not significant (p=0.53). The effect of speaker hearing status was significant ($\chi 2(5) = 15.40$, p=0.0001, r=0.48), with greater category overlap in HI speakers (mean: -1590.3Hz) than in NH speakers (mean: 380.4Hz). This effect can also be seen in Figure 7.7. However, there was no effect of listener hearing status (p=0.63) - neither group changed the overlap between categories in HI-directed speech.

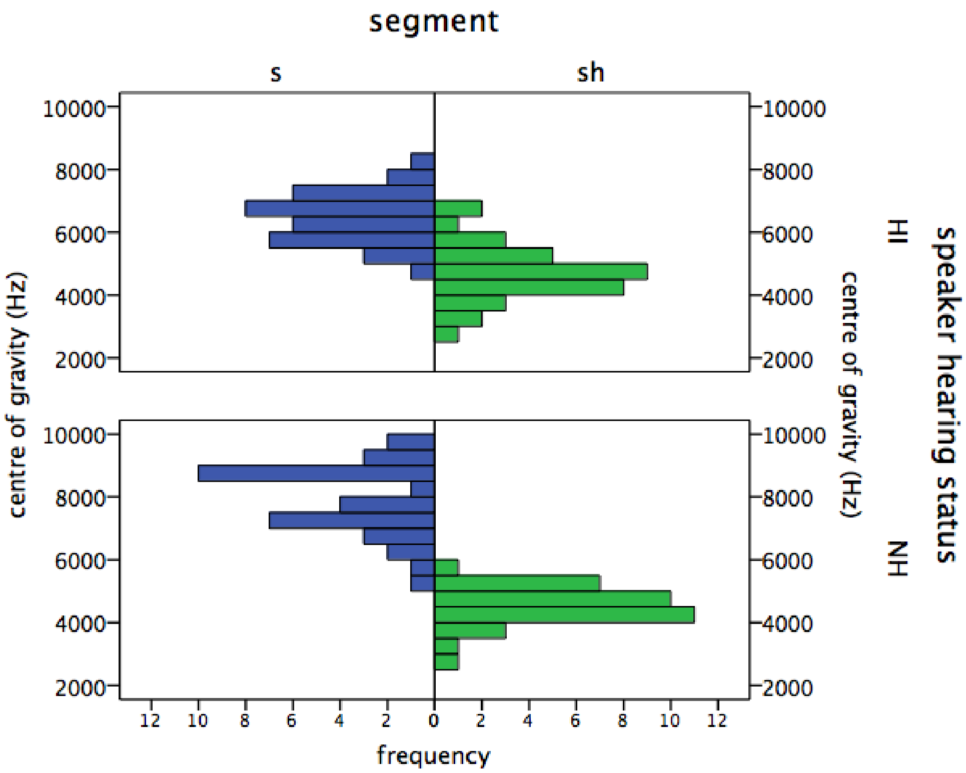



Figure 7.7: The distribution of mean centre of gravity per speaker in /s/ and /ʃ/ for HI speakers (upper) and NH speakers (lower) over both conditions.

**Discriminability** The final model for the category discriminability measure included speaker hearing status and listener hearing status as fixed factors, and age as a random factor. The interaction was not significant (p=0.69), but the effect of speaker hearing status was significant ($\chi 2(5) = 21.66$, p<0.0001, r=-0.62) - there was greater /s/-/ʃ/ category discriminability for NH speakers (mean: 4.62) than HI speakers (mean: 2.03). The effect of listener hearing status was not significant (p=0.25), implying that neither group increased the category discriminability between /s/ and /ʃ/ when talking with a HI interlocutor.

**Summary** HI speakers' spectral centre of gravity for /s/ was found to be lower than NH speakers', but the centre of gravity for /ʃ/ did not differ between groups. NH speakers also have greater between-category distances, less overlap between categories, and more discriminable categories than HI speakers. Not surprisingly, NH speakers were also found to have less within-category dispersion than HI speakers. However, although a general trend was found towards higher centre of gravity values in both fricatives in HI-directed speech, neither group enhanced their /s/-/ʃ/ categories when talking with a HI interlocutor.

#### 7.3.2.4 Tense/Lax distinction: spectral

**Overall** For the F1-F2 Euclidean distance measure, the fixed factors of speaker hearing status, listener hearing status and vowel, and the random factors of speaker and keyword were included in the final model. The three-way interaction was only near-significant ($\chi2(11) = 3.23$, p=0.072). The interaction of speaker hearing status and listener hearing status was not significant ($\chi2(10) = 0.01$, p=0.92), implying that HI and NH speakers were using similar strategies.

The interaction between speaker hearing status and vowel was significant ($\chi2(9) = 33.88$, p<0.0001). Wilcoxon signed rank and Wilcoxon rank sum tests were used as post-hoc tests to explore this interaction further. HI participants' /i/ vowel was found to have a smaller F1-F2 Euclidean distance (mean: 13.9) than NH participants' /i/ vowel (mean: 14.3) (W=99907, p<0.0001, r=-0.12). Similarly, NH participants' /ɪ/ vowel had a smaller Euclidean distance (mean: 11.2) than HI participants' /ɪ/ vowel (mean: 12.1) (W=54197, p<0.0001, r=-0.20). However, the relatively small effect sizes demonstrate that the differences between the two groups were not large. For both HI (V=71919, p<0.0001, r=-0.34) and NH (V=80029, p<0.0001, r=-0.42) participants, /i/ and /ɪ/ differed from each other significantly. As can be seen from the above effect sizes and from Figure 7.8, the /i/ and /ɪ/ vowels were spectrally closer to each other in HI than NH participants' speech.

The interaction between listener hearing status and vowel was also significant ($\chi2(8) = 5.82$, p=0.016). The vowel /i/ did not differ significantly in NH-directed and HI-directed speech (V=27711, p=0.11, n.s.). However, contrary to expecta-

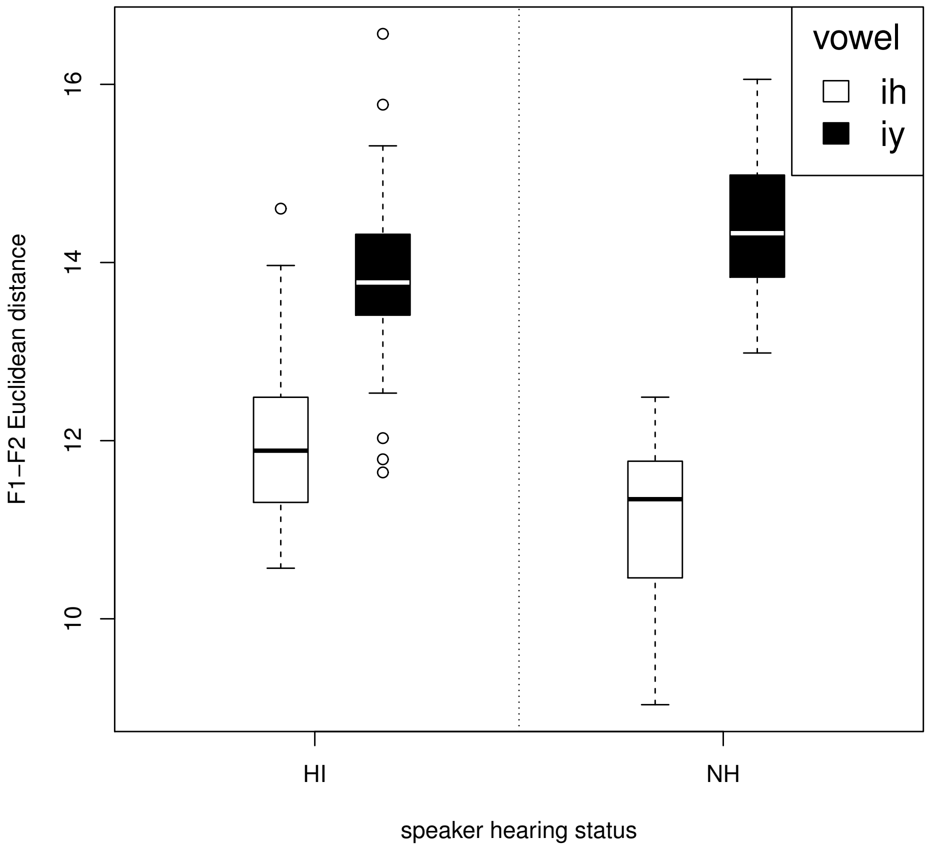



Figure 7.8: Mean F1-F2 Euclidean distance for the /i/ (iy) and /ɪ/ (ih) vowels in the two participant groups.

tions for /ɪ/, the vowel's F1-F2 distance was slightly larger in the HI-directed (mean: 11.79) than the NH-directed (mean: 11.44) condition (V=25402, p=0.01, r= -0.06), albeit with a very low effect size. This suggests that, contrary to expectations, the /ɪ/ vowel becomes spectrally more similar to the /i/ vowel in HI-directed speech. Nevertheless, the difference between /i/ and /ɪ/ was significant in both HI-directed (V=98017.5, p<0.0001, r=-0.40) and NH-directed (V=58564, p<0.0001, r=-0.37) conditions (see Figure 7.9).

In summary, HI speakers were found to have a smaller Euclidean distance for the /i/ vowel, and a greater Euclidean distance for their /ɪ/ vowel compared to

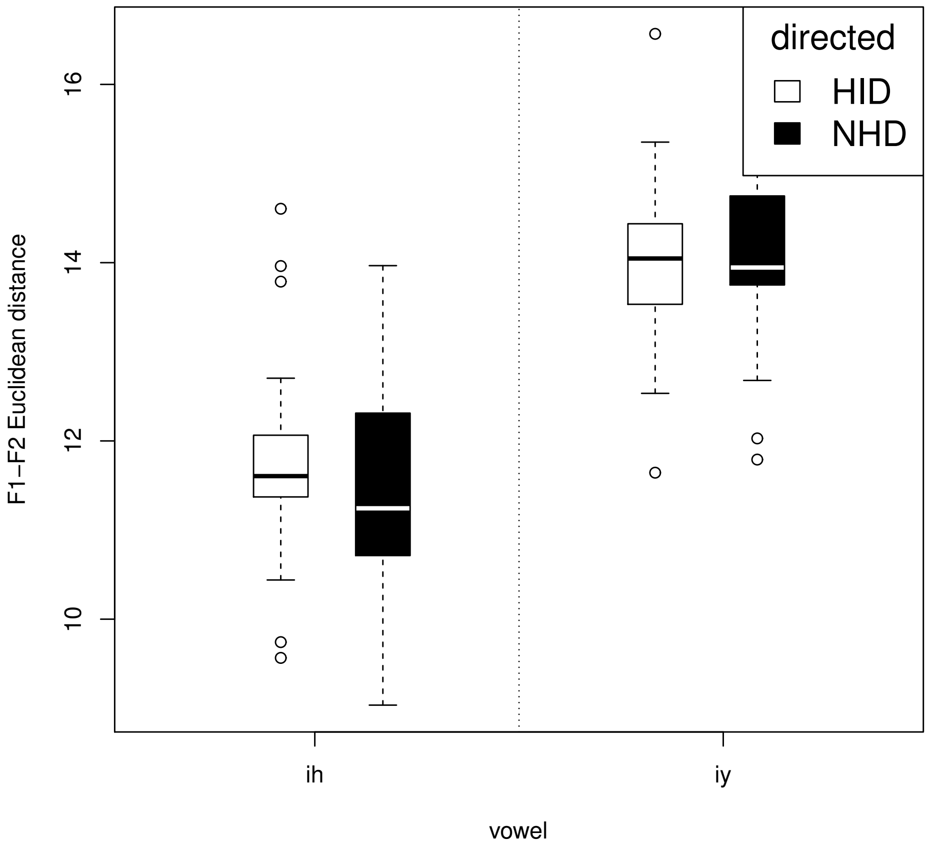



Figure 7.9: Mean F1-F2 Euclidean distance for the /i/ (iy) and /ɪ/ (ih) vowels in the HID and NHD conditions.

NH speakers. Both groups were found to slightly increase the F1-F2 distance for /ɪ/ in the HI-directed condition, although no such difference between conditions was found for /i/. This implies that the two vowels become spectrally more similar to each other in HI-directed speech. The analyses below will investigate in more detail whether the /i/-/ɪ/ distinction is enhanced when talking with a HI interlocutor.

**Between-category distance** The final model for the between-category distance measure included speaker hearing status and listener hearing status as fixed factors, and age as a random factor. The interaction between speaker hearing status and listener hearing status was not significant ($p=0.14$). The main effect of speaker hearing status was significant ($\chi 2(5) = 14.10$, $p<0.001$, $r=0.58$), with greater spectral /i/-/ɪ/ category distances for NH speakers (mean: 3.24) than for HI speakers (mean: 1.68).

There was also a significant main effect of listener hearing status ($\chi 2(4) = 6.91$, $p<0.01$, $r=0.43$) - surprisingly, between-category distances were found to be smaller in HI-directed (mean: 2.28) than NH-directed (mean: 2.64) conditions. This effect is likely to be due to the /ɪ/ vowel becoming spectrally closer to /i/ in HI-directed speech. The lack of a significant interaction between speaker hearing status and listener hearing status implies that both groups make this change when speaking to HI interlocutor.

**Within-category dispersion** The final model for the measure of within-category dispersion included speaker hearing status and listener hearing status as fixed factors, and age as a random factor. The interaction was not significant ($p=0.89$). The main effect of speaker hearing status was significant ($\chi 2(5) = 6.41$, $p<0.05$, $r=-0.33$) - HI speakers were found to have more spectral within-category dispersion for /i/ and /ɪ/ (mean: 1.12) than NH speakers (mean: 0.91). There was also a significant main effect of listener hearing status ($\chi 2(4) = 8.75$, $p<0.01$, $r=-0.35$), with greater within-category dispersion in HI-directed speech (mean: 1.11) than in NH-directed speech (mean: 0.92), likely due to participants changing their speech to their interlocutor during miscomprehensions, but having to make fewer adjustments during parts of the conversation without miscommunications. The lack of a significant interaction shows that this was done by both speaker groups.

**Category overlap** For /i/-/ɪ/ spectral overlap, the final model included speaker hearing status and listener hearing status as fixed factors, and age as a random factor. The interaction was not significant ($p=0.58$). The effect of speaker hearing status was significant ($\chi 2(5) = 13.84$, $p<0.001$, $r=0.50$) - on average, there was no spectral overlap in /i/-/ɪ/ for NH speakers (mean: 0.054), unlike for HI speakers (mean: -1.97) (see Figure 7.10). The effect of listener hearing status was

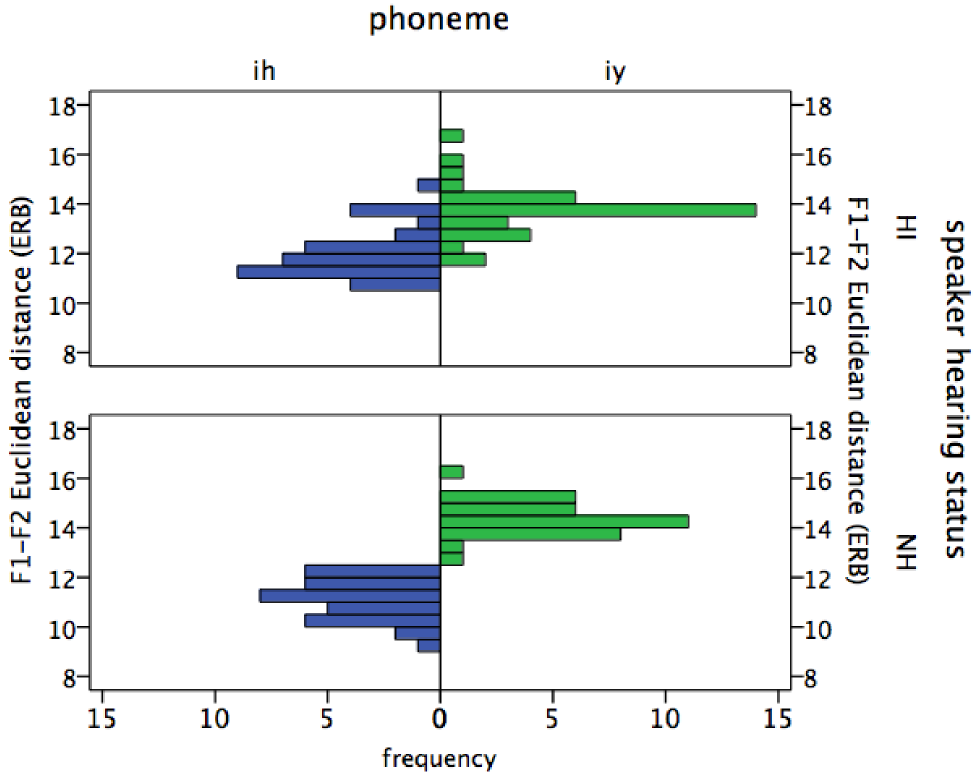



Figure 7.10: Mean F1-F2 Euclidean distance per speaker for the /i/ (iy) and /ɪ/ (ih) vowels for each speaker group.

also significant ($\chi 2(4) = 13.45$, p<0.001, r=0.57), with greater spectral category overlap in the HI-directed condition (mean: -1.60) than in the NH-directed condition (mean: -0.32) . This is likely to be due to the /ɪ/ vowel becoming spectrally closer to /i/ in the HI-directed condition.

**Category discriminability** The final model for /i/-/ɪ/ spectral discriminability included speaker hearing status and listener hearing status as fixed factors, and age as a random factor. There was a significant interaction of speaker hearing status and listener hearing status ($\chi 2(6) = 4.94$, p<0.05). Post-hoc t-tests showed that for NH speakers, category discriminability was significantly lower when talking with a HI interlocutor (mean: 3.08) than when talking with a NH interlocutor (mean: 4.58) (t(16)=-2.96, p<0.01, r=0.60). For HI speakers, there was no difference in spectral category discriminability in HI-directed (mean: 1.56) and NH-directed (mean: 1.77) conditions ($p=0.33$, n.s.), implying that they did not make the same spectral changes as NH speakers. However, in both NH-directed ($t(27.068)=4.79$, $p<0.0001$, $r=0.68$) and HI-directed ($t(31.824)=3.50$, $p<0.01$, $r=0.53$) conditions, NH speakers had greater spectral /i/-/ɪ/ category discriminability than did HI speakers.

**Summary** Altogether, HI speakers were found to have smaller F1-F2 distances in their /i/ vowels, and greater F1-F2 distances in their /ɪ/ vowels than NH speakers. HI speakers also displayed spectrally smaller /i/-/ɪ/ between-category distances, greater within-category dispersion for these vowels, greater /i/-ɪ/ category overlap and less spectral category discriminability than NH speakers. Surprisingly, when talking with a HI interlocutor, both groups of speakers were found to increase the F1-F2 Euclidean distance for /ɪ/, making the vowel spectrally closer to /i/ in that condition. This may be due to speakers ensuring their /ɪ/ and /e/ vowel categories stay spectrally separate, even though minimal pair /ɪ/-/e/ keywords did not occur in the closed word set in the current task. The increase of /ɪ/ spectral values in HI-directed speech presumably led to the observed smaller between-category distances and greater category overlap in the HI-directed condition. NH speakers also exhibited a smaller amount of category discriminability in the HI-directed condition than in the NH-directed condition. However, this effect was not found for HI speakers. For both groups, within-category dispersion for these vowels was greater when talking with a HI interlocutor, implying that speakers made more dynamic changes to their vowel categories, adjusting the clarity of their vowels according to their listener's needs.

#### 7.3.2.5 Tense/Lax distinction: durational

**Overall** For tense/lax vowel duration, the final model included speaker hearing status, listener hearing status and vowel as fixed factors, and age and keyword as random factors. The three-way interaction was significant ($\chi2(11) = 4.31$, $p<0.05$). This interaction was further examined by conducting separate linear mixed effects model analyses on the /i/ and /ɪ/ vowels, using the fixed factors of speaker hearing status and listener hearing status, and the same random factors as in the main model.

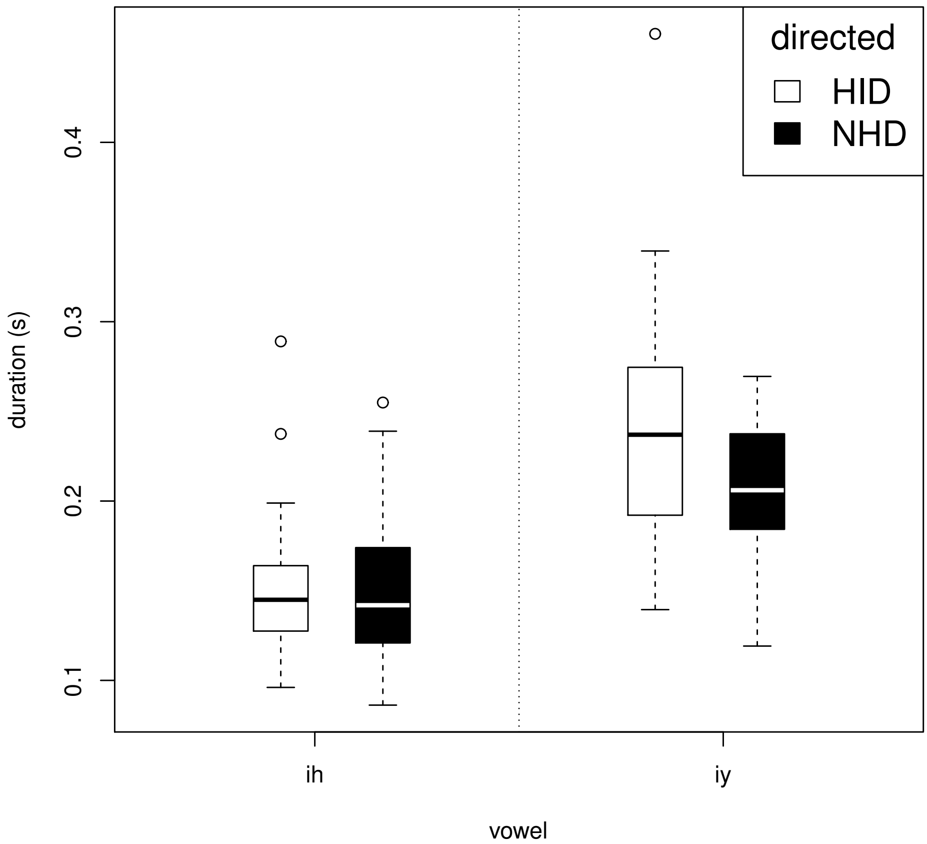



Figure 7.11: Mean durations for the /i/ (iy) and /ɪ/ (ih) vowels in the NH- and HI-directed conditions.

However, the exploration of the three-way interaction showed that there were no significant interactions between speaker hearing status and listener hearing status for either /i/ (p=0.063) nor /ɪ/ (p=0.23). Therefore the next-highest significant interaction from the main model, listener hearing status and vowel ($\chi2(8)$ = 10.74, p=0.001), was explored (see Figure 7.11). As Shapiro-Wilk's normality test showed that the data for both HI and NH participants was non-normally distributed, Wilcoxon's signed rank tests and Wilcoxon rank sum tests were used as post-hoc tests. The /i/ vowel was found to be longer in duration in the HI-directed condition (mean: 235.2ms) than in the NH-directed condition

(203.4ms) (V=18970.5; $p<0.0001$, r=-0.10), albeit with only a small effect size. However, the /ɪ/ vowel was not lengthened when talking with a HI friend (mean: 149.7ms) compared to talking with a NH friend (mean: 149.5ms) (V=22778; p=0.40, n.s.). The two vowels significantly differed from each other in duration in both the NH-directed (mean /i/: 203.4ms; mean /ɪ/: 149.5ms) (V=7598.5, $p<0.0001$, r=-0.24) and the HI-directed conditions (mean /i/: 235.2ms; mean /ɪ/: 149.7ms) (V=8589.5, $p<0.0001$, r=-0.32). The effect of speaker hearing status was only near-significant (p=0.063).

In summary, in terms of tense-lax vowel duration, no differences between NH and HI participants were found. Both groups of participants were found to distinguish between the two vowels in terms of length, and both groups slightly increased the duration of the /i/ vowel when talking with their HI friend compared to when talking with their NH friend. However, the duration of the /ɪ/ vowel was not altered depending on the listener's hearing status, implying that the durational contrast was increased in HI-directed speech compared to NH-directed speech. The analyses below will focus on exploring whether the durational distinction between the two vowels was indeed enhanced in HI-directed speech.

**Between-category distance** The final model for the measure of /i/-/ɪ/ durational between-category distance included speaker hearing status and listener hearing status as fixed factors, and age as a random factor. The interaction between speaker hearing status and listener hearing status was not significant (p=0.088), and neither was the main effect of speaker hearing status (p=0.69). The main effect of listener hearing status was significant ($\chi 2(4) = 10.48$, p=0.001, r=-0.39), with greater durational between-category distances in the HI-directed condition (mean: 86ms) than in the NH-directed condition (mean: 58ms).

**Within-category dispersion** For the within-category dispersion measure, the final model included speaker hearing status and listener hearing status as fixed factors and age as a random factor. The interaction was not significant (p=0.15). The main effect of speaker hearing status was significant ($\chi 2(5) = 5.49$, $p<0.05$, r=0.34) - with greater durational within-category dispersion found in HI speakers (mean: 78ms) than in NH speakers (mean: 58ms). There was also a significant effect of listener hearing status ($\chi 2(4) = 6.66$, $p<0.01$, $r=0.42$) - greater durational within-category dispersion was found in HI-directed (mean: 76ms) than in NH-directed speech (mean: 60ms). As above in section 7.3.2.4, this finding is likely to be due to speakers making dynamic adjustments to their speech according to the particular needs of the communication partner.

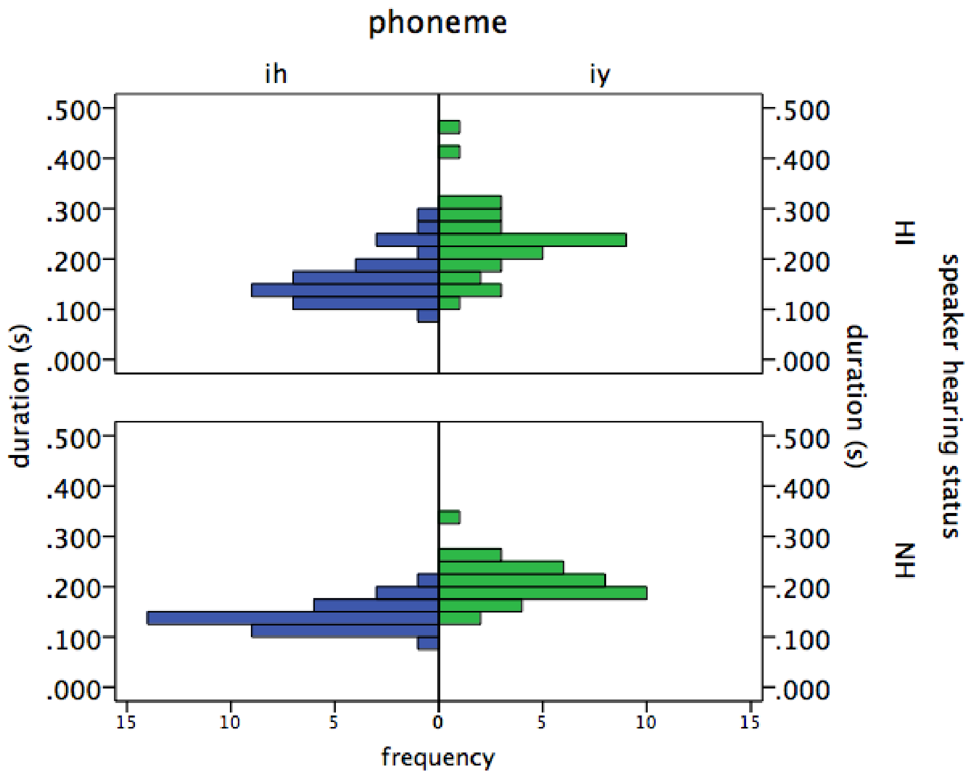



Figure 7.12: Mean durations for the /i/ (iy) and /ɪ/ (ih) vowels per speaker group.

**Overlap** The final model for durational overlap between categories included the fixed factors of speaker hearing status and listener hearing status, and age as a random factor. The interaction between speaker hearing status and listener hearing status was not significant ($p=0.96$). The effect of speaker hearing status was significant ($\chi 2(5) = 6.45$, $p<0.05$, $r=0.45$) - HI speakers had greater durational overlap between /i/ and /ɪ/ (mean: -151ms) than did NH speakers (-94ms). There was no effect of listener hearing status ($p=0.65$), suggesting that neither group of speakers decrease their category overlap when speaking with a HI interlocutor.

**Discriminability** For durational category discriminability, the fixed factors of speaker hearing status and listener hearing status, and the random factor of speaker were included in the final model. The interaction was not significant, and neither were the main effects (all $p>0.19$) (see Figure 7.12). This finding suggests that the discriminability of durational categories did not significantly differ between NH and HI speakers, and neither group made the /i/-/ɪ/ distinction durationally more discriminable in HI-directed speech.

**Summary** In terms of the durational distinction of the /i/-/ɪ/ vowels, HI speakers were found to have similar vowel durations, between-category distances and category discriminability as NH speakers. However, HI speakers' /i/ and /ɪ/ vowel durations had more overlap than did NH speakers'. As expected, HI speakers also exhibited greater within-category dispersion than NH speakers. When talking with a HI interlocutor, both groups of speakers were found to slightly increase the duration of the /i/ vowel, and to make between-category distances durationally greater than in NH-directed talk. There was also increased within-category dispersion, implying that dynamic changes were being made to the vowels according to feedback from the interlocutor. Durational discriminability between vowels was not, however, increased in HI-directed speech.

#### 7.3.2.6 Spectral and durational discriminability of tense/lax vowels

Wilcoxon signed rank tests were performed on the spectral and durational discriminability data to examine whether speakers used spectral or durational cues as the primary feature in discriminating their /i/-/ɪ/ vowels, and whether this was consistent across speaking conditions. Table 7.3 shows that NH speakers' tense/lax vowels were more discriminable spectrally than durationally in both NH- and HI-directed conditions. For HI speakers, the tense/lax vowels were also more discriminable spectrally than durationally in the NH-directed condition. However, in the HI-directed condition, the two vowels were equally discriminable in terms of duration and the spectrum. This suggests that, for NH speakers, the primary cue for distinguishing /i/ and /ɪ/ is spectral, regardless of speaking condition. HI speakers also tend to use spectral cues in distinguishing the vowels in NH-directed speech, but use both cues equally in HI-directed speech.

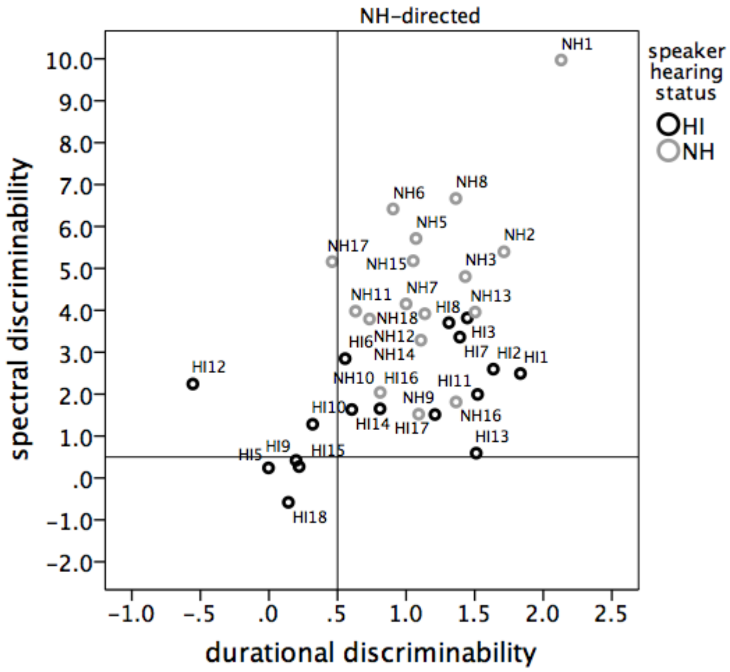
| data | V | p | r | means: spectral vs. durational |
|---|---|---|---|---|
| HI-NHD | 137 | <0.01* | -0.37 | 1.77 > 0.83 |
| HI-HID | 111 | 0.11; n.s. | - | 1.29 = 1.07 |
| NH-NHD | 153 | <0.0001*** | -0.52 | 4.58 > 1.15 |
| NH-HID | 153 | <0.0001*** | -0.52 | 3.08 > 1.14 |

Table 7.3: Wilcoxon signed rank tests performed on spectral and durational discriminability measures for /i/-/ɪ for the NH- and HI-directed conditions (n=68).

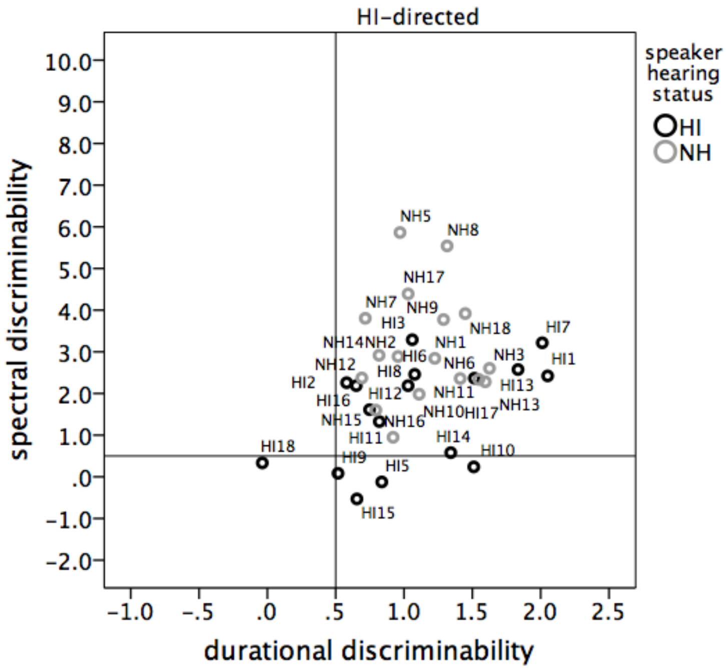
However, there is some individual variability in terms of cue use. Figure 7.13 displays the mean spectral discriminability and mean durational discriminability for the /i/-/ɪ/ distinction per speaker in each condition. The horizontal line distinguishes between speakers whose spectral discriminability is higher or lower than 0.5, and the vertical line distinguishes between speakers whose durational discriminability is higher or lower than 0.5[1]. Therefore the lower left quadrant defined by the lines displays the speakers who, both spectrally and durationally, have very low discriminability of /i/-/ɪ/. Those in the upper left quadrant have low durational discriminability of vowels, but higher spectral discriminability. Speakers in the upper right quadrant have higher spectral and durational discriminability, and those in the lower right quadrant have higher durational discriminability but very low spectral discriminability of vowels.

As the figure shows, in the NH-directed condition, there are four HI participants (HI18, HI5, HI9 and HI15) in the lower left quadrant who do not seem to reliably distinguish between the two vowels either in durational or spectral terms (both spectral and durational discriminability are close to 0). In the HI-directed condition, only HI18 still does not distinguish between the vowels using either cue, and the remaining three HI participants, now in the lower right quadrant, attempt to use duration to discriminate between the vowels. A further HI participant, HI12, only uses spectral information to distinguish between his vowels

[1] Discriminability below 0.5 denotes a very large amount of overlap between the distributions of /i/ and /ɪ/.




(a) NH-directed condition.




(b) HI-directed condition.

Figure 7.13: Scatterplot per speaker of mean /i/-/ɪ/ spectral and durational discriminability for the NH-directed condition (a) and the HI-directed condition (b).

in the NH-directed condition (upper left quadrant), but in the HI-directed condition, additionally uses durational information to discriminate between vowels (upper right quadrant). An additional two participants (HI14 and HI10) are located in the lower right quadrant in the HI-directed condition, and seem to use mainly duration to discriminate between their tense/lax vowels, at least in this condition. Interestingly, of the six HI participants mentioned here who mostly use durational cues to distinguish their /i/-/ɪ/ vowels, five are CI users with profound hearing loss (HI18, HI5, HI9, HI14 and HI10). HI3 and HI8, the remaining two CI users, use both spectral and durational discriminability to a greater extent.

### 7.3.3 Correlations between measures

Table 7.4 shows the correlations between the segmental discriminability measures for the /p/-/b/, /s/-/ʃ/ and /i/-/ɪ/ contrasts for the across speaker groups in both conditions. As some of the data was found to be non-normally distributed, Spearman's rho was used. The correlations between all measures were either non-significant or only weakly correlated, with the exception of fricative spectral discriminability and vowel spectral discriminability, which were moderately positively correlated. This finding implies that speakers' phonetic category discriminability is not necessarily a speaker-specific feature over all phonetic categories, but differs depending on the particular phonetic contrasts examined (see Romeo *et al.*, 2013, for similar findings with regards to VOT and fricative contrasts in adults and children).

| data | rho | p |
|---|---|---|
| fricatives-VOT | 0.05 | n.s. |
| fricatives-tense/lax duration | 0.15 | n.s. |
| fricatives-tense/lax spectral | 0.47 | <0.0001*** |
| VOT-tense/lax duration | 0.11 | n.s. |
| VOT-tense/lax spectral | 0.33 | <0.01** |
| tense/lax duration-tense/lax spectral | 0.35 | <0.01** |

Table 7.4: Correlations between the segmental discriminability measures from the Grid task, using Spearman's rho (n=68).

### 7.3.4 Summary

In summary, findings suggest that, regardless of interlocutor hearing status, HI speakers are able to make adequate phonetic contrasts in spontaneous speech when using durational cues (such as VOT and tense/lax durational distinctions). Although their vowel spaces are generally similar to NH speakers', they have problems with distinguishing between some more fine-grained spectral contrasts, i.e. in terms of vowel front-backness, subtle formant differences in tense/lax vowels, and fricative spectral distinctions. For these contrasts, HI speakers were also found to have greater within-category dispersion, indicating less robust phonetic categories.

Importantly, however, the results from the current chapter suggest that HI speakers use similar strategies as NH speakers in enhancing their segmental contrasts in HI-directed speech. Both groups of speakers made changes to their vowels in response to a HI interlocutor by slightly increasing their vowel space area and F1 range, and by making their /i/ vowels longer and durationally more distinct from their /ɪ/ vowels. Surprisingly, in HI-directed speech, both groups of speakers were found to increase their /ɪ/ vowel values spectrally, which made their /ɪ/ categories more similar to their /i/ categories. This may have been due to speakers attempting to make their /ɪ/ vowels more distinct from other neighbouring vowels. For the secondary cue between the tense-lax vowels, duration, both NH and HI speakers increased between-category distances by increasing the duration of the /i/ vowel, but not the /ɪ/ vowel, in HI-directed speech compared to NH-directed speech. This could have been due to articulation rate effects, which were not analysed here. However, due to greater within-category dispersion in the HI-directed condition than in the NH-directed condition, the discriminability of the durational /i/-/ɪ/ contrast did not change. Neither NH or HI speakers were found to enhance consonantal phonetic contrasts (stop voicing or spectral fricative contrasts) in HI-directed speech – although speakers did increase the fricative centre of gravity, and a subset of speakers prevoiced their /b/, when speaking to a HI peer.

In both spectral and durational aspects, speakers' tense/lax vowels were found to have greater within-category dispersion in HI-directed speech than in NH-directed speech. This is presumably due to speakers dynamically changing their vowel formants and durations according to the ease of communication at each specific time point in the conversation.

## 7.4 Discussion

### 7.4.1 Normally-hearing children

**Segmental adaptations** Most studies examining children's category discriminability are based on speech from picture-naming or read sentences (e.g., Romeo *et al.*, 2013), but this study used a spontaneous speech task to elicit multiple repetitions of minimal pair phonetic contrasts. Our findings suggest that, despite the increased difficulty in completing the task with a HI friend (see section 5.3.2), NH children did not enhance most phonetic category distinctions in HI-directed speech. It is striking that these contrasts were not enhanced even in the Grid task, in which accurate phoneme perception is correlated with task success (c.f., 5.3.2), and therefore a speaker's increased category discriminability would have been very useful for successful task completion. Previous studies on children's enhancement of fine phonetic detail have mostly exclusively examined modifications made to vowel space (e.g., Pettinato *et al.*, 2016; Syrett and Kawahara, 2014), or have used 'clear speech' instructions for elicitation in preschool-aged children (Redford and Gildersleeve-Neumann, 2009). This study is therefore novel in exploring phonetic category enhancement in older children in an interactive spontaneous speech task.

It is possible that most of the segmental contrasts in this study were not enhanced by the NH speakers due to the HI listeners being able to perceive the contrasts correctly even without category enhancement. For the stop voicing contrast, this may indeed have been the case – in section 4.3.2 the HI participants were found to perceive, on average, 84% of all /p/ and /b/ tokens accurately in the AV perception test, and section 5.3.1.2 demonstrated that of the errors made in the final finished grids in the Grid task very few were /p/-/b/ voicing errors. Additionally, the results in the current chapter showed that HI participants' production of the voicing contrast did not significantly differ from that of

NH participants, implying that HI participants' /p/-/b/ phonological categories had been adequately acquired despite their hearing loss. The prevoicing in /b/ produced by a subset of participants in the HI-directed condition but not the NH-directed condition implies that, nonetheless, a few speakers attempted to enhance this contrast – perhaps to the few HI interlocutors with greater difficulty in perceiving this contrast. However, it seems likely that most HI listeners were impaired in perceiving the /s/-/ʃ/ and /i/-/ɪ/ contrasts, as more than 50% of all errors in the final completed grids in the Grid task were either /s/-/ʃ/ or /i/-/ɪ/ keyword confusions, and both of these distinctions were found to be produced less contrastively by the HI speakers compared to NH speakers – implying that HI participants had not acquired these contrasts adequately. The finding of little contrast enhancement in the Grid task is, therefore, unlikely to be due to HI listeners having no difficulty with the phonetic distinctions.

Although, as in Bradlow (2002) and Ferguson and Kewley-Port (2002, 2007), a tendency towards vowel space enhancement was found in this study, spectral tense/lax distinctions for /i/-/ɪ/ were *reduced* – due to the vowel /ɪ/ becoming spectrally closer to /i/ in HI-directed speech compared to NH-directed speech. These results are similar to those found in Granlund *et al.* (2012) for late bilingual speakers producing clear speech and Cristia and Seidl (2014) for adult IDS, as well as Picheny *et al.*'s (1986) findings of a greater expansion of the lax than the tense vowel space in clear speech compared to casual speech. This decrease in discriminability between contrasting vowel pairs suggests that spectral changes in vowels in clear speaking styles may in fact be a side-effect of other modifications made by the speaker, such as an increased vocal intensity (Ferguson and Quené, 2014) or a general increase in vocal effort (Cristia and Seidl, 2014) when adapting to a listener's needs. Indeed, inspection of vowel space area expansion in Figure 7.1 shows that the HI-directed vowel space differed from the NH-directed vowel space mainly in the decreased F1 values of the vowel /æ/ in HI-directed speech, which may have occurred due to speakers opening their jaws to a greater extent in that condition to enable them to speak more loudly (Ferguson and Quené, 2014). Similarly, the slightly increased centre of gravity values for both /s/ and /ʃ/ in the HI-directed condition compared to the NH-directed condition in this study may merely be a side-effect of greater articulatory effort being made (i.e., an increased airflow velocity at the fricative constriction) (Silbert and de Jong, 2008).

The only segmental contrast for which contrast enhancement in HI-directed speech was observed in the current study was the secondary cue to the /i/-/ɪ/ distinction, duration. Although the discriminability of the durational contrast was not increased, the enhanced between-category distance in the HI-directed condition compared to the NH-directed condition implies that speakers were attempting to enhance the /i/-/ɪ/ distinction using the durational cue. It is possible that durational enhancement was used instead of spectral enhancement due to this being a more useful strategy when speaking to HI interlocutors with impaired frequency selectivity but fairly intact temporal processing (see section 1.2.1). Alternatively, despite its secondary cue status, duration may be an easier cue to enhance in speech, as demonstrated by similar findings of durational enhancement of the cue in clear speech by adults (Granlund *et al.*, 2012; Uchanski, 1988).

The participants in the current study were children, and therefore the lack of contrast enhancement could be attributed to their still-developing speech production system (e.g., Lee *et al.*, 1999). The greater variability in children's speech compared to adults' (Romeo *et al.*, 2013) may lead to them having less control over the fine phonetic detail required to produce enhanced phonetic contrasts. Also, as children already produce greater distances between phonetic categories compared to adults (Romeo *et al.*, 2013), children may be unable to enhance their categories any further – for example, in Pettinato *et al.* (2016), it was found that children's vowel space area decreased significantly with age between 9 and 14 years, and only the older children, whose vowel space areas were more reduced than younger children's, were able to enhance the size of their vowel space area when talking to a friend through a vocoder compared to the no-barrier condition. Therefore phonetic reduction in speech may facilitate greater phonetic enhancements being made in adverse speaking conditions. Alternatively, even older children may not have accurate knowledge of the cues that would be useful in making phonetic contrasts more discriminable, as evidenced by children's less consistent categorisation of phonemic contrasts even at age 12 years compared to adults (Hazan and Barrett, 2000; Hoonhorst *et al.*, 2011).

The Dual Process model (Bard *et al.*, 2000; Bard and Aylett, 2005) predicts the results of the current study by stating that enhancing the segmental level of speech would be too cognitively demanding for speakers, as they would need to continuously consult a listener model while planning the production of each phoneme (Bard and Aylett, 2005). Studies which apply direct miscomprehension feedback to speakers (e.g., Schertz, 2013) have shown that speakers can enhance the segmental level of their speech if given specific minimally contrastive feedback, and if word repetition is the only strategy that speakers are allowed to use. However, instances in which miscommunications occur at only a minimally contrastive phoneme, and in which only word repetition can be used as a strategy are likely to be rare outside of clear speech studies. In spontaneous speech studies such as the current study, it seems likely that other strategies, such as linguistic enhancement strategies or gestures, will be used instead of segmental adaptations. Indeed, when encountered with a HI interlocutor with specific impairments in phoneme perception, it is possible that modifications to the linguistic level may be more likely to be successful in interaction than segmental enhancements. These linguistic adaptations are examined further in the next chapter, chapter 8.

The NH speakers in the current study produced greater within-category dispersion for the /i/-/ɪ/ contrast both spectrally and durationally in the HI-directed condition compared to the NH-directed condition, contrary to the hypothesis that in adverse speaking conditions, speakers produce more internally consistent categories, which may increase the intelligibility of a speaker (Newman *et al.*, 2001). This finding implies that speakers modified their speech dynamically in the HI-directed condition in an attempt to enhance their speech in the interaction to a greater extent when encountered with miscommunications compared to when no miscommunications occurred, as predicted by the H&H model (Lindblom, 1990) – an aspect of segmental adaptation that the instructed clear speech literature (e.g., Ferguson, 2004; Picheny *et al.*, 1986; Smiljanić and Bradlow, 2005) is unlikely to be able to emulate.

### 7.4.2 Hearing-impaired children

**Segmental speech characteristics** The HI speakers in this study exhibited very few differences to NH peers in their production of the temporal contrasts, the VOT /p/-/b/ contrast and the secondary, durational cue to the /i/-/ɪ/ distinction. This is likely to be due to HI children's temporal processing being fairly unaffected by their hearing loss (see section 1.2.1). Nonetheless, it is surprising that HI children are able to produce these contrasts adequately, especially in spontaneous speech – previously, for the /p/-/b/ VOT distinction, HI children wearing HAs have been found to exhibit problems in perceiving voicing contrasts (e.g., Borg *et al.*, 2007; Tsui and Ciocca, 2000), likely due to the unavailability of visual cues to voicing (Kishon-Rabin *et al.*, 2002). However, Elfenbein *et al.* (1994) discovered that for children with less than 75 dB hearing loss, the pronunciation of consonants other than fricatives was not greatly affected by their hearing loss. Additionally, Uchanski and Geers (2003) found that 70-80% of CI children's production of voicing in the /t/-/d/ contrast was within NH range – thus it seems that modern hearing aids and early intervention have allowed HI children with even severe losses to produce accurate voicing categories.

Unlike in many previous studies (e.g., Horga and Liker, 2006; Löfqvist *et al.*, 2010; Neumeyer *et al.*, 2010), HI children's vowel space area was generally very similar in size to NH peers'. Nonetheless, HI children exhibited a slightly reduced F2 range compared to NH controls. This result is similar to that of previous studies (e.g., Chuang *et al.*, 2012; Löfqvist *et al.*, 2010; Neumeyer *et al.*, 2010), and may be due to the higher frequency of F2 formants. Despite similar vowel space areas between HI and NH speakers, HI children produced significantly less spectrally discriminable /i/-/ɪ/ vowel contrasts than NH speakers – thus demonstrating that the size of the vowel space area may nonetheless not reflect the accuracy of the HI child's entire vowel system. Similarly, the HI children's production of the spectral /s/-/ʃ/ distinction differed both in between-category distance and in discriminability compared to NH peers, mostly due to HI children's centre of gravity for /s/ being closer in frequency to /ʃ/, as also found by previous studies (Uchanski and Geers, 2003). Thus, as expected from previous studies showing that even mild hearing loss leads to impaired fricative production (Elfenbein *et al.*, 1994), the production of fine spectral detail seems to present difficulty to HI children. This is likely to be due to HI children's impaired frequency selectivity, which prevents both HA and CI users from distinguishing between fine spectral contrasts. This result may also be partly due to some of the HI children wearing frequency compression HAs – which may lead the /s/ and /ʃ/ phonemes to be perceived, and thus also produced, spectrally closer together[1].

Using the Grid task to elicit phonetic contrasts in spontaneous speech enabled a detailed analysis of HI children's within-category variability to be made – as reported in previous studies based on read or imitated speech (e.g., Uchanski and Geers, 2003; Vandam *et al.*, 2011), HI children produced greater variability than NH peers in both the spectral and temporal domain for /i/-/ɪ/, and in the spectral domain for the /s/-/ʃ/ category. This increased variability indicates that the HI children may have immature speech motor control for producing these fine phonetic distinctions. The numerous elicited tokens of each phoneme also allowed an analysis of category overlap to be made – Haley *et al.* (2010) and Kim *et al.* (2011) have previously shown that speakers producing overlapping phonetic categories are significantly less intelligible than those without category overlap – and may be an indication of disordered speech (Haley *et al.*, 2010). Indeed, HI children exhibited greater overlap in spectral and temporal /i/-/ɪ/, as well as greater spectral overlap in /s/-/ʃ/, than NH peers, thus indicating that at least some HI children produce less precise and less separable phonetic categories compared to NH peers.

**Segmental adaptations** Like NH peers, HI children were not found to enhance segmental contrasts when talking to a HI peer, with the exception of an increase in the temporal between-category distance between /i/ and /ɪ/ in HI-directed speech compared to NH-directed speech. Interestingly, as demonstrated by Table 7.3, this durational enhancement by HI speakers in HI-directed speech lead to their /i/-/ɪ/ vowels being equally discriminable in both spectral and temporal domains in that condition – unlike NH speakers' tense/lax vowels, which retained their primary spectral cue. Additionally, most of a subset of HI speakers who were

[1]However, as we did not collect information from participants on whether frequency compression was activated in their HAs, we are unable to analyse this further.

unable to distinguish between the tense/lax vowels in the NH-directed condition were able to use the durational cues to enhance the contrast in HI-directed speech (see Figure 7.13). These findings lend further support for the hypothesis that duration is a somewhat more important cue to tense/lax vowels for HI children than for their NH peers (c.f., Monsen, 1974), and that it may be an easier cue to enhance than spectral distinctions, as discussed in the previous section.

Although it could be argued that NH children do not need to enhance segmental contrasts in HI-directed speech due their phonetic categories already being sufficiently discriminable, this is unlikely to be the case for HI speakers – at least for the /i/-/ɪ/ and /s/-/ʃ/ contrasts for which HI speakers' productions were substantially less discriminable than their NH peers'. It is possible that HI children's less precise phonetic categories and reduced speech motor control prevent them from enhancing these distinctions as necessary. Alternatively, HI children are likely to be unable to monitor the intelligibility of their own production of especially the fine spectral contrasts, which may prevent them from modifying this aspect of their speech when needed. However, as neither NH nor HI participants were found to enhance segmental contrasts when talking with a HI friend, it seems possible that both groups of participants were either unable to do so due to their continued speech production development, or due to using other, such as linguistic, strategies to clarify their speech to a HI peer – this aspect will be further explored in the next chapter, chapter 8. Additionally, chapter 10 examines the individual strategies in segmental adaptation used by speakers, and chapter 9 investigates which segmental characteristics of speech in NH and HI speakers are related to their speech intelligibility.

# Chapter 8

# Linguistic and conversational adaptations

## 8.1 Introduction

In this chapter, we investigate whether NH and HI speakers adapt to their HI interlocutor by altering the linguistic complexity, lexical content or turn-taking strategies used in their Diapix and Grid interactions with their HI friend compared to their NH friend, by measuring the mean length of phrase, the lexical frequency of content words, the diversity of language and the amount of overlap used by speakers.

The mean number of words used per phrase, roughly corresponding to the mean length of utterance (MLU), has been previously shown to correlate with the acquisition of grammatical morphemes, and has been used as a measure of produced syntactic complexity in both NH and HI children (see Flipsen and Kangas, 2014, for a review). The lexical frequency of words is here taken as a measure of the complexity of the vocabulary used by the speakers in each condition, as frequently occurring lexical items have been found to be inversely correlated with word difficulty, and tend to be acquired earlier (e.g., Breland, 1996; Tamayo, 1987). Lexical diversity is used to investigate the range of vocabulary produced by the speakers in this study, and has been previously used to assess the development and use of different vocabulary items in children and L2 learners (Lu, 2012; Richards, 1987) – a speaker needs to use several different words and repeat the same words rarely to obtain a high lexical diversity score (Johansson, 2008). The extent to which participants produce speech overlap is also examined – it is related both to the fluency of the conversation and to the amount of speech interruption by speakers (Boyle, Anderson, and Newlands, 1994; Heldner and Edlund, 2010).

There is a vast literature demonstrating that many HI children exhibit delays in receptive vocabulary and syntax (e.g., Blamey *et al.*, 2001; Spencer, 2004; Uziel *et al.*, 2007), although some studies suggest that even 50% of HI children with moderate to severe hearing loss (Yoshinaga-Itano *et al.*, 2010) or with CIs (Geers and Sedey, 2011) display age-appropriate receptive vocabulary. In the current study, 13 out of 18 HI participants were assessed by their teachers as having at least somewhat of a delay in their language and communication skills compared to their NH peers. Therefore we would expect that the production of less complex syntax (i.e., fewer words per phrase) and less complex vocabulary (i.e., more lexically frequent words) by the HI participants' interlocutors in the Diapix and Grid tasks would be beneficial to the HI listeners. Similarly, as discussed in section 1.3.1, when miscommunications occur, HI children are likely to benefit more from rephrasing rather than repeating an utterance – thus HI children's interlocutors may use more diverse language when talking with a HI peer. Additionally, as HI participants may be less able to monitor the production of their own and their interlocutor's speech simultaneously, it would be expected that speakers would try to minimise the overlap between interlocutors in HI-directed speech.

However, it is unclear whether NH or HI speakers are able to alter the linguistic or conversational aspects of their speech to an interlocutor's needs, as it may require sophisticated language skills as well as sufficient linguistic flexibility. NH children's vocabulary and syntactic knowledge is not yet adult-like in late childhood and adolescence (e.g., Duncan *et al.*, 2014; Nippold *et al.*, 2005). Children's pragmatic skills in turn-taking and conversational fluency are also found to greatly increase from approximately 7 to 17 years of age (Dorval, Eckerman, and Ervin-Tripp, 1984). HI children, on the other hand, are delayed in their expressive language skills (e.g., Wake *et al.*, 2004, 2005), produce less complex syntax and vocabulary than NH peers (Flipsen and Kangas, 2014; Koehlinger *et al.*, 2013), and may be delayed in their general pragmatic skills compared to peers (e.g., Goberis *et al.*, 2012) – and may therefore not be able to make adaptations on these linguistic and conversational levels.

## 8.2 File Processing

### 8.2.1 Number of words per phrase

The mean number of words per phrase was calculated 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[1], and words which were spoken in overlap with the interlocutor were excluded from analysis. Analyses were done on the mean number of words per phrase for each file in each condition.

### 8.2.2 Lexical frequency of content words

A Python script was used to calculate the mean lexical frequency of words per speaker per file. Partially spoken words, words spoken in overlap with the interlocutor, while laughing or in noise[2], or which were unintelligible to the transcriber were excluded. Lexical frequency was determined using word frequency per one million words from a list of word frequencies derived from American subtitles ($SUBTLEX_{US}$) (Brysbaert and New, 2009)[3]. Part-of-speech information was

[1] These were excluded as the aim was to include only utterances with meaningful content. The criteria are similar to those used in Shatz and Gelman (1973).

[2] These were excluded to ensure the words used for the linguistic measures were the same as those used for the acoustic measures.

[3] In Brysbaert and New's (2009) study, it was shown that using word frequency data based on subtitles more accurately accounted for listeners' variance in word recognition times than word frequency obtained from written texts, as used in many previous studies (e.g., Kučera and Francis, 1967). This is likely to be due to subtitles from film and TV programmes being more representative of everyday language use and interaction than edited written documents (see e.g., Brysbaert and New, 2009; Soares, Machado, Costa, Iriarte, Simões, de Almeida, Comesaña, and Perea, 2014, for a discussion of these issues). Recently, a version of SUBTLEX based on British English (van Heuven, Mandera, Keuleers, and Brysbaert, 2014) has also been published. The

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 $SUBTLEX_{US}$ word list were excluded from the frequency calculation (mean number of word exclusions per file: 1). In the HID condition, lexical frequency was calculated over a mean of 79 different content words per file, while in the NHD condition it was calculated over a mean of 67 content words per file. On average, there were 80 different content words in the Diapix task, and 69 different content words in the Grid task.

### 8.2.3 Lexical diversity (VOCD)

Lexical diversity was calculated using the Lingua-Diversity package (Xanthos, 2011) on Perl. Words spoken in overlap with the interlocutor, unintelligible words, partially spoken words, words spoken while laughing or in noise, agreement, hesitation and exclamation words (such as 'okay', 'umm' and 'yay') were excluded from analysis. The package calculates lexical diversity using an implementation of VOCD from McKee, Malvern, and Richards (2000). Traditional methods for measuring lexical diversity typically rely on the type-to-token ratio (TTR), which divides the number of different words (types) by the number of total words (tokens) in a file. However, the TTR has been shown to be 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[1], and then finding the curve that best fits the generated TTR x sample size curve. The parameter value D gives the best fit to the measured curve and reflects lexical diversity over the entire file. The method is repeated three times and the average

current analyses were conducted prior to the publication of $SUBTLEX_{UK}$, which was therefore not used in the current study. Notably, $SUBTLEX_{UK}$ includes word frequencies calculated on the basis of children's programmes, which could, in the future, be used as a likely more accurate frequency count for the speech of the children in the current study.

[1]Subset sample size is n=35,36,...,49,50 of words taken randomly from the entire file. The TTR for each sample size is measured 100 times, after which the mean TTR for that sample size is calculated.

D is calculated as the final VOCD measure for each file, with higher values indicating high lexical diversity, and lower values indicating low lexical diversity. The VOCD measure is reported to correlate with other child language measures and is not dependent on file length, making it more reliable than the TTR (McKee *et al.*, 2000).

### 8.2.4 Overlap between speakers

A Praat script counted the number of words transcribed as being spoken in overlap with the interlocutor. Agreement, hesitation and exclamation words were excluded from analysis, as overlap in these words is likely to reflect interactional processes such as back-channelling rather than overlap in meaningful utterances, which was the focus of the study. This number was then divided by the number of all words in the file[1] to obtain the % words spoken in overlap per participant per file.

## 8.3 Results

### 8.3.1 Statistical approach

The same statistical approach as detailed in section 6.3.1 was used in this chapter. As in chapter 6, speaker hearing status ('spHstatus', [HI, NH]), listener hearing status ('directed', [HI-directed, NH-directed]) and type of task ('task', [Grid, Diapix]) were included as fixed factors in the model, and speaker, task number and age as a continuous variable were treated as random factors. If an interaction between speaker hearing status and listener hearing status is found, it would imply that NH and HI speakers are using different linguistic enhancement strategies when speaking with their HI friend. Any significant task effects would suggest that the participants' speech was influenced by the type of task used.

[1]Excluding the agreement, hesitation and exclamation words, but including unintelligible words and words spoken in noise, while laughing, or in overlap with the interlocutor.

| measure | task | NH-NHD | NH-HID | HI-NHD | HI-HID |
|---|---|---|---|---|---|
| phrase length | Grid | 4.6 (1.1) | 4.3 (1.0) | 3.8 (0.8) | 3.6 (0.7) |
| | Diapix | 5.7 (1.0) | 5.4 (1.2) | 4.6 (1.1) | 4.7 (0.8) |
| lexfreq | Grid | 740.9 (205.9) | 784.8 (188.9) | 896.3 (228.5) | 893.2 (162.9) |
| | Diapix | 838.2 (130.2) | 849.5 (153.6) | 971.3 (181.8) | 974.8 (153.5) |
| VOCD | Grid | 34.1 (8.0) | 38.2 (10.3) | 39.4 (11.2) | 47.0 (12.7) |
| | Diapix | 50.0 (11.5) | 48.5 (9.0) | 48.2 (10.2) | 46.2(12.4) |
| overlap (%) | Grid | 6.6 (3.9) | 6.4 (3.8) | 6.5 (4.1) | 7.2 (3.9) |
| | Diapix | 5.4 (2.8) | 6.6 (3.8) | 7.3 (4.0) | 9.0 (3.7) |

Table 8.1: Means and standard deviations (in brackets) for each linguistic measure.

### 8.3.2 Mean number of words per phrase

The final model included speaker hearing status, listener hearing status and task as fixed factors, and age and task number as random factors. None of the interactions were significant. The main effect of speaker hearing status was significant ($\chi2(6) = 7.24$, $p<0.01$, $r=0.36$), with NH participants using more words per phrase (mean: 4.86) than HI participants (mean: 4.06). The main effect of listener hearing status was also significant ($\chi2(5) = 7.86$, $p<0.01$, $r=0.12$) - when talking with a HI interlocutor, speakers used slightly fewer words per phrase (mean: 4.38) than when talking with a NH interlocutor (mean: 4.57). There was also a main effect of task ($\chi2(7) = 169.32$, $p<0.0001$, $r=0.65$): more words per phrase were used in the Diapix task (mean: 5.10) than in the Grid task (mean: 4.14). These results suggest that although HI participants use shorter phrases than NH participants, both groups of speakers tend to reduce the length of their phrases, and therefore perhaps to use less complex phrases, when talking to a HI interlocutor. Table 8.1 displays the means and standard deviations for each participant group per condition and task.

### 8.3.3 Mean lexical frequency of content words

The fixed factors of speaker hearing status, listener hearing status and task, and the random factor of age were included in the final model. None of the interactions were significant. A main effect of speaker hearing status was found ($\chi2(5) =12.19$, $p<0.001$, $r=0.34$), with NH participants using less frequent content words (mean:

789.89) than HI participants (mean: 923.90). The main effect of listener hearing status was not significant - therefore speakers did not use more frequent words when speaking with a HI interlocutor. The main effect of task was significant ($\chi 2(6) = 15.86$, p=0.0001, r=0.38): content words in the Diapix task were more frequent (mean: 906.83) than those used in the Grid task (mean: 822.88) (see Table 8.1). This result is likely to be due to the keywords in the Grid task including some lexically less frequent words. Contrary to expectations, this result suggests that NH and HI participants do not choose more common vocabulary items when interacting with a HI interlocutor.

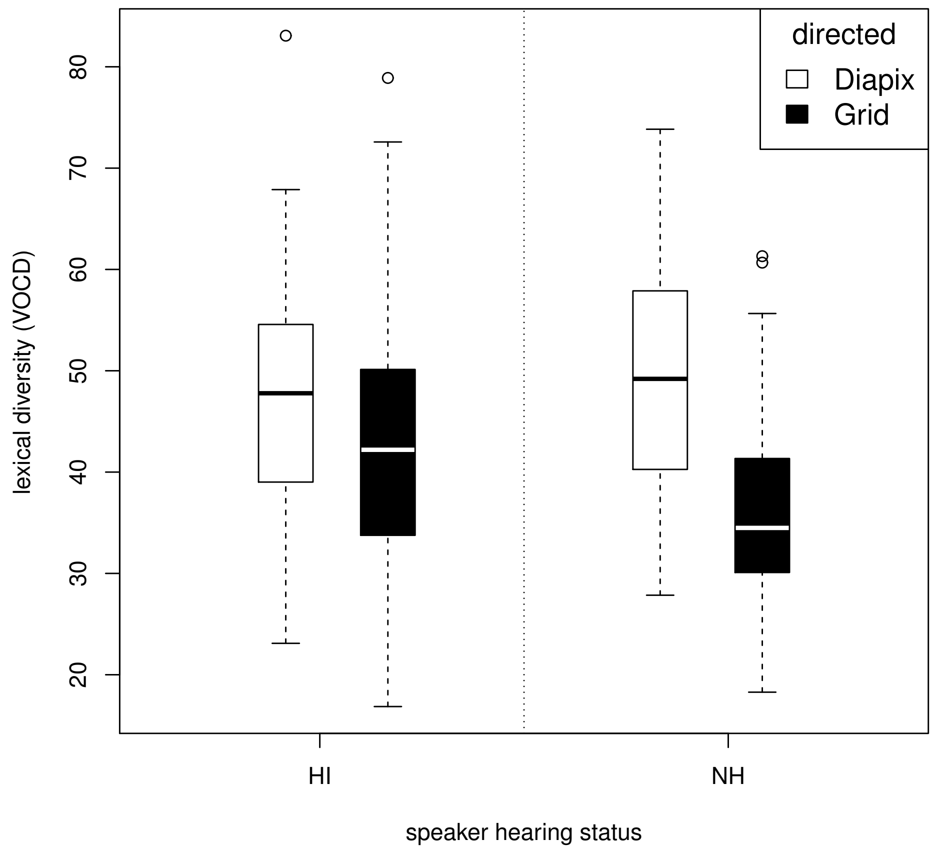



Figure 8.1: Lexical diversity for the NH and HI participants in the Grid and Diapix tasks.

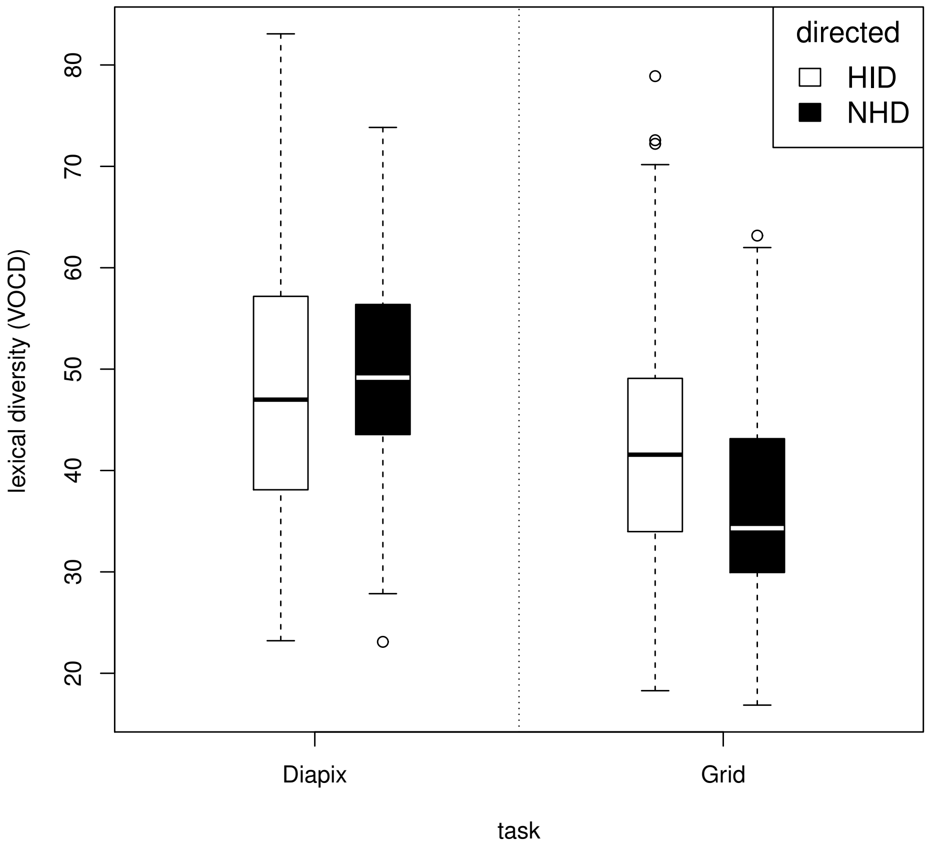



Figure 8.2: Lexical diversity in the NHD and HID conditions in the Grid and Diapix tasks.

### 8.3.4 Lexical diversity (VOCD)

The final model included speaker hearing status, listener hearing status and task as fixed factors, and age and task number as random factors. There was no significant three-way interaction. The interaction between speaker hearing status and task was significant ($\chi 2(9)$ =16.03, p=0.0001) (see figure 8.1). Post-hoc t-tests revealed that while HI and NH participants did not differ in their lexical diversity in the Diapix task (t(118.70)=1.03, n.s.) (mean HI: 47.24; mean NH: 49.24), the two groups did differ in the Grid task (t(182.77)=-4.66, p<0.0001, r=0.33), with higher lexical diversity used in the HI group (mean: 43.01) than in the NH group (NH: 35.97). The NH group also had a significantly higher lexical diversity in the Diapix task (mean: 49.24) than in the Grid task (mean: 39.97) ($t(49)=5.31$, $p<0.0001$, $r=0.60$), while the lexical diversity of the HI participants did not depend on the task (mean Diapix: 47.24; mean Grid: 43.01) ($t(35)= 0.33$, n.s.) (see Figure 8.1). In the Grid task, therefore, the HI participants used more diverse vocabulary than NH participants, but the two groups did not differ in their lexical diversity the Diapix task. This may be due to HI participants having to revise their utterances more often in the Grid task.

The interaction between listener hearing status and task was also significant ($\chi2(8) =10.14$, $p<0.01$) ( see figure 8.2). Post-hoc t-tests revealed that in the Grid task, participants had a higher lexical diversity score when talking to a HI interlocutor (mean: 42.38) than when talking to a NH interlocutor (mean: 36.51) ($t(95)=4.23$, $p<0.0001$, $r=0.40$), but in the Diapix task, there was no condition effect (mean HID: 47.39; mean NHD: 49.09) ($t(56)=-1.14$, n.s.). Additionally, in the HI-directed condition, there was no effect of task (mean Grid: 42.38; mean Diapix: 47.39) ($t(42)=-0.54$, n.s.), but in the NH-directed condition, lexical diversity was higher in the Diapix task (mean: 49.09) than in the Grid task (mean: 36.51) ($t(48)=-4.83$, $p<0.0001$, $r=0.57$) (see Figure 8.2 and Table 8.1). These results demonstrate that participants used lexically more diverse language when talking with a HI interlocutor, but only in the Grid task. This may point towards speakers having to revise their utterances more in the HI-directed condition, due to an increase in misunderstandings. The Diapix task probably elicits more diverse vocabulary and language than the Grid task, leading to higher lexical diversity scores in the NHD condition in the Diapix task than in the Grid task.

There was no interaction of speaker hearing status and listener hearing status, and therefore the two groups did not differ in their strategy of increasing lexical diversity in response to a HI interlocutor in the Grid task.

### 8.3.5 Proportion of words spoken in overlap

The final model included speaker hearing status, listener hearing status and task as fixed factors, and speaker and task number as random factors. The interaction between speaker hearing status and task was significant ($\chi2(9)$ = 4.78,

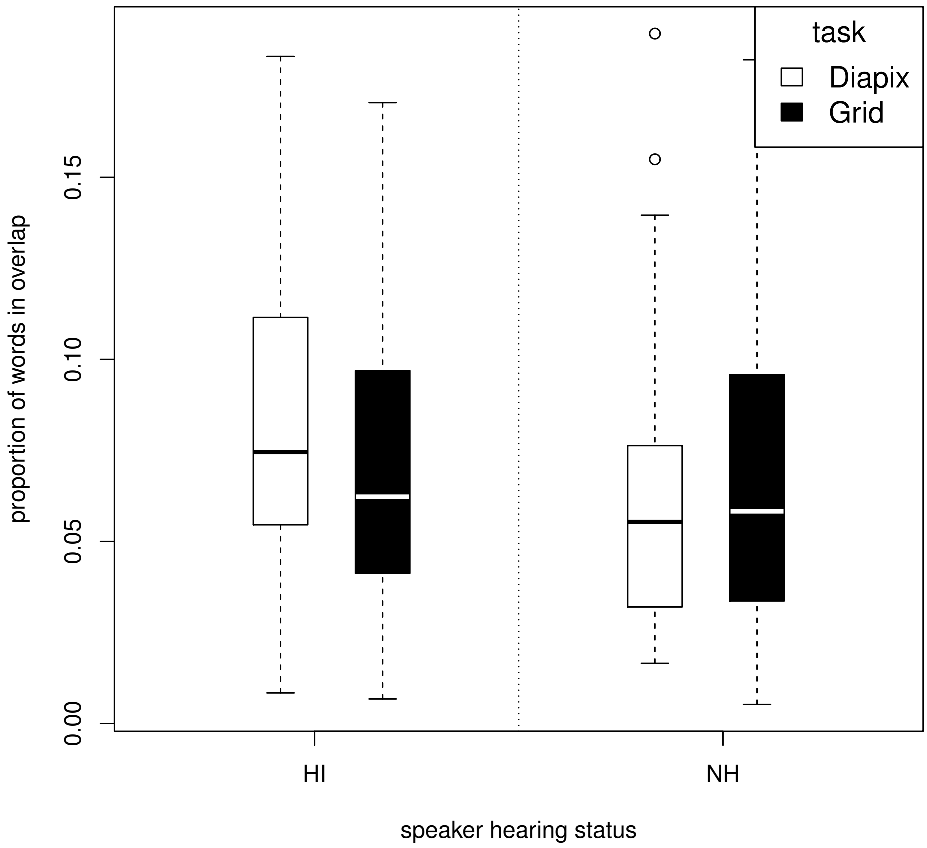



Figure 8.3: Proportion of words produced in overlap with the interlocutor in the two tasks.

$p<0.05$). The Shapiro-Wilk test showed that both NH and HI groups' data was non-normally distributed (HI: $W= 0.97$, $p=0.003$; NH: $W= 0.95$, $p<0.0001$), and therefore the Wilcoxon rank-sum test and the Wilcoxon signed-rank tests were used as post-hoc tests. The post-hoc tests demonstrated that HI participants had proportionately more words in overlap with their interlocutor in the Diapix (mean: 8.1%) than in the Grid task (mean: 6.8%) ($V=1326$, $p<0.01$, $r=-0.16$). However, the amount of NH participants' overlap with their interlocutor did not depend on the task ($V=1216$, n.s.). In the Grid task, the NH and HI participants did not differ in the amount of overlap they used with their interlocutor (W=

5652.5, n.s.), but in the Diapix task, HI participants spoke proportionately more in overlap with their interlocutor (mean: 8.1%) than did NH participants (mean: 6.0%) (W=1256, p=0.001, r=-0.17) (see Fig. 8.3). The main effect of listener hearing status only approached significance ($\chi2(5) = 3.31$, p=0.069).

This result suggests that the type of task used to elicit speech may affect the amount of overlap that the speakers use with their interlocutor - here, the Diapix task may be less constrained in terms of turn-taking than the Grid task, and therefore the HI participants, but not the NH participants, used more overlap in the Diapix task than in the Grid task. They also exhibited more overlap with their interlocutor in the Diapix task than did NH participants. It may be that HI participants have difficulty in turn-taking, and that this is especially evident in tasks in which turn-taking is not specifically assigned to each person, as in the Diapix task. By interrupting and overlapping with their interlocutor, HI participants may try to take more 'control' of the situation when they have difficulty understanding their interlocutor. The fact that NH participants did not do this suggests that they are aware that talking simultaneously with their interlocutor will likely hinder communication. However, the lack of an effect of listener hearing status implies that neither group changed the amount of overlap they use in response to a HI interlocutor.

### 8.3.6 Summary

Compared to the NH participants, HI speakers were found to use shorter phrases and lexically more frequent content words in both tasks, as well as more diverse vocabulary in the Grid task. HI participants were also found to speak more in overlap with their interlocutor than NH participants in the Diapix task. However, when talking with their HI friend compared to their NH friend, both groups used shorter phrases in both tasks, and increased the diversity of their vocabulary in the Grid task. Neither group used more lexically frequent content words, or less speech overlap, in response to a HI interlocutor.

The effect of the type of task was more evident in the linguistic measures than in the acoustic-phonetic ones. Probably due to the Grid task being an easier, more constrained and less complex task, speakers used fewer words per phrase and less diverse vocabulary in the NH-directed condition in the Grid task compared to the Diapix task. HI children also used a greater amount of speech overlap in the Diapix task than the Grid task – supporting the notion that HI children used speech overlap to control the interaction, especially in the more difficult and less constrained speaking conditions. In addition, higher frequency words were used in the Diapix than the Grid task – this is likely to be due to the keywords in the Grid task including some less frequent vocabulary items. In addition, we found that the type of task affected the enhancement strategies used by speakers: in the Grid task, speakers were found to use more diverse vocabulary with a HI than with a NH interlocutor, but such a strategy was not evident in the Diapix task. It seems likely that this is due to the constrained nature of the Grid task – it may enable us to measure even small differences in the lexical diversity used by the speakers in the task. In the Diapix task, speakers have greater flexibility to describe their pictures using many different strategies even in the NH-directed condition, and therefore the lexical diversity measure may not be as successful in detecting subtle differences in the task.

## 8.4 Discussion

### 8.4.1 Normally-hearing children

This chapter examined the linguistic modifications to speech made by speakers when talking to a HI peer compared to a NH peer, using measures of syntactic complexity, lexical diversity and complexity, and turn-taking. Previous literature on children's speaker-listener adaptations has frequently examined children's use of syntactic simplification and repetition in IDS (e.g., Dunn and Kendrick, 1982; Sachs and Devin, 1976; Shatz and Gelman, 1973). No known studies have examined whether children are able to modify the complexity of their vocabulary or the amount of speech overlap according to listener needs, and none have previously studied whether these linguistic and conversational modifications are made when speaking with a HI peer.

Although NH children's syntactic abilities are unlikely to be adult-like by late childhood (Nippold *et al.*, 2005), NH speakers were found to simplify their speech by slightly decreasing the number of words per phrase used in HI-directed speech compared to NH-directed speech. As in this study phrases were defined solely based on within-speaker pause length, rather than according to intonational phrasing as in, e.g., Shatz and Gelman (1973), the measure may be somewhat unreliable. Nonetheless, the strategy of reducing the number of words in a phrase in HI-directed speech is similar to that found in FDS by adult speakers (e.g., Long, 1983), and in IDS even by preschool-aged children (Dunn and Kendrick, 1982; Sachs and Devin, 1976; Shatz and Gelman, 1973), and therefore it is not very surprising that NH children are able to use this strategy as well.

NH children also increased the lexical diversity of their speech in the Grid task – therefore they used a greater number of different words when talking with a HI peer compared to a NH peer. Although VOCD is a fairly crude measure of the repetitiveness of the vocabulary produced by the participants in this study, and reflects the general diversity of vocabulary used by a speaker over the entire conversation rather than during points of miscommunication, it gives a general indication of whether the participants use repetition (lower lexical diversity) or revision (higher lexical diversity) behaviour during the interaction. Therefore our finding implies that speakers were using revision behaviour when describing the picture on their grid in the HI-directed condition than in the NH-directed condition. Indeed, as discussed in section 1.3.4, according to the previous literature, NH children over the age of 9 years are be able to use adaptive strategies, such as message revision rather than the developmentally easier message repetition strategy, in response to listener feedback (e.g., Brinton *et al.*, 1986).

Interestingly, NH speakers did not produce lexically more frequent words when talking to a HI friend compared to a NH friend, in either the Grid or the Diapix task – although, in another linguistic barrier condition, FDS, adults are found to use this strategy (Long, 1983). For a speaker to be able to use higher frequency vocabulary items 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 likely to be still developing in late childhood. Alternatively, it may be the case that the restrictive nature of the referential communication tasks used in this study do 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.

The amount of overlap used by the interlocutors was also not modified by NH children according to listener hearing status – contrary to findings of Aubanel, Cooke, Foster, García Lecumberri, and Mayo (2012), who discovered that adult speakers reduced their overlap with their interlocutor when talking in audio-only conditions compared to audiovisual conditions. Changing interlocutor overlap may however be difficult for an individual speaker to modify, as it will also be dependent on their interlocutor's speech behaviour, and may require sophisticated pragmatic skills. Additionally, it may be that the NH children in this study were already producing very little overlap with their interlocutor, which would be difficult to minimse further – Heldner and Edlund (2010) report that in the adult Map Task corpus (Anderson *et al.*, 1991), approximately 40% of time is spent in overlap compared to 6% of the time by the NH children in the current study.

### 8.4.2 Hearing-impaired children

**Linguistic speech characteristics** As expected based on teacher assessments of their language and communication skills, the HI children in this study were found to differ from their NH peers on all four linguistic measures taken from their spontaneous speech. Consistent with previous findings using MLU (e.g., Flipsen and Kangas, 2014; Geers *et al.*, 2003), the HI children used significantly fewer words per phrase than did their NH peers, suggesting that the HI children used less complex syntax in both the Diapix and Grid tasks than their NH friends. They also produced more frequent vocabulary items compared to their NH peers – confirming many previous studies' findings of a delay in expressive vocabulary development in HI children (e.g., Blamey *et al.*, 2001; Uziel *et al.*, 2007). Additionally, HI children were found to use a more diverse lexicon than their NH friends in the Grid task – implying that although the complexity of the HI children's expressive vocabulary is lower than their NH peers', they are nonetheless able to use lexical alternatives to make themselves understood. Finally, the HI children also spoke more in overlap with their interlocutor than did the NH children. It is possible that the HI children interrupted and spoke simultaneously with their interlocutor as a strategy to control the interaction by avoiding being the listener, and therefore circumventing any potential hearing difficulties – similarly to post-lingually deafened CI adults have been found to do (Caissie, Dawe, Donovan, Brooks, and MacDonald, 1998; Tye-Murray and Witt, 1996).

**Linguistic adaptations** Despite the findings of delayed expressive syntactic skills in HI children compared to NH peers, the HI children, like their NH friends, slightly reduced the number of words used in each phrase in HI-directed speech compared to NH-directed speech. It may be that syntactic simplification is a relatively easy feat, and does not require a great amount of syntactic skill, as, in spontaneous conversation, all speakers are likely to produce utterances which are both at the higher and lower ends of their syntactic abilities (Costa *et al.*, 2008). Similarly, the HI children increased the lexical diversity of their speech in the Grid task when talking with a HI friend compared to a NH friend – indicating that they further increased the number of different words in the conversation, and may have used message revision rather than repetition strategies in response to miscommunications. This interpretation is supported by findings from repair analyses done using Conversation Analysis on a subset of the data – Chu (2015), Dunn (2014), Harris (2014) and Ní Almhain (2014) report that many of the HI children examined were competent and frequent users of adaptive repair strategies, including revision and cue responses. As discussed in section 1.3.4, previous studies have found both that HI children display deficits in using adaptive repair strategies (e.g., Most, 2002), as well as that they are able to use adaptive repair strategies even to a greater extent than NH peers (e.g., Ciocci and Baran, 1998) – nevertheless, it is surprising that HI children are able to modify their speech by using a greater number of different vocabulary items in their interactions despite having smaller expressive vocabularies compared to NH children (Blamey *et al.*, 2001). These results therefore suggest that, even with delayed language and communication skills, and possible delays in theory of mind and perspective-taking, HI children are able to model their listener and monitor their needs by using these two linguistic strategies to both simplify their speech and to increase the number of different vocabulary items used in their speech – both of which likely require high levels of cognitive processing and may be cognitively fairly demanding.

Despite HI children exhibiting a greater amount of speech overlap with their interlocutor in the NH-directed condition than the NH children, they did not reduce the extent of their speech overlap in the HI-directed condition – although this strategy may have been helpful to HI interlocutors. It seems likely that, as discussed above, HI children use speech overlap to control the interaction and reduce their own hearing difficulties – but this may lead to problems in interactions involving two HI interlocutors, as neither HI child may be able to monitor both their own and the interlocutor's speech simultaneously, therefore potentially leading to greater miscommunications occurring in HI-HI dyads.

Overall, then, despite their deficits in language and communication compared to NH peers, HI children were found to be surprisingly adept at adapting to their interlocutor, as they used the same linguistic adaptation strategies as those used by their NH peers.

The previous three chapters have discovered that both NH and HI children made several global acoustic-phonetic, and some segmental and linguistic adaptations to their speech, when talking with a HI interlocutor compared to a NH interlocutor. However, it is unclear whether the changes made accounted for an intelligibility increase for listeners – this will be explored in the next chapter, chapter 9.

# Chapter 9

# Speech clarity

## 9.1 Introduction

In chapters 6 and 7, both NH and HI speakers were found to increase their speech intensity and F0 range and median, slightly decrease their speech rate, and increase their vowel space area and the between-category temporal distinctions between /i/ and /ɪ/ in HI-directed speech compared to NH-directed speech. Although many of these modifications are similar to those found in instructed clear speech studies, in which adult speakers have been found to increase their intelligibility to both NH and HI listeners in 'clear' compared to 'casual' speech (c.f., Smiljanić and Bradlow, 2009), it is unclear whether similar changes in speech are effective in enhancing the intelligibility of NH children in adverse listening conditions, due to their developing neuromotor control (see section 1.2.2). Despite making similar adaptations to NH children, HI children's speech modifications may also be less successful than NH children's in increasing their speech intelligibility in HI-directed speech, as speech modifications may be applied inconsistently by HI speakers due to the delays they exhibit in speech production compared to peers. Alternatively, as HI children's speech intelligibility is likely to be lower than NH children's already in non-adverse conditions (e.g., Montag *et al.*, 2014), HI children's speech modifications may increase their speech intelligibility in HI-directed speech to a greater extent than NH children's speech enhancements.

This chapter therefore assesses whether the NH and HI children's perceived speech clarity increases in HI-directed compared to NH-directed conditions. Because the speech samples in this study are of spontaneous speech, which was fairly uncontrolled in both content and context, perceived speech clarity as judged by naive adult listeners was used as the measure of speech intelligibility. Although most instructed clear speech studies use intelligibility tests in noise, rather than rating experiments, to examine whether 'clear' speech is more intelligible than 'casual' speech (e.g., Ferguson, 2004; Payton *et al.*, 1994), a clarity rating study has been successfully used previously to assess the intelligibility of speech modifications in spontaneous speech (Hazan and Baker, 2011), and clarity judgments are also frequently used to rate the intelligibility of both NH (e.g., Redford, 2014) and HI (e.g., Elfenbein *et al.*, 1994; Flipsen, 2008) children's spontaneous speech. Additionally, a study by Ferguson and Kerr (2008) has shown there to be a correlation between the subjective rated clarity of speech and the intelligibility of the same samples in an objective listening test. However, there is also some evidence that objective speech intelligibility and subjective speech clarity judgments are not necessarily completely equivalent to each other (e.g., Eisenberg, Dirks, Takayanagi, and Martinez, 1998; van Buuren, Festen, and Plomp, 1995), and therefore speech clarity ratings as reflecting speech intelligibility should be treated with caution. In this study, using perceived clarity ratings and regression analyses enables us to examine the individual differences between speakers in their perceived speech clarity, and the relative contribution of acoustic measures to perceived clarity in the NH and HI speakers.

## 9.2 Method

### 9.2.1 File processing

For each of the 34 speakers, four short speech snippets per condition per task were extracted from the recordings, leading to a total of 544 snippets (34 speakers x 2 conditions (NHD, HID) x 4 snippets x 2 tasks). The snippets were approximately 2-3 seconds long, reasonably self-contained utterances containing a minimum of four words. They were whole intonational phrases or ends of phrases, and did not occur directly after a miscommunication. Snippets with multiple disfluencies or long pauses were avoided. This approach is identical to that used in Hazan and Baker (2011).

In the Grid task, the four snippets were taken only from the speaker's description of their picture, as these were found to contain the longest speech samples. All snippets were extracted from the middle of each conversation; for speakers who completed only two Grid tasks in a condition, the snippets were taken from as close as possible to one-third and two-thirds into each file. For speakers who completed three Grid tasks in a condition, snippets were chosen close to halfway into the first and third files, and one-third and two-thirds in to the second file. For speakers who completed all four Grid tasks in a condition, the snippets were extracted close to the midpoint of each file.

In the Diapix task, for speakers who completed only one Diapix picture in a condition, the snippets were taken as close as possible to the 10th, 15th, 20th and 25th turns in the conversation[1]. For speakers who completed two Diapix pictures in a condition, snippets were taken from as close as possible to the 10th and 20th turns in the file.

All snippets were normalised for intensity using a Praat script.

### 9.2.2 Raters

Twenty monolingual Southern British English speakers (15F, 5M; mean age: 21.6 years; age range: 18.9-26.3 years) participated in the rating study. According to a post-study questionnaire, 12 participants had had some phonetic training, and 4 participants had substantial experience of communicating with a person with a hearing-impairment. They reported having no speech or language disorders.

A hearing screening test conducted at octave frequencies between 250 Hz and 8000 Hz demonstrated that all participants had normal hearing thresholds (25 dB HL or better). All participants were paid for their participation and were naive as to the purpose of the study.

[1]This was done to enable a comparison to be done to the NH children recorded in (Hazan *et al.*, submitted), and the NH adults' Diapix data from Hazan and Baker (2011)

### 9.2.3 Testing procedure

The snippets extracted from the Grid and Diapix tasks were separated into two short rating experiments, each lasting approximately 20-25 minutes. Praat's ExperimentMFC function was used to present the stimuli through Beyerdynamic DT297PV headphones at a comfortable loudness level; the snippets within each experiment were randomised for each rater.

Participants sat facing a computer screen in an acoustically treated booth, and were asked to listen to each snippet and rate them on a rating scale between 1 ('very clear') to 7 ('not very clear'), which was presented on the screen along with the question 'How clear is this?'. Before the start of the study, participants were told that they would hear a range of children's voices, but that they should try to ignore the quality of the voices[1] and not to pay attention to any accent or grammatical correctness, but to only rate the snippets according to how clear they think the speech is, i.e. 'how easy it would be to understand in a noisy environment' (following instructions given to participants in Hazan and Baker, 2011). The raters were also told that the recordings had been made at the children's schools and that therefore there may be some noise in the recordings, which they were asked to ignore. The experimenter did not tell the raters that they would be listening to hearing-impaired children's speech, nor were they told how the speech had been elicited.

All participants completed both rating experiments in one session of about 55 minutes. Half of the participants started the session by rating the Grid task snippets, while the other half began by rating the snippets extracted from the Diapix task. The participants were given a short break between experiments, during which they completed the hearing screening test and were asked to fill out a background questionnaire.

A post-study questionnaire was given to all raters at the end of the study. When asked 'Did you recognise how the voices differed from each other?', five participants answered that they thought some of the speakers had a hearing-impairment. A further 10 participants mentioned speech impairment or speech difficulties in some of the speakers. The remaining five participants only remarked

[1] This instruction was given to avoid raters automatically giving low ratings to children with more obvious speech difficulties.

on between-speaker differences in accent, gender or age.

### 9.2.4 File processing

As each speech snippet was rated once by each rater, a total of 10,880 clarity ratings (20 raters x 544 snippets) were obtained in the experiment. The mean clarity rating per snippet was calculated for statistical analysis.

## 9.3 Results

### 9.3.1 Inter-rater reliability

To examine whether the raters were consistent with their rating of perceived clarity, an Intra-class Correlation Coefficient (ICC) analysis was run on the rating made by each rater for each snippet. A two-way mixed model was run with rater as the fixed effect. The average measures ICC for Cronbach's $\alpha$ was 0.98, which suggests a very strong reliability between raters (Streiner and Norman, 2002)[1]. Therefore the results indicate very good inter-rater consistency for the perceived clarity measure.

### 9.3.2 Speech clarity rating

The same statistical approach using linear mixed effects models as in section 6.3.1 was used, with speaker hearing status, listener hearing status and task as fixed factors. The random effects of speaker, age and snippet number were tested for inclusion in the model. The final model included only the random factor of age.

The three-way interaction between speaker hearing status, listener hearing status and task was not significant (p=0.37). None of the two-way interactions were significant either (all p>0.54). The only significant factor in the model was speaker hearing status($\chi 2(5)$ =30.56, p<0.0001, r=-0.76) - NH participants (mean: 1.88) were consistently rated as clearer than HI participants (mean: 4.08)

[1]All values of Cronbach's $\alpha$ over 0.7 indicate substantial agreement between raters; the maximum value is 1. The average measures ICC value was chosen as the following analyses are all conducted on the average perceived clarity scores given among all raters in this study.

(see Table 9.1). Surprisingly, despite the findings in chapter 6 showing that both NH and HI speakers enhanced the acoustic-phonetic aspects of their speech in response to a HI interlocutor, neither group was rated as more intelligible in the HI-directed condition, compared to the NH-directed condition. The speech from both tasks was also rated as equally clear.

| task | NH-NHD | NH-HID | HI-NHD | HI-HID |
|---|---|---|---|---|
| Grid | 1.86 (0.59) | 1.82 (0.63) | 4.02 (1.54) | 4.13 (1.33) |
| Diapix | 1.91 (0.66) | 1.91 (0.79) | 4.11 (1.30) | 4.05 (1.34) |

Table 9.1: Means and standard deviations for the speech clarity rating experiment.

### 9.3.3 Predicting speech clarity ratings

#### 9.3.3.1 Individual variability in clarity

The findings from the previous section demonstrated that the only significant differences in perceived clarity were related to speaker hearing status - although chapters 6 and 7 found that speakers did enhance several acoustic-phonetic aspects of their speech, the hearing status of the listener had no influence on perceived clarity ratings of the speaker. However, a large amount of variability in clarity ratings was observed between participants. Figure 9.1 shows the 95% confidence intervals for perceived speech clarity for each speaker in the NH-directed condition. The figure indicates that the speakers can be divided into three distinct near non-overlapping groups on the basis of their perceived clarity. Interestingly, the clarity of speakers seems to be related to their hearing level. The first group comprises of only NH participants, all 11 of whom are consistently rated as very clear, with mean ratings between 1 and 2. The second group includes both NH participants (6 speakers) and HI participants (4 speakers), whose mean rating is between 2 and 3. Of the four HI participants in this group (HI6, HI13, HI1 and HI12), three have only moderate or moderate-to-severe hearing loss. The third group, whose mean clarity ratings are over 3, is very variable, and only consists of HI participants (13 speakers). Of the seven participants who are rated the least

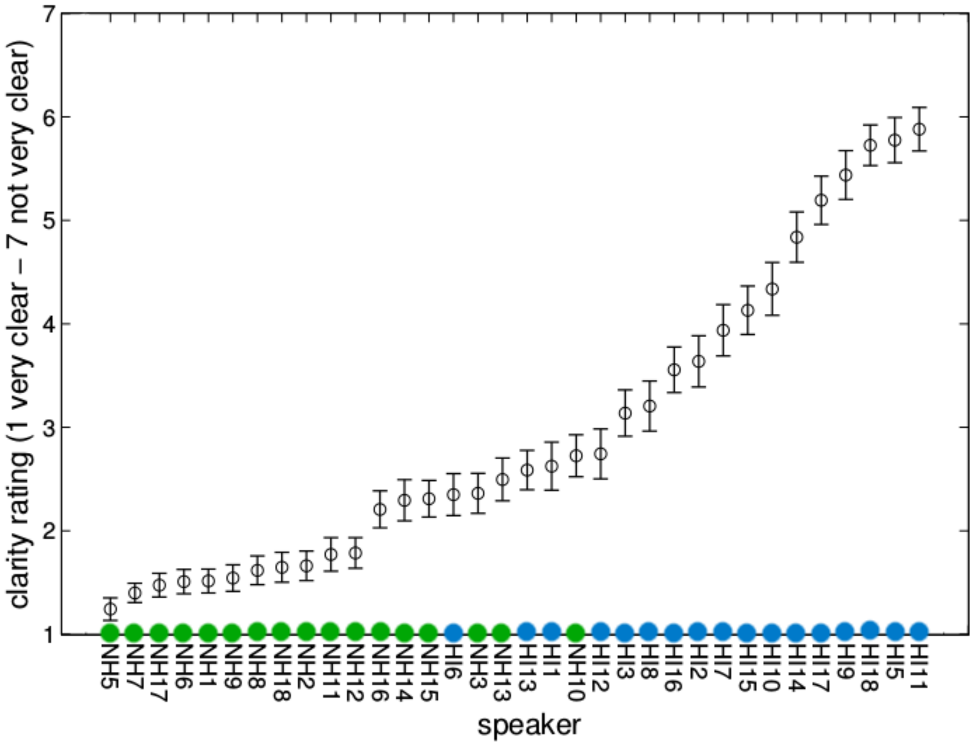



Figure 9.1: Clarity rating in the NH-directed condition for each speaker. NH participants are indicated by a green circle, while HI participants are shown with a blue circle. Error bars demonstrate 95% confidence intervals from the mean.

clear, six have profound hearing loss. Only one participant with profound hearing loss (HI3) is rated as clearer than average in the HI group, indicating that the level of hearing loss nevertheless does not account for all variation in perceived clarity. Two out of the seven HI children who wear CIs (HI3 and HI8) are rated as being perceptually clearer than the average HI child in this group.

#### 9.3.3.2 Predicting clarity from global acoustic measures

The above figure demonstrates particularly high variability in the perceived clarity of the HI participants' speech. It is therefore of interest to explore which acoustic-phonetic aspects of the participants' speech account for this variation.

Following Hazan and Markham (2004), a forward stepwise multiple linear regression analysis[1] was conducted on the NH and HI participants' clarity ratings

[1]Several studies have reported a wide range of acoustic-phonetic factors influencing speech

separately. As no differences in clarity ratings were found between NH- and HI-directed speech, cases from both conditions were included in the analysis, as were acoustic-phonetic measures taken from both the Grid and Diapix tasks. Of the acoustic-phonetic measures, F0 median and F1/F2 range were excluded due to being highly correlated with the intensity and vowel space area measures, respectively[1]. Therefore, measures of speech rate (syllables per second), speech intensity (mean energy 1-3kHz), F0 range (semitones) and vowel space area (ERB2)[2] were included in the analysis, with all measures averaged per speaker per task per condition, resulting in 68 data cases for the HI participants, and 66 data cases for the NH participants[3].

The step function in the 'stats' package for R, which exploits AIC in determining the best model, was used to perform the regression. For NH speakers, the final model included only vowel space area as a predictor (R-squared: 0.05; adjusted R-squared: 0.03) (see table 9.2). However, its contribution to the model is only near-significant, and it accounts for very little of the variance. Therefore none of the four tested acoustic-phonetic measures seem to accurately predict clarity ratings for the NH speakers.

For the HI speakers, the final model included the predictors speech intensity, speech rate and vowel space area (R-squared: 0.31; adjusted R-squared: 0.27) (see table 9.3). The residuals and assumptions of no multicollinearity were checked for the data. This model therefore accounts for 30.7% of the variance in clarity ratings for the HI speakers. Intensity, which has the largest amount of influence on the model, alone accounts for 18.4% of the variation, while the second largest factor influencing the model, speech rate, alone accounts for 17.1% of the variation. Vowel area has only a near-significant contribution to the model, and alone accounts for under 1% of the variance. As can be seen from figures 9.2 and 9.3, the quieter and faster the speech, the clearer it is rated. In chapter 6 it was found

intelligibility, preventing a theoretically-driven hierarchical model being used for analysis.

[1]F0 median was excluded rather than intensity due to previous studies indicating that F0 median changes are likely to occur due to increased speech intensity (Titze, 1989). F1/F2 range measures were excluded as the vowel space area was seen to represent both measures.

[2]The vowel data were included with the global acoustic-phonetic predictors as it has been traditionally used to predict intelligibility in previous studies. Additionally, it was calculated over both Diapix and Grid tasks.

[3]Missing values for one condition were removed for NH2 before conducting the analysis.

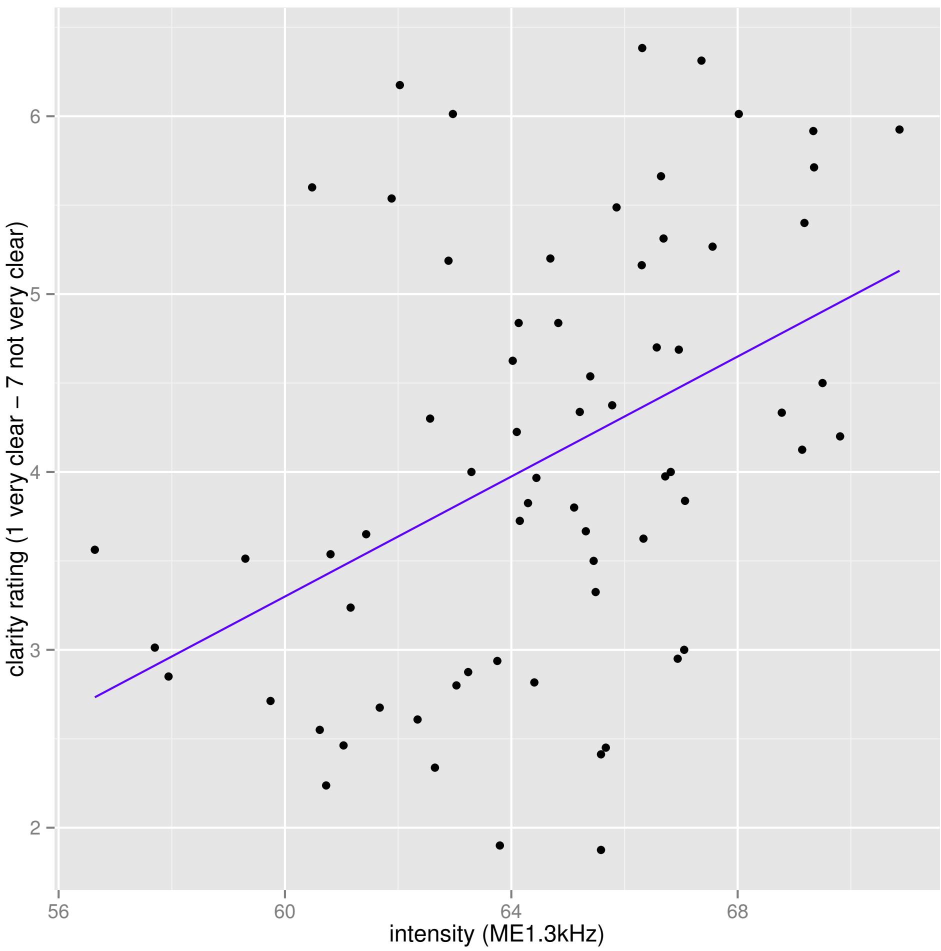



Figure 9.2: The relationship between clarity rating and speech intensity for the HI speakers.

that HI participants spoke more slowly and more loudly than NH participants. Interestingly, although speaking more slowly and more loudly is also a feature of clear speech, in the HI participants' case, it seemed to make speakers less clear. A greater speech intensity and a slower speech rate may be a feature of speech for those HI participants whose speech is most impaired by their hearing loss, such as the participants with profound hearing loss who were rated the least clear in the NH-directed condition (see Figure 9.1). This may explain why a slower speech rate and a greater speech intensity are associated with perceptively less clear speech for the HI talker group.

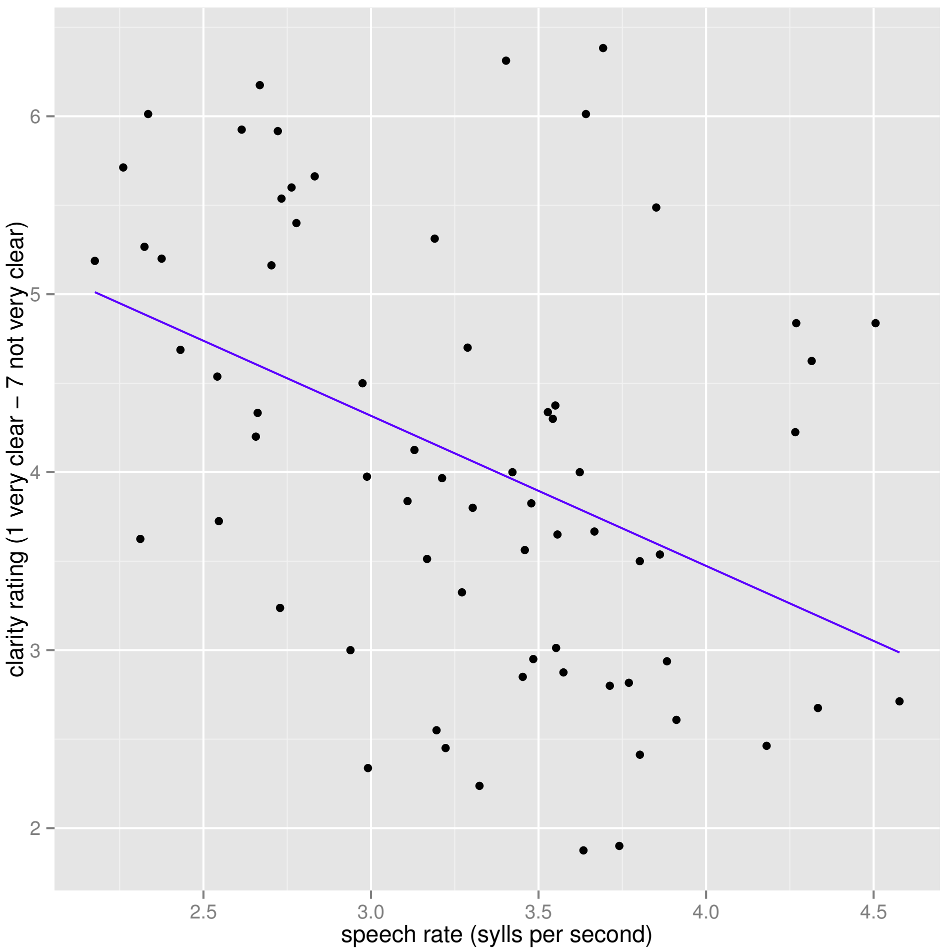



Figure 9.3: The relationship between clarity rating and syllable rate for the HI speakers.

| | B | SE B | t | p |
|---|---|---|---|---|
| (Intercept) | 2.57 | 0.40 | 6.45 | <0.0001*** |
| vowel area | -0.03 | 0.02 | -1.74 | 0.087(*) |

Table 9.2: Result of the multiple regression on global acoustic-phonetic measures for NH speakers.

In summary, for NH speakers, no global acoustic-phonetic measure was found to predict perceived clarity significantly. For HI speakers, louder and slower speech predicted less clear speech - perhaps due to these acoustic-phonetic aspects

| | B | SE B | t | p |
|---|---|---|---|---|
| (Intercept) | -3.27 | 3.18 | -1.03 | 0.31 |
| intensity | 0.16 | 0.05 | 3.50 | 0.0009*** |
| speech rate | -0.66 | 0.23 | -2.91 | 0.005** |
| vowel area | -0.06 | 0.03 | -1.94 | 0.057(*) |

Table 9.3: Result of the multiple regression on global acoustic-phonetic measures for HI speakers.

of speech being a feature of reduced speech motor control and speech planning. The HI speakers who speak more quietly and faster may be better able to control their speech production, which in turn may make their speech perceptually clearer.

#### 9.3.3.3 Predicting clarity from category discriminability

The above analysis indicated that global acoustic-phonetic measures were unable to predict speech clarity for NH speakers, and HI speakers were judged as clearer if they spoke less loudly and faster than others. It may be that more fine-grained acoustic-phonetic aspects of the participants' speech play a role in predicting their perceived clarity. In particular, when judging clarity, raters may be sensitive to the discriminability of phonetic categories in speech.

To examine this question, a further forward stepwise multiple linear regression analysis was conducted on the category discriminability data obtained from the Grid task. Therefore, for this analysis, only the clarity ratings for the Grid task were used. As none of the four category discriminability measures (spectral discriminability between /i/-/ɪ/, durational discriminability between /i/-/ɪ/, fricative /s/-/ʃ/ discriminability and VOT /p/-/b/ discriminability) were highly correlated with one another (see section 7.3.3), all of them were included in the model as potential predictors of perceived speech clarity. As in the previous section, NH and HI speakers were analysed separately, with 34 data cases for each group.

The results of the NH group can be found in Table 9.4 (R-squared: 0.41; adjusted R-squared: 0.35). The model was checked for non-multicollinearity and

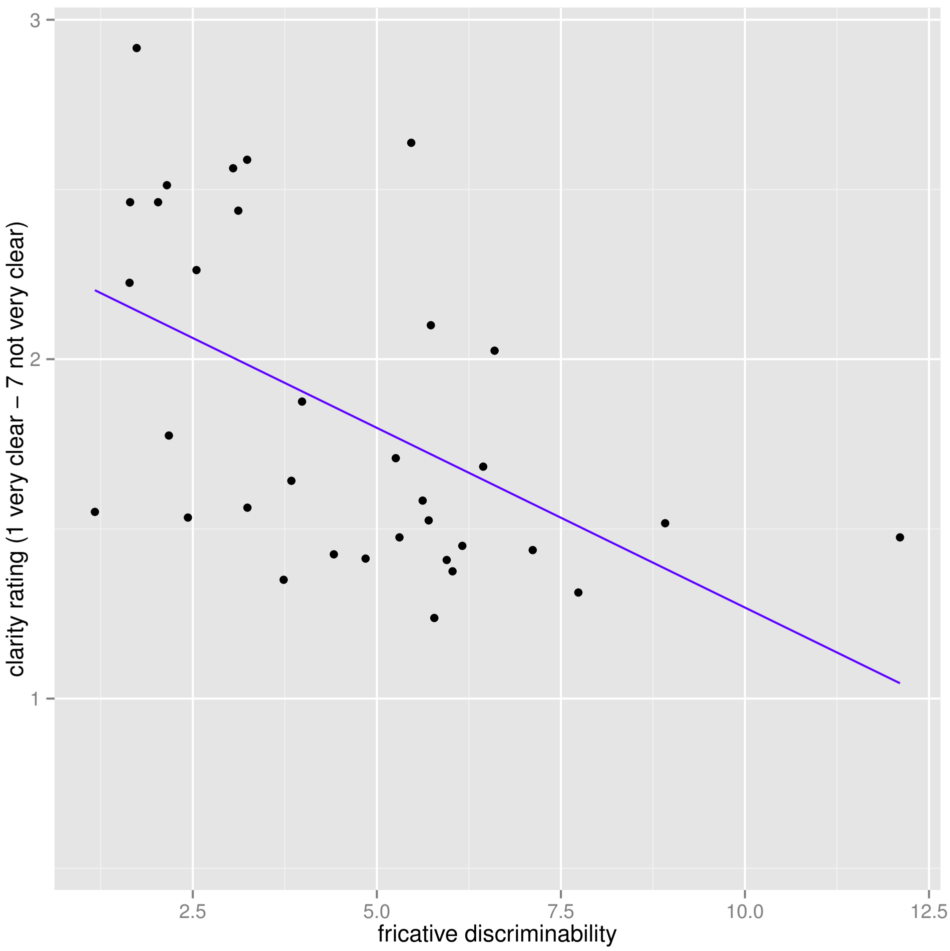



Figure 9.4: The relationship between clarity rating and the spectral discriminability of fricatives for the NH speakers.

residuals. The model accounts for 41% of the variation found, and includes fricative discriminability, the durational discriminability between the tense/lax vowels, and the spectral discriminability between the tense/lax vowels as predictors. Of these, only fricative discriminability and durational discriminability of /i/-/ɪ/ make a significant contribution to the model. On its own, the discriminability in fricatives accounts for 27% (adjusted R-squared: 0.25) of the variance in NH speakers' perceived clarity, while durational discriminability alone accounts for less than 1% of the variation. Therefore greater fricative spectral discriminability seems to be the most important factor in predicting greater perceived speech clarity for the NH speakers (see Figure 9.4).

| | B | SE B | t | p |
|---|---|---|---|---|
| (Intercept) | 2.05 | 0.25 | 8.29 | <0.0001*** |
| fricatives | -0.09 | 0.03 | -2.93 | <0.01** |
| tense/lax duration | 0.45 | 0.20 | 2.25 | 0.03* |
| tense/lax spectral | -0.08 | 0.04 | -1.83 | 0.077(*) |

Table 9.4: Result of the multiple regression on the segmental discriminability measures for NH speakers.

Table 9.5 displays the results of the regression for the HI speakers (R-squared: 0.45; adjusted R-squared: 0.40). Non-multicollinearity and residuals were checked for the model. The model includes spectral discriminability between the tense and lax vowels (/i/-/ɪ/), fricative spectral discriminability and the discriminability between /p/-/b/, and accounts for 45% of the variability in speech clarity. Of the predictors in the model, however, only the tense/lax spectral discriminability significantly adds to the model, and it alone predicts 36% of the variance (adjusted R-squared: 0.34) (see Figure 9.5). Tense/lax spectral discriminability therefore appears to be the most important factor in predicting speech clarity judgments for the HI speakers, with greater spectral discriminability predicting clearer speech. As with greater speech intensity and slower speech rate, which were found to predict less clear speech ratings in the previous section, smaller spectral discriminability between the tense/lax categories is likely to be a feature of less accurate speech motor control, leading to speakers with smaller spectral discriminability between /i/-/ɪ/ being judged as less clear.

| | B | SE B | t | p |
|---|---|---|---|---|
| (Intercept) | 6.24 | 0.60 | 10.33 | <0.0001*** |
| tense/lax spectral | -0.53 | 0.15 | -3.62 | 0.0011** |
| fricatives | -0.29 | 0.15 | -1.96 | 0.059(*) |
| VOT | -0.24 | 0.15 | -1.62 | 0.12 (n.s.) |

Table 9.5: Result of the multiple regression on the segmental discriminability measures for HI speakers.

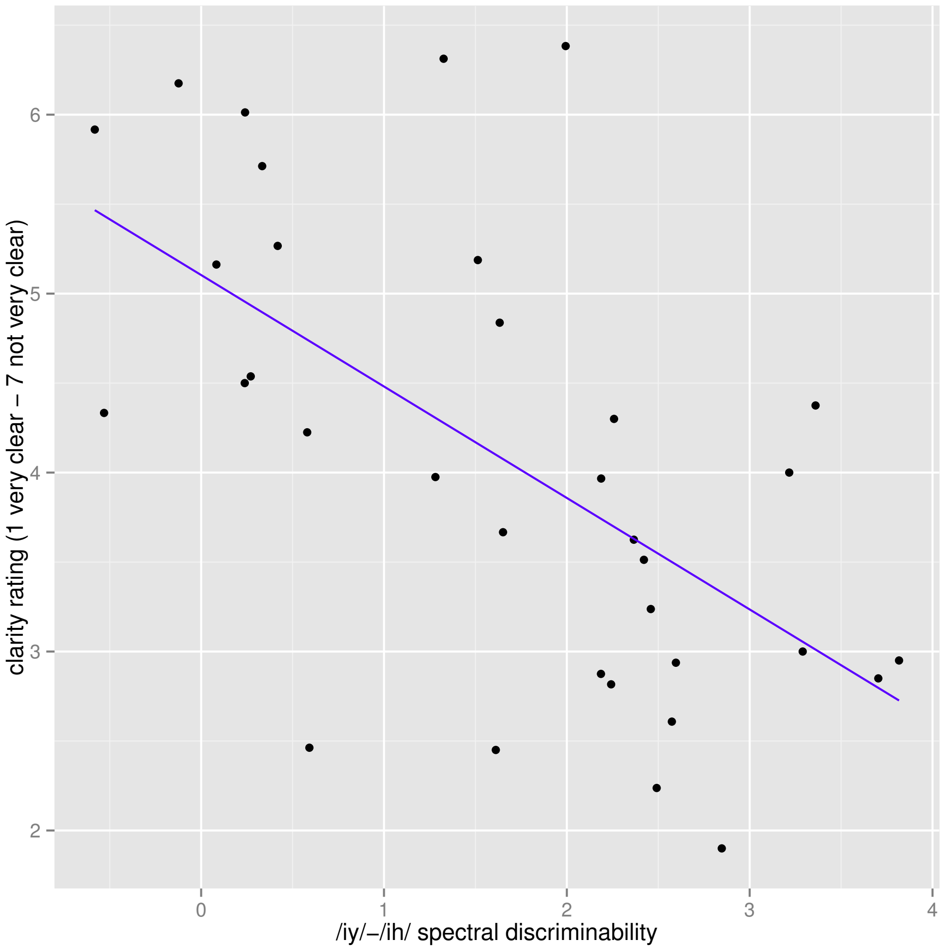



Figure 9.5: The relationship between clarity rating and the spectral discriminability of the tense/lax vowels for the HI speakers.

In summary, in terms of segmental contrasts, NH speakers' speech clarity was predicted to some extent by greater fricative spectral discriminability and greater tense/lax durational distinctions. Greater speech clarity for HI speakers was associated with greater spectral distinctions in both tense/lax vowels and fricatives.

## 9.4 Discussion

Previous studies assessing whether instructed ‘clear’ speech is more intelligible than ‘casual’ speech have demonstrated that the speech enhancements made by adult speakers in clear speech are beneficial to both NH and HI listeners in noise (e.g., Ferguson, 2004, 2012; Helfer, 1997; Picheny *et al.*, 1985), and a clarity rating study similar to that used in the current study has shown that adults’ spontaneous speech spoken to a listener hearing them through a simulated CI is rated as perceptually clearer than their speech spoken to a listener hearing them normally (Hazan and Baker, 2011). However, in the current chapter, we found no differences in perceived clarity between NH and HI children’s speech spoken to a HI friend compared to a NH friend, despite the findings of previous chapters indicating that acoustic-phonetic modifications were made in the HI-directed condition compared to the NH-directed condition.

It is possible that NH children are unable to enhance their intelligibility adequately using acoustic-phonetic speech modifications, perhaps due to their still-developing speech production skills. However, in this experiment, NH children’s speech clarity was rated as very clear even in the NH-directed condition (see Table 9.1), and therefore there was likely to be little scope for raters to rate the HI-directed condition as being perceptually clearer than the NH-directed condition. Additionally, it is possible that NH children only enhance their speech when they encounter miscommunications – therefore, as the speech snippets in this study were not taken after misunderstandings, the differences between the NH- and HI-directed snippets may not have been large. The only significant difference within the experiment was that NH children were rated as perceptually clearer than HI children, whose speech clarity was rated approximately halfway between ‘very clear’ and ‘not very clear’. HI children’s perceived speech clarity would therefore have been able to improve in the HI-directed condition compared to the NH-directed condition – the fact that such a difference between conditions was not observed could be interpreted as HI children’s speech modifications not being effective in enhancing their perceived intelligibility in HI-directed compared to NH-directed conditions.

Alternatively, it seems likely that this result is due to experimental factors.

In the only known previous study to have assessed the speech intelligibility of children's spontaneous speech adaptations, Syrett and Kawahara (2014) found that words produced by preschool-aged children when teaching a puppet were perceptually distinguishable from the same words produced when talking casually with an experimenter – however, in Syrett and Kawahara's (2014) study, listeners' task was merely to discriminate between 'normal' and 'clear' speech conditions – thus they were told from the outset that they would be listening to different speaking conditions, which may have biased them to pay careful attention to certain acoustic-phonetic parameters which they may not otherwise have attended to. In the current study, listeners were not told that they were listening to different speaking conditions or even to children with hearing loss – raters were therefore rating the speech snippets according to their own interpretation of speech clarity. As 15 out of the 20 raters reported having noticed that some of the speakers had a speech impairment and/or speech characteristics associated with a hearing-impairment, raters may have been using judgment parameters relating to perhaps more salient between-speaker differences, such as perceived speech impairment or fluency, rather than within-speaker differences relating to speaking condition. Thus, although the raters would have had scope to rate the HI speakers as clearer in the HI-directed condition than in the NH-directed condition, it is likely that listeners were not paying attention to such subtle within-speaker differences.

Indeed, in the previous study using clarity ratings for adults' spontaneous speech adaptations (Hazan and Baker, 2011), the rated speech was produced by a very homogenous group of speakers, who were all young adults of approximately similar ages and the same accent group – and therefore it may be that the only salient parameter that raters were able to use to rate the clarity of the speech was the within-speaker differences relating to speaking condition. However, clarity rating studies may not be as appropriate for judging differences between speaking conditions in studies, such as the current one, with a more heterogenous group of speakers, whose speech may be judged according to a wide range of parameters. Similarly, the basis on which raters judge the stimuli in clarity rating studies is somewhat unknown, as raters may have different ideas of what may be 'clear' – and in this study, the various instructions of judging according to 'speech clarity' and 'how easy it would be to understand in a noisy environment' may have been confusing to raters. In future work, keywords could be extracted from the Grid task to present to listeners for distinguishing between NH-directed and HI-directed speaking styles, similarly to Syrett and Kawahara (2014), or keywords could be mixed with noise to examine whether speed of identification is greater in different speaking conditions, as in Grynpas *et al.* (2011). Crucially, the raters in the current study were not hearing-impaired, despite the speech being specifically HI-directed speech – ideally, any future study would therefore use HI participants as listeners.

Although it may be difficult to assess whether the speakers' perceived speech clarity differed according to listener hearing status on the basis of the current chapter, the results demonstrate that the HI children were perceived as having perceptually less clear speech than NH children, as expected. Additionally, as seen in Figure 9.1, the HI children's speech varied greatly between speakers, with more clearly perceived speech being related to a faster speech rate, a decreased speech intensity, and an increased spectral discriminability of /i/-/ɪ/ and /s/-/ʃ/ compared to less clearly perceived speech. These findings are consistent with the notion that raters were rating the fluency of the HI children's speech, with faster speech rates, more quietly spoken speech, and greater spectral phoneme discriminability likely implying a greater amount of articulatory control, faster speech planning and more robust phonetic categories in a speaker (c.f., Burkholder and Pisoni, 2003; Redford, 2014).

As discussed in Montag *et al.* (2014), the production of highly intelligible speech is an important indicator of high general expressive language skills in HI children, as it reflects their ability to perceive speech adequately, to have sufficient linguistic skill to plan and formulate utterances, as well as to possess required speech motor control to enable the articulation of utterances. Therefore high speech intelligibility in HI children likely reflects high levels of language proficiency – accordingly, HI children's speech intelligibility has been found to correlate with clinical measures of both receptive and expressive language (e.g., Montag *et al.*, 2014). The mean perceived clarity measure for HI speakers in the NH-directed condition is therefore used in the following chapter, chapter 10, as a general measure of the HI children's language proficiency, when relating the individual variability in adaptation strategies used by speakers to both speaker and listener factors.

# Chapter 10

# Individual Variability and Factors Affecting Adaptation

## 10.1 Introduction

The findings from chapters 6, 7 and 8 suggest that, when talking with a HI friend, both NH and HI speakers make a range of acoustic-phonetic and linguistic adaptations to their speech. As regards the measures investigated in this study, they increase their F0 median, F0 range, and intensity, slightly decrease their articulation rate, somewhat increase their vowel space area and their F1 range, increase the between-category temporal distance between /i/ and /ɪ/ vowels, slightly decrease the number of words used per phrase, and increase the lexical diversity of their speech in the Grid task. However, so far, all of these analyses have been conducted per speaker group – thus it is unclear whether the great majority of participants within the group are using a specific enhancement strategy, or whether strategy use varies between individuals, as typically found in instructed clear speech studies (e.g., Ferguson and Kewley-Port, 2007; Krause and Braida, 2004; Picheny *et al.*, 1986).

Several potential factors may influence whether strategies are used and the extent to which they are applied. In this study, listener-related factors such as the interlocutor's language skills may cause differences between speakers' strategy use – as seen in chapters 4 and 9, the HI interlocutors in this study varied greatly in their speech perception skills and in their perceived speech clarity. Adult speakers have indeed been found to make adaptations according to interlocutor's language proficiency (Warren-Leubecker and Bohannon, 1982) or the child listener's language level (Roy *et al.*, 2009). The models reviewed in section 1.3.2 suggest that adults do this by monitoring the listener for signs of miscomprehension or explicit listener feedback (Brown and Dell, 1987; Fussell and Krauss, 1992; Lindblom, 1990; Schober, 1993). Although there is some evidence that NH preschoolers are be able to change their MLU according to a younger listener's language skill (Masur, 1978; Warren-Leubecker and Bohannon, 1983), Hazan *et al.*'s (submitted) results suggest that older NH children may be less sensitive to the specific needs of a listener than adults. Therefore it is possible that NH and HI children do not enhance their speech specifically to the needs of a HI interlocutor, but use general enhancement strategies regardless of interlocutor communication difficulty.

Strategy variability may also stem from speaker-related factors, such as a speaker's age (as in Pettinato *et al.*, 2016), or their language skills – although as a group, the HI speakers were found to modify their speech similarly to NH peers, those HI speakers with lower speech and language skills may be unable to enhance their speech, due to possible delays in perspective-taking, responding to feedback, and speech and language inflexibility (see section 1.3.4). Another potential factor in strategy use is that speakers may treat different types of adaptations as alternatives – thus, some speakers may prefer to use, for example, linguistic adaptation strategies over acoustic-phonetic ones.

This chapter, therefore, aims to explore individual variability in strategy use between the speakers in the current study. The extent to which adaptations are made by each speaker is related, firstly, to listener-related factors of interlocutor perceived speech level (as measured using the perceived clarity measure from chapter 9), and to speaker-related factors of speaker's perceived speech level, to examine whether these factors account for some of the between-speaker variability in enhancement strategies found. Additionally, we examine whether different types of enhancements are used as alternative strategies when talking to a HI peer, and finally explore the individual strategies used by each speaker.

## 10.2 File Processing

### 10.2.1 Conversion to z-scores

Z-scores were used to measure the extent to which a speaker adapted to their interlocutor relative to other participants in their speaker group (HI or NH). It provides a more reliable estimate of a speaker's overall amount of speech change relative to others than do percentages, as the conversion to z-scores normalises for any inherent differences between measures in the extent to which they can be modified by a speaker[1] (for a similar approach, see Hazan, Messaoud-Galusi, Rosen, Nouwens, and Shakespeare, 2009).

For each type of adaptation (global, segmental and linguistic), an overall z-score per speaker was calculated to reflect the relative extent to which a speaker made adaptations using that type of strategy. This consisted of four steps: first, using the formula (HID-NHD)/NHD, where HID is the value in the HI-directed condition, and NHD is the value in the NH-directed condition, the mean relative change between HI-directed and NH-directed conditions was obtained in each of four measures separately. Then, the z-score of each individual measure was calculated for each speaker within each participant group. Finally, the mean z-score per adaptation type was obtained by averaging over the four individual measure z-scores per speaker. Thus z-scores reflect a broad measure of the extent of change made in relation to others in the speaker group. A positive z-score reflects a higher than average amount of adaptation made by the speaker, while a negative z-score indicates that the speaker used fewer adaptations than average.

### 10.2.2 Global z-score

To obtain a measure of the relative degree of global acoustic-phonetic change for each participant, the mean value per condition of four global acoustic-phonetic measures (syllable rate, F0 range, F0 median and mean energy 1-3k Hz) was taken over all files (both Grid and Diapix). The measure of number of pauses was not used, as values were only available for the Diapix task (see section 6.2.4).

[1]For example, a speaker is likely to be able to make greater changes to their F0 range than to their median F0.

The z-scores were calculated as described in section 10.2.1. However, for syllable rate, relative change was calculated using the formula (NHD-HID)/NHD, as an enhancement in this measure (decreased articulation rate) is associated with fewer syllables articulated per second. For the two NH participants for whom F0 range values (NH2) or F0 median values (NH13) were not available (see section 6.2.1), the mean of the remaining three measures were used to calculate the composite z-score.

### 10.2.3 Segmental z-score

The measure of overall relative change in segmental measures was obtained for four segmental acoustic-phonetic measures (vowel space area, /s/-/ʃ/ discriminability, /p/-/b/ discriminability and /i/-/ɪ/ overall discriminability[1]). The discriminability measures were used as they provide information on both between- and within-category distance, and the vowel space area measure was chosen as it incorporates both F1 range and F2 range. To obtain the mean overall segmental z-scores per participant in each speaker group, the procedure described in 10.2.1 was used.

### 10.2.4 Linguistic z-score

Measures of the number of words used per phrase, lexical diversity, lexical frequency, and the amount of overlap were used to obtain a measure of the relative overall change in linguistic measures by each participant. For overlap and number of words per phrase, the relative change per speaker was calculated in reverse (as for syllable rate in section 10.2.3) – i.e., as (NHD-HID)/NHD – as an enhancement in those measures involves less overlap and a smaller number of words per phrase being produced. The z-scores per participant within each speaker group were calculated as in section 10.2.1 above.

[1]This was calculated as the mean discriminability over the spectral and temporal discriminability measures.

### 10.2.5 Composite overall z-score

To obtain a measure of the overall extent of modifications made by each participant in the HI-directed compared to the NH-directed conditions, the mean of the global change z-score, the segmental change z-score and the linguistic change z-score was taken per speaker. This 'composite' score therefore reflects the mean extent to which each speaker made adaptations to their listener over all measures, relative to the other participants in their speaker group.

### 10.2.6 Perceived speech level measure

As a measure of the HI participants' perceived speech level, their mean perceived clarity in the NH-directed condition over both the Diapix and Grid tasks was calculated (see section 9.4 for discussion)[1].

## 10.3 Results

### 10.3.1 Listener perceived speech level

First, to explore whether the degree to which speakers' enhancement is related to the extent to which their listener had difficulty in speech and language, correlations were run between NH and HI speakers' relative extent of change in measures (z-scores) and their interlocutor's perceived speech level (speech clarity measure).

**NH speakers** The mean overall change across all measures by NH speakers between the NH-directed condition and the HI-directed condition (in mean composite z-scores) significantly correlated with the mean perceived clarity of their HI interlocutor's speech ($n=17$; $p<0.01$, $r=0.665$, $R^2=0.442$) (see Figure 10.1). This finding implies that those NH speakers whose HI interlocutor's perceived speech level was lower made more overall adaptations to their interlocutor than those NH speakers whose HI interlocutor's language level was higher.

[1] Here, the term 'perceived speech level' is taken from the speech clarity measure to mean the perceived competence or proficiency of the speaker.

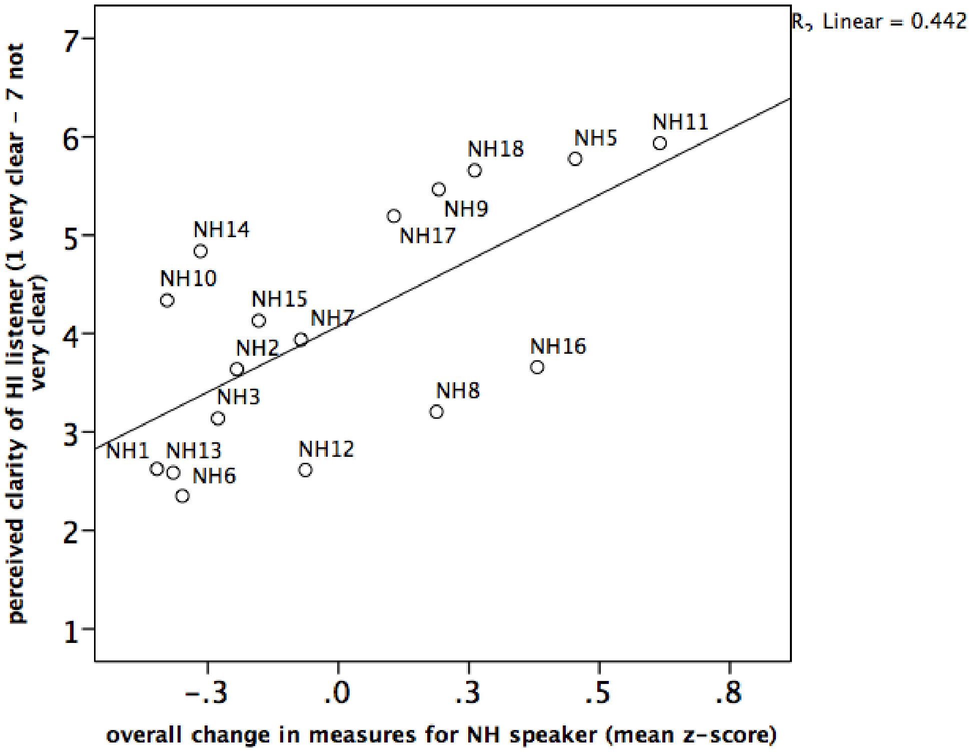



Figure 10.1: Scatterplot of the mean perceived clarity rating of the HI listener and the mean z-score of the overall change in measures from the NH-directed to the HI-directed condition for the NH speakers.

When inspecting the relationship between the extent to which NH speakers modified their speech in each type of enhancement (global, segmental and linguistic) separately and the HI interlocutor's mean perceived clarity, only the extent of global adaptations made was significantly correlated with HI interlocutor perceived clarity (see Table 10.1).

| data | r | p | $R^2$ |
|---|---|---|---|
| global-HI mean clarity | 0.619 | <0.01** | 0.383 |
| segmental-HI mean clarity | 0.402 | 0.11, n.s. | NA |
| linguistic-HI mean clarity | -0.052 | 0.84 | NA |

Table 10.1: Correlations between global, segmental and linguistic enhancement z-scores and HI interlocutor mean clarity rating in the NH-directed condition for NH speakers (n=17).

**HI speakers** As shown by Figure 10.2, the correlation between the amount of change in overall measures by HI speakers and the perceived speech level of their HI interlocutor was significant (n=16, r=0.503, p=0.047, $R^2$=0.253)[1] – those HI speakers who were paired with HI interlocutors with lower perceived speech levels made greater adaptations to their speech and language in HI-directed speech.

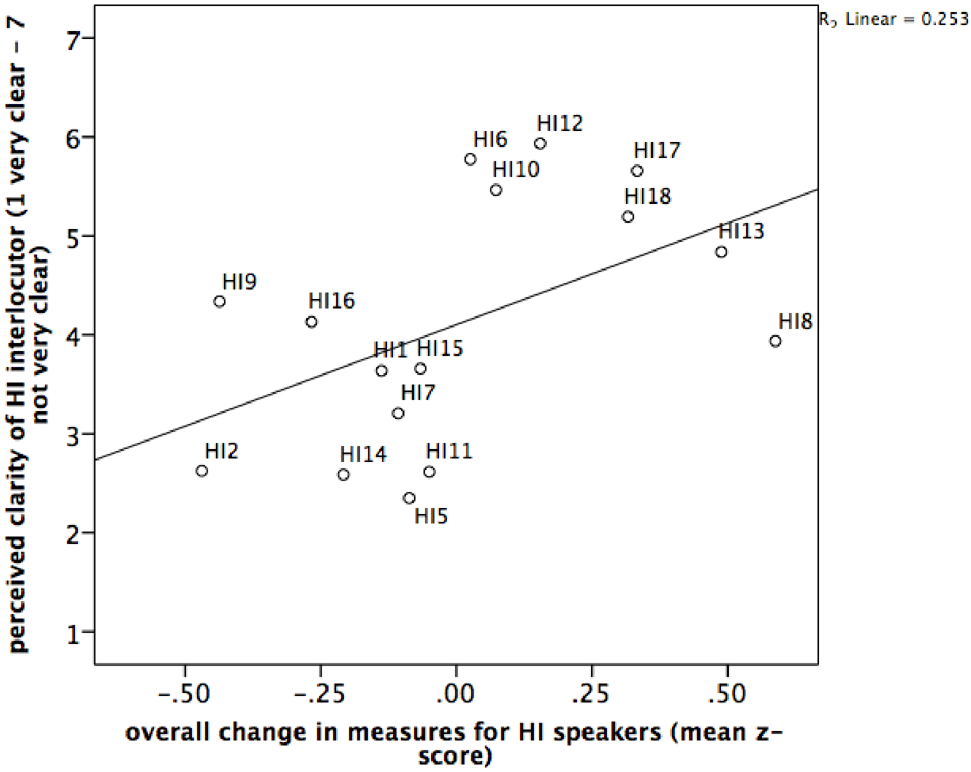



Figure 10.2: Scatterplot of the mean perceived clarity rating of the HI listener and the mean z-score of the overall change in measures from the NH-directed to the HI-directed condition for the HI speakers. The circles represent HI speakers.

When examining the different types of enhancement strategies, the extent of global and linguistic enhancements made by the HI speakers were not significantly correlated with HI interlocutor's speech clarity – however, the extent of segmental enhancements was just-significantly correlated with interlocutor speech clarity, with greater segmental enhancements being made to interlocutors with lower perceived speech levels (see Table 10.2).

[1]HI3 was excluded from this analysis due to his HI interlocutor's (HI4) speech clarity score being unavailable.

| data | r | p | $R^2$ |
|---|---|---|---|
| global-HI mean clarity | 0.366 | 0.163, n.s. | NA |
| segmental-HI mean clarity | 0.487 | 0.054(*) | 0.239 |
| linguistic-HI mean clarity | -0.23 | 0.392, n.s. | NA |

Table 10.2: Correlations between global, segmental and linguistic enhancement z-scores and HI interlocutor mean clarity rating in the NH-directed condition by HI speakers (n=16).

### 10.3.2 Speaker perceived speech level

We also explored whether the perceived speech level of the HI speaker predicts the extent to which adaptations are made to listener needs. We would expect that those HI participants with lower perceived speech levels would be less able to make adaptations to their listener's needs.

However, there was no significant correlation between the perceived clarity rating of the HI speaker and the amount of overall extent of adaptations made by that speaker (measured in the composite z-score) (n=17; r=-0.187, p=0.471, n.s.), indicating that a HI speaker's perceived speech level was not related to the extent of the adaptations made to a HI interlocutor. When examining each adaptation type separately, the HI speakers' speech clarity was also not correlated with the change in adaptation (global adaptations: r=-0.288, p=0.263, n.s.; segmental adaptations: r=-0.0935, p=0.721, n.s.; linguistic adaptations: r=-0.105, p=0.689, n.s.).

### 10.3.3 Relationship between strategies

So far, then, we have seen that listener-related factors, but not speaker-related factors, seem to explain part of the variability in adaptation strategies between speakers in this study. Another reason that may explain the variability in strategy use between speakers is that speakers are using complementary strategies to adapt to their HI interlocutor. For example, some speakers may use only acoustic-phonetic adaptations, but few linguistic adaptations, while others may favour linguistic adaptations over acoustic-phonetic ones when talking to their HI peer.

To explore this possibility, Pearson's correlations were run on the mean z-scores of overall change in global, segmental and linguistic measures for both NH and HI participants separately. The results for NH speakers are displayed in Table 10.3. The table shows that the extent of adaptations made in global and segmental measures, and in global and linguistic measures, were not correlated with each other. However, there was a significant negative correlation between the amount of segmental and linguistic change – the more a NH speaker made linguistic changes to their speech, the fewer segmental changes were made (see Figure 10.3). This finding implies that segmental and linguistic adaptations may be used by NH speakers as alternative strategies in adaptation. The figure also shows that a small subset of speakers (mainly NH8, NH9 and NH11) use segmental contrast enhancement as a strategy in HI-directed speech, despite the findings of chapter 7 suggesting that on the group level, NH speakers do not make these enhancements when talking with a HI peer.

| data | r | p | $R^2$ |
|---|---|---|---|
| global-segmental | 0.311 | 0.224, n.s. | NA |
| global-linguistic | -0.001 | 0.996, n.s. | NA |
| segmental-linguistic | -0.662 | <0.01** | 0.438 |

Table 10.3: Correlations between global, segmental and linguistic change z-scores for NH speakers (n=17).

As displayed in Table 10.4, the extent of adaptations made in global, segmental and linguistic measures by HI speakers were however not correlated with each other, suggesting that HI speakers' variability in strategy use is not explained by their use of global, segmental and linguistic adaptations as alternative strategies in this study.

| data | r | p | $R^2$ |
|---|---|---|---|
| global-segmental | 0.208 | 0.423, n.s. | NA |
| global-linguistic | -0.217 | 0.404, n.s. | NA |
| segmental-linguistic | -0.253 | 0.328, n.s. | NA |

Table 10.4: Correlations between global, segmental and linguistic change z-scores for HI speakers (n=17).

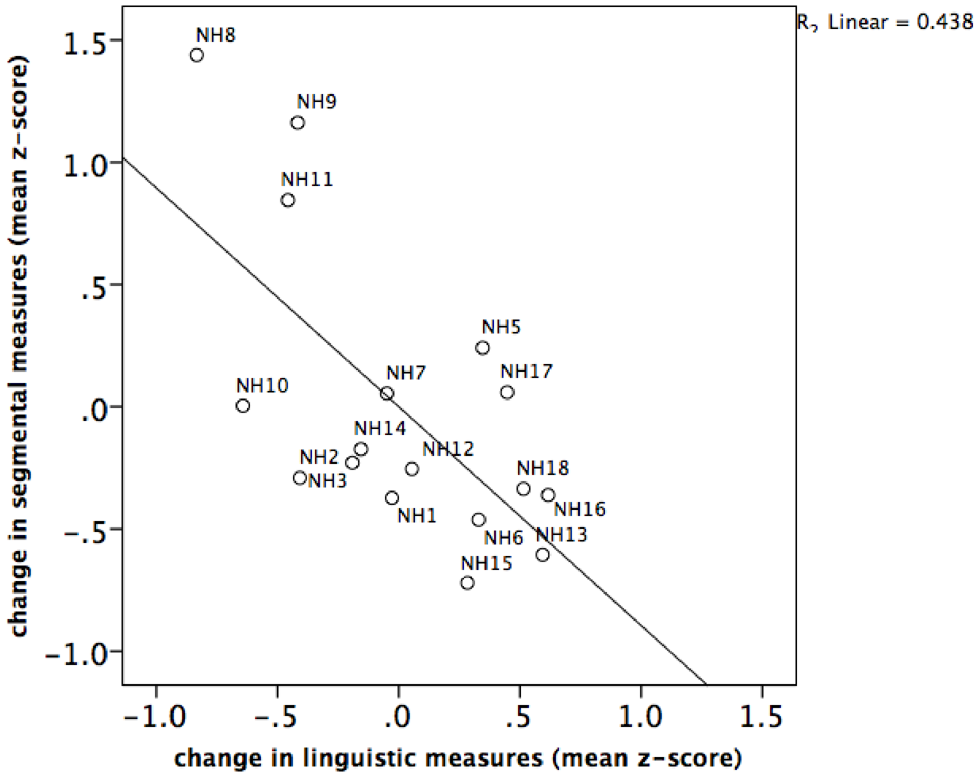


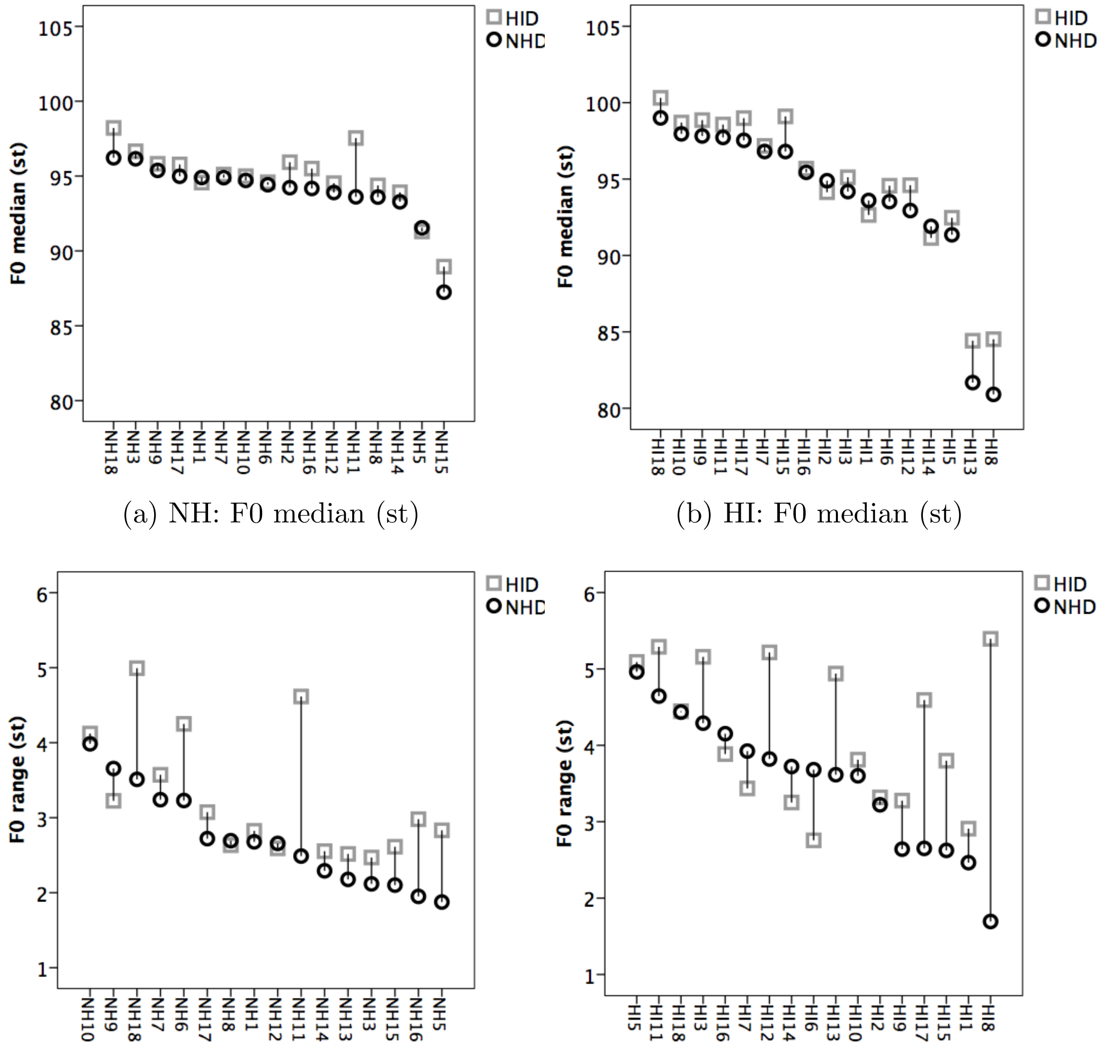

Figure 10.3: Scatterplot of the mean change in segmental and linguistic measures (in z-scores) for the NH speakers.

### 10.3.4 Individual variability in strategies

**Variability in strategy use** To graphically demonstrate the variability between participants' adaptation strategies, the mean values per speaker per condition over both tasks[1] were calculated for each global, segmental and linguistic measure which previous analyses in chapters 6, 7 and 8 showed speakers to use as a significant strategy in HI-directed speech. Figures 10.4, 10.5, 10.6 and 10.7 display the means per speaker in both HI- and NH-directed conditions for those measures.

As implied by the sections above, it is evident that there is a great deal of variability in most of the measures as to the extent to which the speakers enhance

[1]With the exception of lexical diversity, for which only the Grid task measures were used, as the results from chapter 8 demonstrated that speakers only increased their lexical diversity in HI-directed speech compared to NH-directed speech in the Grid task, but not in the Diapix task.



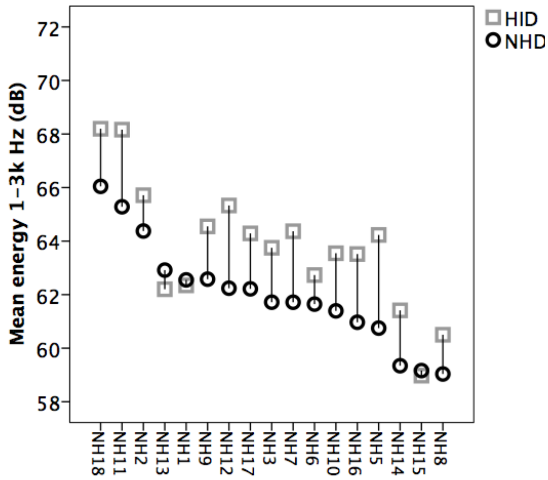

(a) NH: F0 median (st)

(b) HI: F0 median (st)

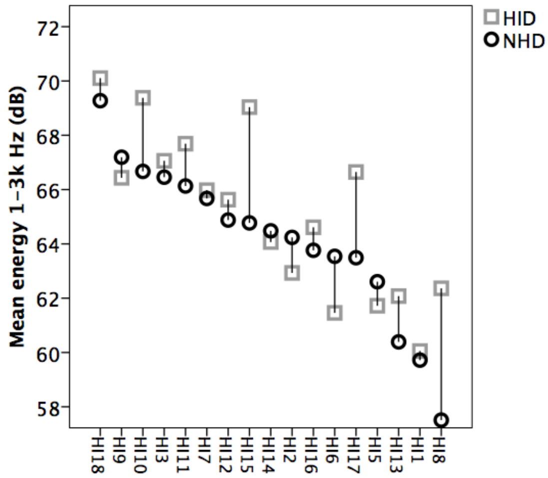
(c) NH: F0 range (st)

(d) HI: F0 range (st)

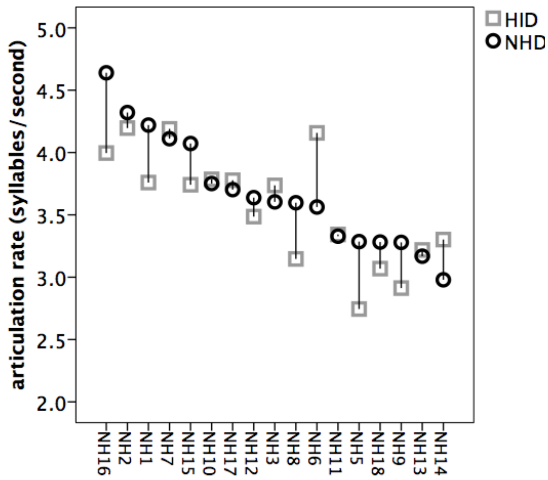
Figure 10.4: Global adaptations 1: F0 median (in semitones) and F0 range (in semitones) for NH and HI speakers in the HI-directed (grey square) and NH-directed (black circle) conditions. Speakers are ordered according to values in the NH-directed condition.



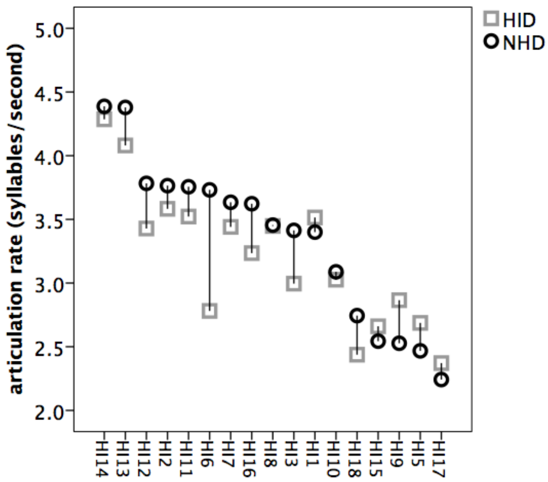

(a) NH: Mean energy 1-3k Hz (dB)




(b) HI: Mean energy 1-3k Hz (dB)

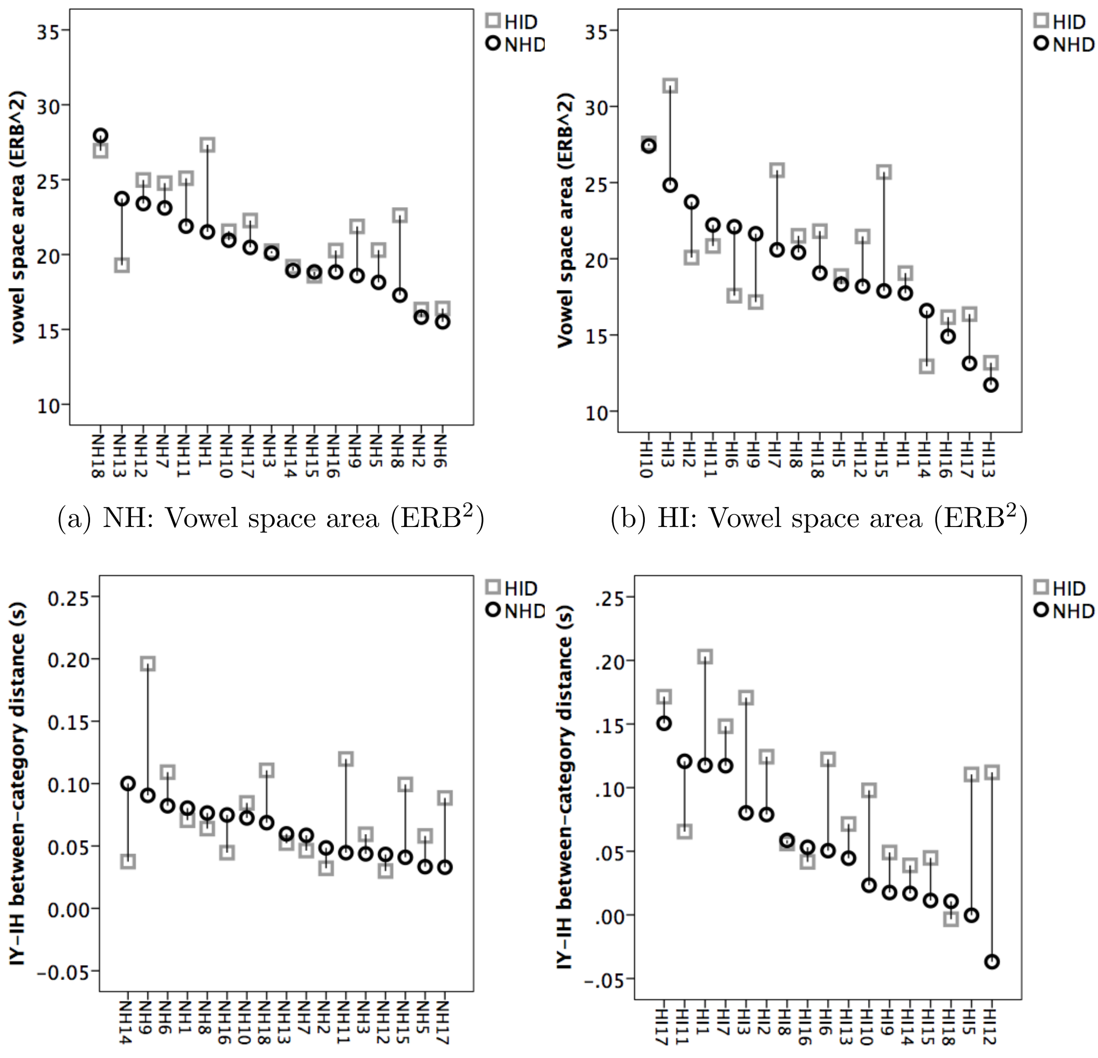



(c) NH: articulation rate (syll/s)




(d) HI: articulation rate (syll/s)

Figure 10.5: Global adaptations 2: Intensity (mean energy 1-3k Hz, in dB) and articulation rate (in syllables per second) for NH and HI speakers in the HI-directed (grey square) and NH-directed (black circle) conditions. Speakers are ordered according to values in the NH-directed condition.



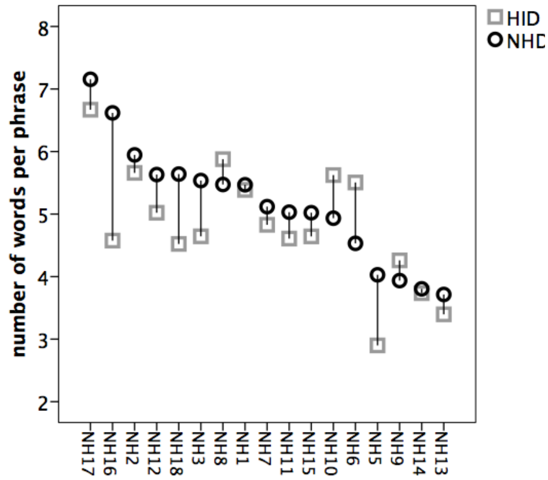

(a) NH: Vowel space area (ERB$^2$)

(b) HI: Vowel space area (ERB$^2$)

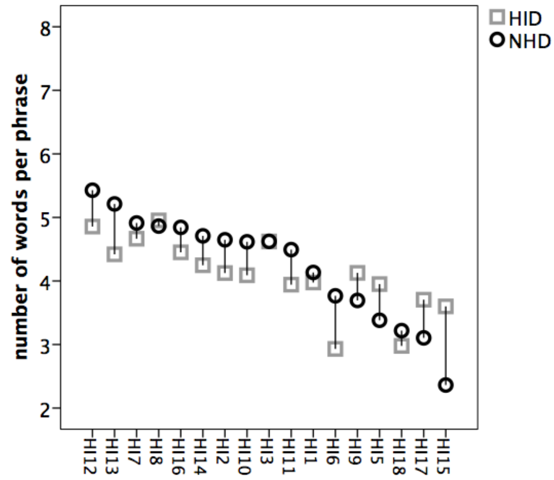
(c) NH: /i/-/ɪ/ temporal distance (s)

(d) HI: /i/-/ɪ/ temporal distance (s)

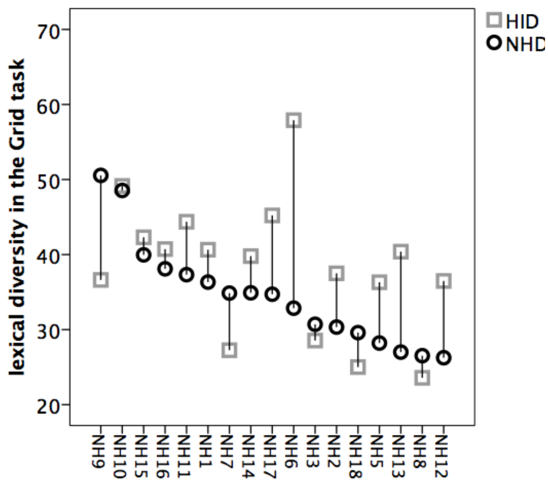
Figure 10.6: Segmental adaptations: Vowel space area (in ERB$^2$) and between-category distance for the /i/-/ɪ/ temporal between-category distance (in seconds) for NH and HI speakers in the HI-directed (grey square) and NH-directed (black circle) conditions. Speakers are ordered according to values in the NH-directed condition.



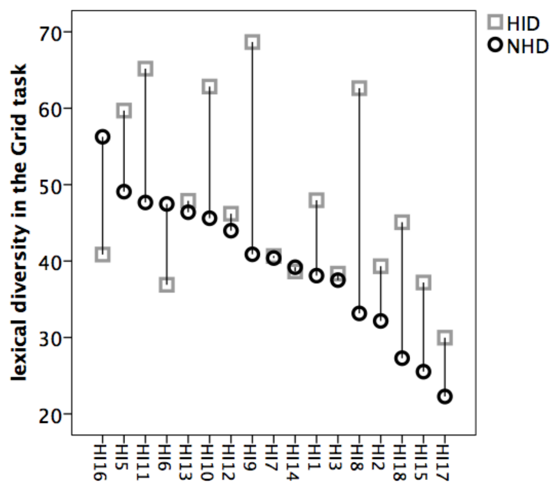

(a) NH: words per phrase




(b) HI: words per phrase




(c) NH: lexical diversity




(d) HI: lexical diversity

Figure 10.7: Linguistic adaptations: mean number of words per phrase, and lexical diversity in the Grid task, for NH and HI speakers in the HI-directed (grey square) and NH-directed (black circle) conditions. Speakers are ordered according to values in the NH-directed condition.

them. The figures also demonstrate that even in the NH-directed condition, the speakers vary greatly in their values for each measure. For most of these measures, however, the speakers do not seem to be affected by their NH-directed 'baseline' measure in the extent to which they can use that measure to adapt – for example, for the /i/-/ɪ/ temporal distance measure, some speakers appear to increase the temporal distance between the vowels regardless of their location on the group distribution in the NH-directed condition (Figure 10.6), i.e., both those with greater and less temporal distance between their /i/-/ɪ/ categories in the NH-directed condition are nonetheless able to change this measure in HI-directed speech. However, two of these measures, articulation rate (Figures 10.5c and 10.5d) and number of words per phrase (Figures 10.7a and 10.7b), may be slight exceptions to this – it seems that speakers who had the slowest articulation rate or who used the fewest words per phrase per group did not enhance their speech on these parameters – for example, for the number of words per phrase measure, both NH and HI speakers who used approximately four or fewer words per phrase in the NH-directed condition did not seem to decrease the number of words used per phrase further in HI-directed speech. Thus the scope of some speakers to modify their speech in the expected direction may be somewhat limited by their NH-directed values for these two measures.

**Speaker-listener effects** As seen above in section 10.3.1, the strategy use by speakers was significantly related to the characteristics of their listener. Thus, to explore speaker-listener factors in adaptation in more individual detail, the HI participants were split into two groups based on their level of hearing loss – those with severe-to-profound and profound hearing loss (8 participants), and the remaining HI participants with milder hearing loss (10 participants). Table 10.5 displays the two groups of HI participants along with their NH and HI friends, and the strategies used by their NH and HI friends when speaking to them. Here, a strategy is termed as being used if the speaker used that strategy to a relatively extreme extent – if they were in the upper 38% of speakers (mean z-score of 0.3 or above) in their speaker group using that strategy.

In general, speakers talking to HI participants with severe-to-profound or profound hearing loss used somewhat more extreme strategies than did those

| listener | list.id. | NH friend | NH's strategy | HI friend | HI's strategy |
|---|---|---|---|---|---|
| HI3$^{CI}$ | 75% | NH3 | none | HI4$^{S}$ | – |
| HI5$^{CI}$ | 50% | NH5 | G,L | HI6$^{M}$ | none |
| HI8$^{CI}$ | 75% | NH8 | S | HI7$^{MS}$ | none |
| HI9$^{CI}$ | 42% | NH9 | S | HI10$^{CI}$ | L |
| HI10$^{CI}$ | 83% | NH10 | none | HI9$^{CI}$ | none |
| HI11$^{P}$ | 58% | NH11 | G,S | HI12$^{S}$ | G |
| HI14$^{CI}$ | 67% | NH14 | none | HI13$^{M}$ | G,S |
| HI18$^{CI}$ | 67% | NH18 | G,L | HI17$^{M}$ | G,S |

| listener | list.id. | NH friend | NH's strategy | HI friend | HI's strategy |
|---|---|---|---|---|---|
| HI1$^{MS}$ | 92% | NH1 | none | HI2$^{MS}$ | none |
| HI2$^{MS}$ | 83% | NH2 | none | HI1$^{MS}$ | L |
| HI4$^{S}$ | 92% | NH4 | – | HI3$^{CI}$ | none |
| HI6$^{M}$ | 92% | NH6 | L | HI5$^{CI}$ | S |
| HI7$^{MS}$ | 92% | NH7 | none | HI8$^{CI}$ | G |
| HI12$^{S}$ | 75% | NH12 | none | HI11$^{CI}$ | none |
| HI13$^{M}$ | 100% | NH13 | L | HI14$^{CI}$ | none |
| HI15$^{S}$ | 83% | NH15 | none | HI16$^{M}$ | none |
| HI16$^{M}$ | 92% | NH16 | G,L | HI15$^{S}$ | G |
| HI17$^{M}$ | 92% | NH17 | L | HI18$^{CI}$ | S |

Table 10.5: The 8 HI participants with severe-to-profound and profound hearing loss (upper table), and the 10 HI participants with moderate, moderate-to-severe, and severe hearing loss (lower table). Next to each HI participant is their NH and HI friend's participant numbers, and the main adaptation strategies used by those friends in HI-directed speech (i.e., the adaptation type/s in which they are in the upper 38% for their speaker group (z-score over 0.3), if applicable) (G=global, S=segmental, L=linguistic, – =no data). The HI listener's mean /p/-/b/ and /s/-/ʃ/ identification score (list.id., in % correct) is displayed in the second column. The superscript next to each HI participant's name denotes the level of their hearing loss (for those without CIs), or their CI status (M=moderate, MS=moderate-to-severe, S=severe, P=profound, CI=wears CIs).

speakers interacting with HI participants with less severe hearing loss levels. In particular, it seems that NH speakers mainly used extreme linguistic adaptation strategies for those with milder hearing loss levels, and more extreme global, segmental and linguistic strategies to those with more profound hearing losses.

As can be seen from the table, all the NH speakers who use the segmental contrast enhancement strategy to a great extent (NH8, NH9 and NH11) (as also seen in Figure 10.3) interact with a friend with severe-to-profound or profound hearing loss in the HI-directed condition. Likewise, two of the HI participants ($HI13^{M}$, $HI17^{M}$) who use segmental contrast enhancement to a greater extent compared to other HI participants, interact with a HI friend with profound hearing loss, while themselves having only moderate hearing loss. The segmental enhancement strategy therefore seems to be mostly used only by a subset of speakers (see Table Table 10.6) interacting with HI participants with the most profound hearing loss levels.

Indeed, the HI participants with profound losses, to whom segmental contrast enhancement is used as a strategy, obtained a mean 62% score for /s/-/ʃ/ and /p/-/b/ identification in the VCV (AV) perception task (see chapter 4 and Table 10.5), compared to the mean of the remaining HI speakers' identification score of 85% – suggesting that these listeners had greater difficulty than the other HI listeners in perceiving the segmental contrasts in the Grid task, which may have lead some of their interlocutors to use segmental contrast enhancement strategies in HI-directed speech.

| speaker | /p/-/b/ | /s/-/ʃ/ | /i/-/ɪ/ |
|---|---|---|---|
| NH8 | 134% | 94% | -15% |
| NH9 | 7% | 144% | 94% |
| NH11 | -34% | 662% | -16% |
| $HI5^{CI}$ | -26% | 8% | 201% |
| $HI13^{M}$ | 37% | 159% | 110% |
| $HI17^{M}$ | 20% | 144% | 43% |
| $HI18^{CI}$ | 41% | 76% | 167% |

Table 10.6: The % increase in segmental contrast discriminability for the 7 speakers who were found to increase segmental contrasts in HI-directed speech compared to NH-directed speech.

Although $HI5^{CI}$ also obtained low scores on the VCV (AV) test, her interlocutors did not use segmental contrast enhancement to a great extent but her NH interlocutor, NH5, used greater global and linguistic strategies instead. Similarly, it is not the case that all those speakers who interacted with the five HI listeners who had greater difficulty in phoneme perception enhanced segmental contrasts in HI-directed speech – for example, $HI7^{MS}$ and NH14 did not use any extreme strategies when talking with their profoundly hearing-impaired friend compared to their NH friend, whereas $HI10^{CI}$, NH18 and $HI12^{S}$ used global and/or linguistic enhancement strategies.

Interestingly, the two other HI participants who used segmental contrast enhancement as a strategy, $HI5^{CI}$ and $HI18^{CI}$, interact with HI friends with only moderate hearing loss – however, they themselves have profound hearing loss, and perform among the worst in the speech perception tests reported in chapter 4 (see Figures 4.1 and 4.3). It seems likely, therefore, that their segmental contrast enhancement is linked to their own reduced speech intelligibility – they are attempting to enhance the contrasts to their HI friends due to having trouble producing the contrasts themselves.

## 10.4 Discussion

The findings of this chapter indicate that, as previously implied by instructed clear speech studies (e.g., Ferguson and Kewley-Port, 2007), a great deal of variability exists between speakers in the strategies they use when talking with a HI peer. One potential factor affecting between-speaker variability in the strategies used was proposed to be the specific characteristics of the listener. Indeed, the results of this study indicated that both NH and HI participants who interacted with a HI peer with lower perceived speech levels made greater overall enhancements to their speech than did NH and HI participants who were talking to a HI friend with higher perceived speech skills. It is particularly interesting that NH and HI children appear to be able to make specific adaptations to listener needs, as it has been suggested that older children are not necessarily very sensitive to the specific needs of a listener (Hazan *et al.*, submitted). This implies that both NH and HI children may be using similar monitoring of their listener as adults do (e.g., Lindblom, 1990; Schober, 1993) – alternatively, as each speaker interacted with only one HI interlocutor, and the interlocutors were friends and had experience in interacting with each other, speakers may have learned to use a greater number of adaptation strategies when talking with a particular HI friend.

The finding that segmental contrast enhancement strategies were used only by a subset of NH and HI speakers who interacted with a HI interlocutor with poor phoneme identification skills implies that, contrary to conclusions of chapter 7, children can and do adapt on the segmental level. However, not all speakers who interacted with those HI interlocutors with poor phoneme identification skills made adaptations on the segmental level, and therefore segmental contrast enhancement seems to be an optional strategy, which is perhaps used only in circumstances where it is very much required. This result is similar to that found in more controlled computer-directed production tasks (Maniwa *et al.*, 2009; Schertz, 2013), in which speakers only enhanced minimal pair distinctions after being provided with specific misrecognition feedback of a target phoneme. Thus phoneme contrast enhancement may be restricted to specific local misrecognition errors. Nonetheless, it is surprising that at least a few NH and HI children are able to make these enhancements. Related to this finding, it is possible that a group-level effect of segmental contrast enhancement would have been found in chapter 7 had more of the HI participants been more profoundly hearing-impaired.

Additionally, NH speakers were found to use segmental and linguistic modifications as alternative strategies when talking with their HI friend – speakers who modified the segmental aspects of their speech made few linguistic changes, while those making greater linguistic adaptations rarely modified the segmental level. This may partly explain why only few speakers used segmental contrast enhancement in HI-directed speech – as an alternative to enhancing contrasts for the minimal pair keywords in the Grid task, speakers may have selected an alternative word or explained their keyword differently to their interlocutor. These results reflect the importance of using a spontaneous interactive speech task to elicit natural speech adaptations, as clear speech studies using read sentence materials are unable to assess the extent to which other, for example linguistic, strategies are used in the adaptation to listener needs.

Although it was predicted that part of the between-speaker variability in strategy use could be accounted for by the speaker's perceived speech level, no such effect was found. Although the measure of perceived clarity may be a somewhat crude measure of HI participants' speech and language level, it is nonetheless surprising that those HI participants with greater difficulty in speech and language are able to make adaptations to their interlocutor similarly to peers with higher language levels. In particular, two HI participants, HI5 and HI18 even used segmental contrast enhancement strategies in HI-directed speech, despite having low speech and language skills, as measured both by the clarity rating study in chapter 9 and by the perception tests in chapter 4. These findings therefore support the conclusions of previous chapters in indicating that lower speech and language skills do not prevent speakers from modifying their speech as needed, even on the fine phonetic level – similarly to results from clear speech studies (Goberman and Elmer, 2005; Tjaden *et al.*, 2014) showing that speakers with dysarthria are nonetheless able to make acoustic-phonetic modifications to their speech when instructed to do so.

Overall, the results of this chapter therefore suggest that at least some of the variability in strategy use by speakers talking to HI peers can be explained by the listener's perceived speech level, but not by the speaker's speech level. They also demonstrate the importance of examining the adaptations made on an individual level, as a subset of NH and HI speakers were found to make segmental adaptations which were not made by the speakers as a group.

# Chapter 11

# Discussion

## 11.1 General Discussion

An essential aspect of effective speech communication is the ability to make speech and language adaptations to the needs of a listener. Previous research has mainly concentrated on assessing the speech adaptations made by adults when talking to an interlocutor in adverse conditions (e.g., Cooke and Lu, 2010; Hazan and Baker, 2011), or by preschool-aged children when interacting with infants or puppets (Masur, 1978; Syrett and Kawahara, 2014; Warren-Leubecker and Bohannon, 1983). Due to their difficulties in receptive speech and language, HI children are also likely to require acoustic-phonetic and linguistic modifications from their interlocutor. However, no known previous studies have examined the abilities of older NH and HI children in adapting to their HI listener's needs, even though a large proportion of HI children attend mainstream schools in the UK (CRIDE, 2014), and children therefore face frequent interactions with both NH and HI peers. Both NH and HI older children's speech production and higher linguistic, pragmatic and cognitive skills are still developing, and therefore they may not have the ability to spontaneously adapt to the needs of a HI listener.

The aim of this thesis was therefore to examine the speech communication strategies used by both NH and HI 9- to 15-year-old children in their talk to a HI peer. Unlike in most previous studies, in which a hyperarticulated speaking style is elicited by instructing speakers to 'speak as if to a hearing-impaired person' compared to ‘casually as if to a friend’ (e.g., Ferguson and Kewley-Port, 2007; Smiljanić and Bradlow, 2005), speech was elicited using a more ecologically valid method – with real interlocutors, interacting face-to-face on two problem-solving referential communication tasks similar to those used in an everyday school setting. As well as using the more difficult Diapix task (Baker and Hazan, 2011) in speech elicitation, the Grid task was successfully devised to enable the analysis of both between- and within-category segmental contrast enhancement as well as a wide range of other acoustic-phonetic and linguistic measures from the participants’ speech (see chapters 2 and 5). As described in chapter 3, children were not given instructions on how to communicate with each other on these tasks, but interacted spontaneously, once with a HI friend and once with a NH friend, to elicit both NH- and HI-directed speech from each interlocutor.

In chapter 6, it was found that HI children’s speech characteristics differed from NH peers’ in having a greater F0 range, greater speech intensity in the 1 to 3k Hz range, and a slower articulation rate. The following chapter, chapter 7, demonstrated that, despite having similar vowel space areas and similarly discriminable /i/-/ɪ/ temporal contrasts and /p/-/b/ voicing categories, HI children differed from their NH peers in exhibiting a smaller vowel F2 range, and less discriminable spectral /i/-/ɪ/ and /s/-/ʃ/ contrasts. Additionally, chapter 8 suggested that HI children used shorter phrases and lexically more frequent content words in both tasks compared to NH peers. They also used more diverse vocabulary in the Grid task – reflecting perhaps their frequent need to revise utterances following miscommunications – and showed greater speech overlap in the Diapix task, than did NH peers. Additionally, a clarity rating study conducted on adult listeners in chapter 9 showed that HI children’s speech was perceived as being less clear than NH children’s. These results are generally in agreement with findings of previous studies on HI children’s speech characteristics (e.g., Burkholder and Pisoni, 2003; Elfenbein *et al.*, 1994; Kosky and Boothroyd, 2003), based mostly on read speech, in suggesting that HI children exhibit reduced speech motor control, slower speech production planning, difficulty in producing distinctions between fine spectral contrasts, delayed expressive syntax and vocabulary skills, and perhaps more controlling conversational behaviour, compared to their NH peers.

Likely due to these difficulties in speech production, and the speech perception deficits shown by HI interlocutors in chapter 4, HI-HI and HI-NH dyads were found to take longer to complete both the Grid and Diapix tasks than NH-NH dyads, implying an increase in communication difficulty in conditions involving HI participants (chapter 5). This suggested that, in those conditions, HI children's interlocutors needed to make greater effort and modify their speech and language to enable successful communication.

Chapters 6, 7 and 8 also indicated that, despite the findings of delayed speech and language in the HI children compared to their NH peers, both speaker groups used similar acoustic-phonetic and linguistic strategies in adapting to their HI interlocutor, at least for the measures analysed here. In chapter 6, both NH and HI children were found to increase their F0 median, F0 range and speech intensity in the 1 to 3k Hz range, and slightly decrease their articulation rate, in HI-directed speech compared to NH-directed speech. These findings closely reflect instructions given to HI persons' communication partners, which often ask them not to shout, but to use emphasis and intonation, and a slower speech rate, when talking with a HI person (Action on Hearing Loss, 2012; Caissie and Tranquilla, 2010; Marschark and Hauser, 2012). The global adaptations made to a HI interlocutor were qualitatively similar to those found in instructed clear speech studies by adult speakers (e.g., Picheny *et al.*, 1986; Smiljanić and Bradlow, 2005), but differed somewhat in the extent to which the strategies were applied – likely due to the communicative nature of the current study.

The /p/-/b/, /s/-/ʃ/ and /i/-/ɪ/ phonetic contrasts played a specific communicative role in the Grid task, as listeners had to perceive and produce minimal pair keywords accurately to be able to complete the task – crucially, these segmental contrasts are also typically found to be difficult for HI children to produce and perceive (e.g., Giezen *et al.*, 2010; Mildner *et al.*, 2006). Thus a speaker's increased phonetic category discriminability in the HI-directed condition would have been very useful for completing the task successfully. Indeed, HI children's communication partners may be instructed to "enunciate carefully" (Caissie and Tranquilla, 2010, p.100) to a HI person – and some previous studies on instructed clear speech show speakers to make segmental enhancements in clear speech (e.g., Ferguson and Kewley-Port, 2002; Smiljanić and Bradlow, 2008a). However, most of the NH and HI children were not found to use segmental contrast enhancement to a great extent in HI-directed speech, with the exception of a slight increase in vowel space area and F1 range, and an increase in the durational between-category distance of /i/-/ɪ/ in HI-directed speech compared to NH-directed speech. The spectral /i/-/ɪ/ categories were even found to become less contrastive in HI-directed speech, leading to a suggestion that vowel formant changes in HI-directed speech may be a side-effect of increased speech effort in that condition.

The use of shorter and less complex utterances, rephrasing rather than repetition of a message, and increased sentence contextual information are also encouraged when talking with HI persons (Action on Hearing Loss, 2012; Marschark and Hauser, 2012; NDCS, 2012b). Few studies have investigated these aspects of adaptation, beyond measuring utterance length or message repetition, or using objective measures (e.g., Masur, 1978). On the linguistic level, chapter 8 found that both NH and HI speakers used slightly shorter phrases in both tasks, and increased lexical diversity in the Grid task, when talking with their HI friend – implying that syntactic simplification and message revision were used as a communicative strategy in HI-directed conditions. However, the lexical frequency of the children's content words did not differ between NH- and HI-directed speech – perhaps due to limitations imposed by the tasks on the vocabulary used, or due to the high level processing requirements for deciding on and changing the difficulty of the vocabulary being used by a speaker. The speakers also did not modify the amount of speech overlap used when talking with a HI friend, perhaps due to requiring advanced pragmatic skills of the speakers.

The findings of the final chapter, chapter 10, indicated that some of the variability in strategies found in the current study could be explained by listener-related, but not speaker-related, factors. Both NH and HI speakers who were paired with HI interlocutors with lower perceived speech levels used speech and language enhancement strategies to their HI peer to a greater extent than did speakers interacting with HI peers with higher perceived speech levels – implying that both NH and HI speakers are fairly sensitive to the specific speech and language needs of their listener, contrary to suggestions in Hazan *et al.* (submitted). Additionally, a subset of speakers were found to make enhancements to segmental contrasts in HI-directed speech, despite findings of chapter 7 suggesting that no such enhancements were made on the group level. Some of the speakers who interacted with a HI interlocutor who had specific speech perception difficulties with the contrasts used in the Grid task did enhance those segmental contrasts in HI-directed speech – but even so, not all speakers used segmental contrast enhancement to clarify their speech to these interlocutors. At least NH speakers may have used linguistic enhancement strategies as an alternative to segmental contrast enhancement – those speakers who made their phonetic contrasts more discriminable in HI-directed speech did not enhance linguistic aspects of their speech, while those with fewer segmental contrast enhancements used linguistic enhancements to a greater extent.

On the other hand, the ability of HI speakers to adapt to a HI friend did not depend on their own perceived speech level – even those HI children with lower perceived speech levels and below-average speech perception skills were nonetheless able to make modifications to their speech and language when talking with their HI friend, suggesting that speakers' low speech and language skills do not preclude them from making adaptations to a listener. This is surprising, as the literature reviewed in chapter 1 suggested that a speaker needs to have developed fairly sophisticated skills to be able to adapt to their listener.

In summary, the main finding of this thesis is that despite older NH children's continued speech and language development, and despite the delays exhibited by older HI children in their speech and language skills, both speaker groups made global acoustic-phonetic and linguistic adaptations to their speech when talking with a HI peer compared to a NH peer – and a subset of speakers also enhanced the segmental level of their speech. These adaptations did not appear to depend on the HI speakers' speech and language level. Moreover, speakers seemed to be sensitive to the specific needs of their listener, with both groups of speakers making greater adaptations to HI interlocutors whose perceived speech level was lower than to those with higher perceived speech levels.

## 11.2 Theoretical considerations

**Skills needed to adapt** These results bring to light two important theoretical points. Firstly, it is unclear which skills are required to enable a speaker to adapt to their listener's needs, and how these skills develop. Studies on adults have shown that speakers are able to make adaptations to their speech based on the listener's listening environment (e.g., Aubanel and Cooke, 2013), listener feedback (e.g., Smith and Trainor, 2008; Warren-Leubecker and Bohannon, 1982) and a listener's visible characteristics (e.g., Sims and Chin, 2002), and models propose that speakers both adapt according to the situational needs of the listener (Brown and Dell, 1987; Lindblom, 1990), at least if resources allow (Bard and Aylett, 2005), as well as according to the speaker's initial assessment of listener needs (Fussell and Krauss, 1992; Schober, 1993). However, to our knowledge, none of the theories mentioned in section 1.3.2 have proposed a developmental theory of their model.

In section 1.3.3, it was hypothesised that several different developmental changes must occur for children to be able to take a listener's needs into account in communication (Schmidt and Paris, 1984) – children would need to develop flexible speech and language systems, as well as adequate skills in perspective-taking and in inferring listener needs according to listener feedback. Although 9- to 15-year-old NH children are not likely to have yet developed adult-like speech production control, it seems probable that they possess most of the basic skills required for listener adaptation. However, the HI children in this study are known to have problems in speech motor control, speech production planning and delayed language and, as a possible consequence of missing out on social situations and incidental learning due to their hearing loss (Löfqvist *et al.*, 2010), may also exhibit delays in pragmatic skills, such as perspective-taking and detecting listener feedback (see section 1.3.4). Nonetheless, even with these deficits in skills that were hypothesised to be vital for being able to adapt to the listener, HI children were found to make similar adaptations to their HI interlocutor as NH peers, and to even be sensitive to the amount of speech and language adaptation required for the language level of their particular interlocutor.

As the speech adaptation measures in the current study were taken from the participants' entire interaction, we do not know whether the adaptations made were done dynamically according to listener feedback, or based on experience or initial assessment. However, it is possible that the HI children in this study have learned to make these adaptations to their listener due to their own experiences in having, perhaps regular, communication breakdowns (Tye-Murray, 2003) – and thus learning from their communication partners which strategies may be most helpful when interacting with a person with hearing loss. Indeed, findings by Ciocci and Baran (1998), which show that 4- to 8-year-old HI children made more adaptive repair responses to their interlocutor than did age-matched NH peers, counts towards this view. Similarly, HI children may not always need to model another HI listener to as great an extent as NH children do, if they use their own experience of having hearing loss as a guide to infer the listener's needs – although this would not explain how the HI children seemed to be able to adjust the extent of adaptations made according to the perceived speech level of their HI listener. Alternatively, it may be that the HI children in this study, who were friends with their HI interlocutor, and had known each other on average for several years, had, through trial and error, developed specific strategies that they used with their particular friend. Thus their adaptation strategies may not reflect listener modelling per se, as they may not generalise to even other HI listeners, let alone other types of situations in which listener needs have to be taken into account. For example, Lederberg *et al.* (1986) report on findings of 5- to 6-year-old NH children who used visual communication strategies with a familiar HI friend, but were not able to generalise these strategies to an unfamiliar HI peer.

The above explanations may account for the HI children's abilities in modelling and inferring the needs of their HI friend, despite HI children's likely delays in language and pragmatic skills. However, they do not explain the reason for many of the HI children's abilities in using global acoustic-phonetic enhancements in speech to a HI peer despite their likely deficits in speech motor control – as seen in chapter 10, the perceived speech level of the HI children, which is likely to be strongly linked to poorer speech motor control, did not prevent HI children from making adaptations – similarly to Syrett and Kawahara's (2014) findings of global acoustic-phonetic speech adaptations being made even by preschool-aged children. As suggested in chapter 6, it may be that some types of adaptations, especially global acoustic-phonetic ones, are linked to more automatic processes associated with speech effort rather than speech motor control. Alternatively, speech adaptations may be linked to interactive social processes within the communicative situation, such as suggested in the Interactive alignment model of speech communication (Pickering and Garrod, 2004, 2013). According to the model, in dialogue, interlocutors automatically adjust their speech and language to become more similar ('aligned') to their interlocutor in an interaction, to benefit production and comprehension processes in both interlocutors. This is done on several different levels of communication in an interaction, thus making it easier for interlocutors to be understood by each other – and requiring little listener modelling from each interlocutor. Such a model could potentially account for both preschool and language-delayed children's abilities in listener adaptation, but is still lacking in giving a developmental account of the processes involved (e.g., Aitken, 2013; Krishnan, 2013).

**Levels of adaptation** Secondly, the results of the current study imply that models of speech adaptation need to take into account the different acoustic, phonetic and linguistic levels on which speakers can adapt to their interlocutor in spontaneous interaction. As discussed in section 1.3.2, the H&H theory (Lindblom, 1990) was primarily postulated to account for phonetic variation. According to the theory, the listener's comprehension of the message is a priority to the speaker and therefore, despite the increased speaking effort needed, speakers endeavour to approximate phonetic targets whenever communicative demands so require. Thus, although Lindblom's (1990) theory seemingly accounts for listener adaptation in interaction, it does not take into account the many alternative strategies available to speakers in a real communicative situation. Therefore it is unable to account for the finding of this study suggesting that NH children used segmental and linguistic enhancements as alternative strategies – and the fact that, despite the importance of accurate phoneme perception in the Grid task, both NH and HI children only rarely used segmental contrast enhancement when speaking to a HI friend.

On the other hand, the Dual Process approach (Bard *et al.*, 2000; Bard and Aylett, 2005) postulates that speakers do not adjust the phonetic level of speech to a listener's needs due to the high cognitive demands required in consulting a listener model for each produced phoneme. Instead, it hypothesises that speakers can only use higher linguistic levels in taking a listener's needs into account – and even then, adaptations are only made if cognitive resources allow. The current findings are therefore somewhat in disagreement with this model in suggesting that speakers can and do increase the discriminability of phonetic contrasts – although evidently only in situations in which such an enhancement is crucial for listener understanding. Similarly, although according to the Dual Process approach, speakers would make fewer adaptations to their listener when task demands are high, few effects of task difficulty were found in the current study. Although it is possible that the Diapix task did not elicit high enough task demands to prevent speakers from making adaptations, these results imply that speakers may be able to make adaptations to listeners regardless of cognitive demands.

Most instructed clear speech studies interpret their findings through Lindblom's (1990) H&H framework. However, the findings of the current study highlight the need for both the adaptation and clear speech literature to consider more holistic and realistic approaches to listener adaptation, which would integrate accounts of both the development of these interactive skills as well as the multiple levels of adaptation strategies available to speakers when modifying their speech to their listener.

## 11.3 Limitations and future research

As discussed in chapter 3, this study differs from the majority of previous studies examining speech production in HI children – which typically use a highly-controlled cohort of HI participants – in that the focus of the study was not on assessing the factors such as type of hearing aid or communication mode contributing to good speech production development, but rather on examining communication in a regular inclusive setting found in many mainstream schools in the UK. This study was therefore designed to be as ecologically valid as possible, both in its participants and in its speech elicitation methods. However, some methodological issues remain which may limit the scope of the current findings.

One such limitation is that in each school, the teachers ultimately selected both the HI and NH children who participated in the study. It seems possible that, at least in some of the schools, teachers only chose those NH and HI children to participate who they thought would perform well in a communicative task with a HI peer. It is not clear, therefore, how representative the NH and HI participants in the current study are of the general population in these schools.

Referential communication tasks were used in this study to enable the elicitation of spontaneous, but controlled, interactive speech (see chapter 2). Although children are likely to encounter such tasks frequently in problem-solving contexts both at school and at home, the tasks are nonetheless somewhat artificial (Lloyd, 2003). In particular, referential communication tasks have a high understanding criterion (Dunn, 2014; 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 tasks used in this study 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 social situation.

Additionally, due to the clarity rating study in chapter 9 being unable to differentiate between speaking conditions, we do not know whether the strategies used by speakers were helpful to their HI listeners – this somewhat limits the conclusions that can be drawn from the study. Future work needs to carefully consider the methodology to use in assessing the intelligibility of NH and HI children's spontaneous speech in different speaking conditions, as discussed in chapter 9.

The current study is novel in using a corpus of NH and HI children's spontaneous speech to elicit different speaking styles. However, the acquisition of a corpus is a time-consuming process, with a large amount of time taken up by recruiting and testing participants, and corpus transcription and checking. Thus, time constraints limited the extent to which the children's interactions could be analysed in this study. In particular, although the study attempted to analyse the main communication strategies used by NH and HI speakers in their peer-to-peer interaction, many of the participants may have been using other communication strategies to their HI friend which were not analysed here. For example, when asked by the researcher about the strategies they use during miscommunications with their friend, many participants mentioned using visual strategies, such as gesture or sign language (see Appendix G). 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. The video data has already proven a fruitful source for Conversation Analysis work on repair (Chu, 2015; Dunn, 2014; Ebrahim, 2015; Harris, 2014; Ní Almhain, 2014).

Similarly, as the Grid task elicited a large amount of miscommunications relating to minimal pair keywords in NH-HI and HI-HI pairs, it may be very valuable in future work to analyse the acoustic-phonetic properties of the participants' speech occurring before and after miscommunication in conjunction with the repair strategies used, similarly to case studies on adult NH-HI communication by Lind *et al.* (2010). This would enable us to explore whether NH and HI children modify their speech dynamically during the interaction, as predicted by Lindblom's (1990) H&H model. It would also allow an examination to be done on whether acoustic-phonetic strategies interact with higher-level pragmatic strategies used by the speakers. For example, in their Conversation Analysis studies of two of the groups of four participants used in the current study, Ní Almhain (2014) and Harris (2014) suggest that global acoustic-phonetic enhancements were sometimes used in participants' repair responses, but often other adaptive repair strategies, such as cue responses or adding gesture, were used instead.

This study analysed the adaptation strategies used in peer-to-peer interaction by NH and HI children, as effective communication with peers is likely to be very important to school-aged NH and 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 both the NH and HI children's continued speech and language development, as no similar data has been collected from NH and HI adults' interactions – a further potential focus of future work.

## 11.4 Conclusion

In summary, the findings of this thesis suggest that both NH and HI older children are surprisingly adept at using acoustic-phonetic and linguistic adaptation strategies when interacting with a HI interlocutor. While most previous studies attempting to assess the characteristics of speech directed to HI persons have used read speech tasks, in which speakers are instructed to speak 'as if to a hearing-impaired person' compared to 'casually as if to a friend' (e.g., Ferguson, 2004; Picheny *et al.*, 1986), this study is novel in assessing spontaneous speech adaptations made in a real communicative situation in peer interaction, to allow for the elicitation of more realistic strategies used by speakers. In particular, a referential communication task, the Grid task, was successfully developed to enable an analysis to be done of several different types of speech measures – including global acoustic-phonetic, segmental, and linguistic measures. As the majority of HI children in the UK attend mainstream schools with both NH and HI peers, it is vital to explore the speech communication strategies being used by children in these school environments, to ensure that both NH and HI children obtain the maximum benefit from inclusive education.

# Appendix A

# Grid task materials

## A.1 Task introduction

The Grid task was developed (1) to elicit multiple repetitions of three types of phonetic contrasts (/p/-/b/, /s/-/ʃ/ and /i/-/ɪ/) in a spontaneous communicative task (which may also enable the potential elicitation of several miscommunications in the task in NH and HI interaction by using phonetic contrasts which may be difficult to produce and perceive by HI children), (2) to elicit multiple repetitions of several different vowels to enable measures of vowel space to be taken, and (3) to enable an assessment of the influence of task difficulty on adaptation measures by creating an interactionally simpler and possibly easier task as an alternative to the Diapix task. Elicitation of multiple repetitions of phonetic contrasts in (1) was achieved by including several minimal pair keywords in the task, and multiple repetitions of different vowels in (2) was achieved by including colour-number words including these vowels as part of the task. As required in (3), the Grid task is a simpler task than Diapix as it is more controlled, requires less complex linguistic knowledge and has predetermined turn-taking. For greater detail on the aims and implementation of the Grid task, see chapter 2.

## A.2 Grid randomisation process

In the task, each participant is given a picture-grid and an empty grid (see Figures in A.3) on a board (see Figure 2.1a), and a tray of labelled cards containing five picture versions of each of 16 keywords (see the Figures in section A.4 for the versions of each picture and Figure 2.1b for an example of the tray given to each participant). The aim of the task is for each interlocutor, without being able to see each other's 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 for the keyword (i.e., the correct colour-number square). The participants were instructed to work from the top left hand corner horizontally, row by row, with the interlocutors taking it in turns to describe the squares on their picture-Grid to each other (see Appendix B for the exact instructions given to participants). Each interlocutor's empty grid has the same colour-numbers on it as the other's picture-grid, but in a randomised order to prevent the pair of participants from identifying the location of the correct square without mentioning the colour-number word. Additionally, the keywords on each participant's tray were randomised, so that the two participants could not identify the keyword in the tray without referring to the keyword. To prevent participants from identifying keywords as being minimal pairs, the keywords in the four trays used during the experiment were distributed so that no minimal pair keywords were located next to each other.

Grids 1A and 2A, 1B and 2B, 1C and 2C, 1D and 2D, 3A and 4A, 3B and 4B, 3C and 4C and 3D and 4D are linked – one participant in the pair works on one grid, while the other works on the other grid. In one condition, a pair of participants is given either the grids numbered 1 and 2, or those numbered 3 and 4. Each participant in the same condition works through the grids of the same number – therefore, for example, in one condition, one of the participants would be given grids 1A, 1B, 1C and 1D, while the other participant would work through grids 2A, 2B, 2C and 2D. Ideally, four pairs of grids would be done per condition per pair.

Within each grid number, the grids labelled A and B, and those labelled C and D, together contain all 16 keywords and all 16 colour-number words (i.e., two consecutive grids completed by one participant do not share keywords or colour-number words). Thus, if one participant completes grids 1A and 1B in a condition, they will need to mention all 16 keywords and all 16 colour-number words to their interlocutor, but if they work through grids 1A to 1D, they will describe each keyword and colour-number twice to their interlocutor. Within the 8 keywords that occur in each grid, three are VOT keywords, three are /s/-/ʃ/ keywords, and at least two are /i/-/ɪ/ keywords. These keywords must include instances of different sounds – for example, in one grid, for the VOT keywords, at least one must be a /b/ keyword and at least one must be a /p/ keyword. To ensure that potential minimal pair keyword confusions would not be excluded by the interlocutor on the basis of expectancy, as far as possible, each grid does not include minimal pair keywords – i.e., if a grid contains 'sheep', it will not also include 'ship'. However, this was not always possible due for example to the 'triple' minimal pairs of 'bean'-'bin'-'pin'. For the 8 numbers that occur in each grid, there are two of each number and two of each colour. The location of the colour-numbers on the grid were restricted so that the same colour or the same number cannot be located next to each other in the picture-grid, to reduce the likelihood of contrastive accent being applied to the colour-number word.

Additionally, to counteract any position effects, the eight squares in grids A and D (and B and C), within each grid number contain the same keywords and colour-number words. However, the exact location of the keyword or colour-number will differ from grid A to grid D (and from Grid B to Grid C) – those keywords/colour-numbers which occur in the first row of one grid occur in the second row of the other grid.

Within each condition, for the 1-2 grids and the 3-4 grids, the same picture version of a keyword never occurs twice – this was done to reduce familiarity with different keyword versions. To further reduce expectancy, half of the keywords are shared between the grids of the two interlocutors. It was also attempted not to have the same keyword or a minimal pair keyword described after each other between the two grids of a pair – however this was not always possible due to the prior randomisation of keywords as described above.

To randomise the grids, participants were given the grids in the order of A to D or D to A. If they worked on the grids labelled 1-2 in one condition, they would be given the grids labelled 3-4 in the other condition, and vice versa.

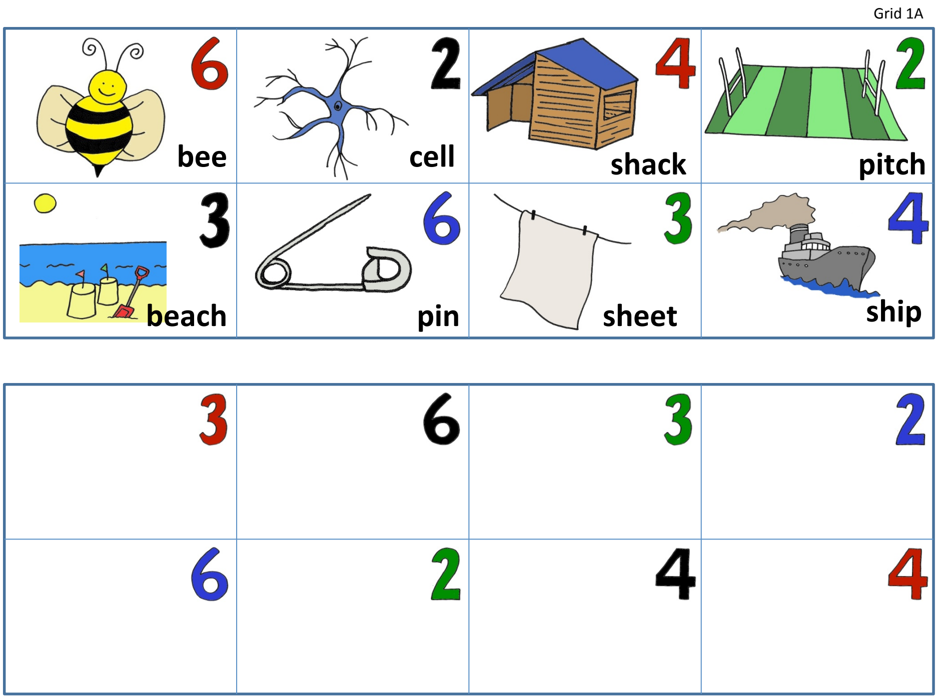
Correct answer sheets were provided to participants after completion of each Grid (see section A.3).

## A.3 Grids and correct answers

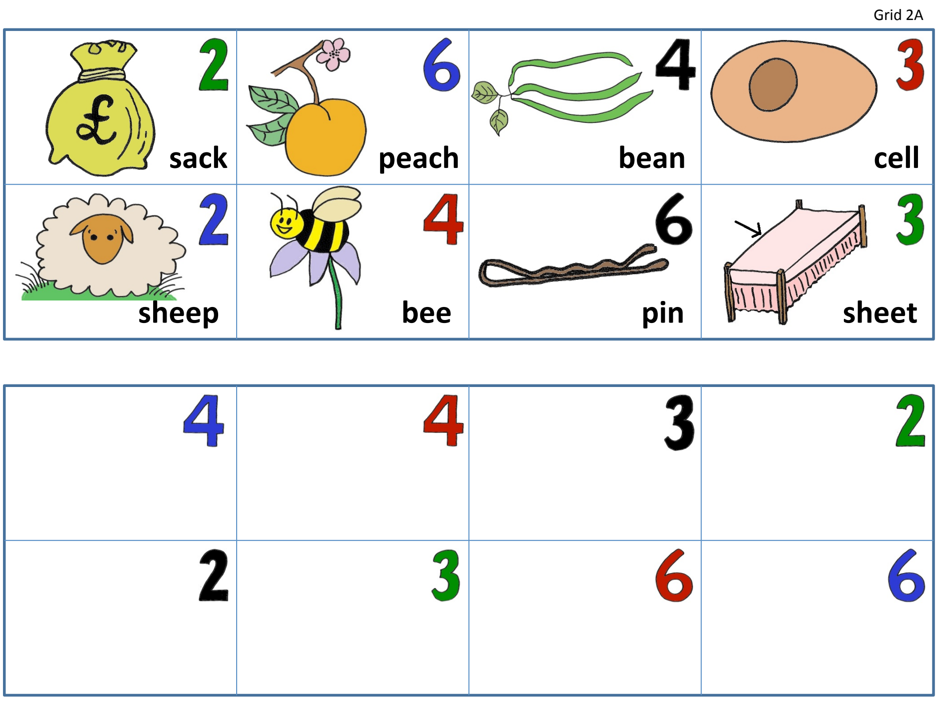
Figures A.1, A.3, A.5, A.7, A.9, A.11, A.13 and A.15 represent the paired grids given to each participant pair in the grid task. Each grid pair is followed by the correct answer sheets (figures A.2, A.4, A.6, A.8, A.10, A.12, A.14 and A.16).

## A.4 Keyword pictures

Figures A.17, A.18 and A.19 display the five picture versions of each keyword present in the Grid task. Keyword versions differed from each other in representing either (1) different types of a certain object (e.g., rugby, football, baseball and cricket pitches), (2) the same types of object but differing in details (e.g., bees with different numbers of stripes, different kinds of faces and wings), or (3) a mixture of both types of differences.

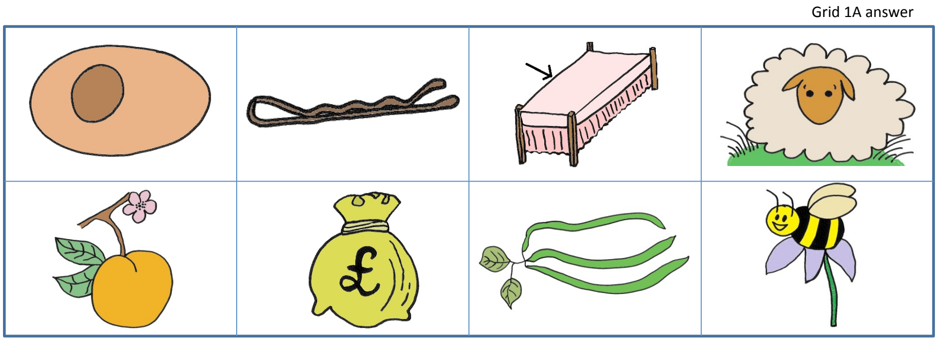



(a) 1A




(b) 2A

Figure A.1: Grids 1A-2A



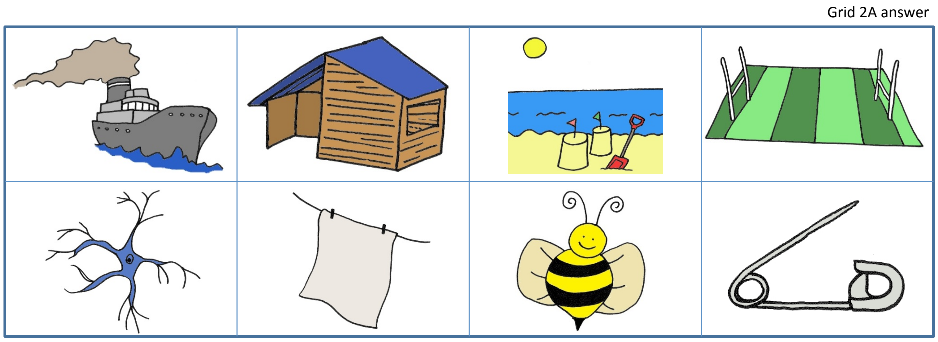
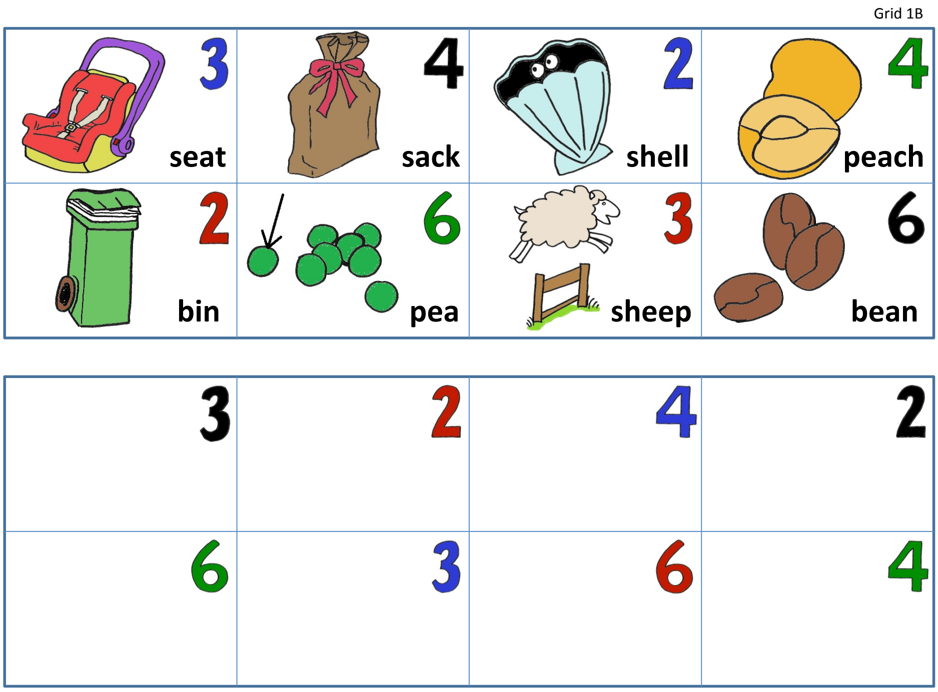
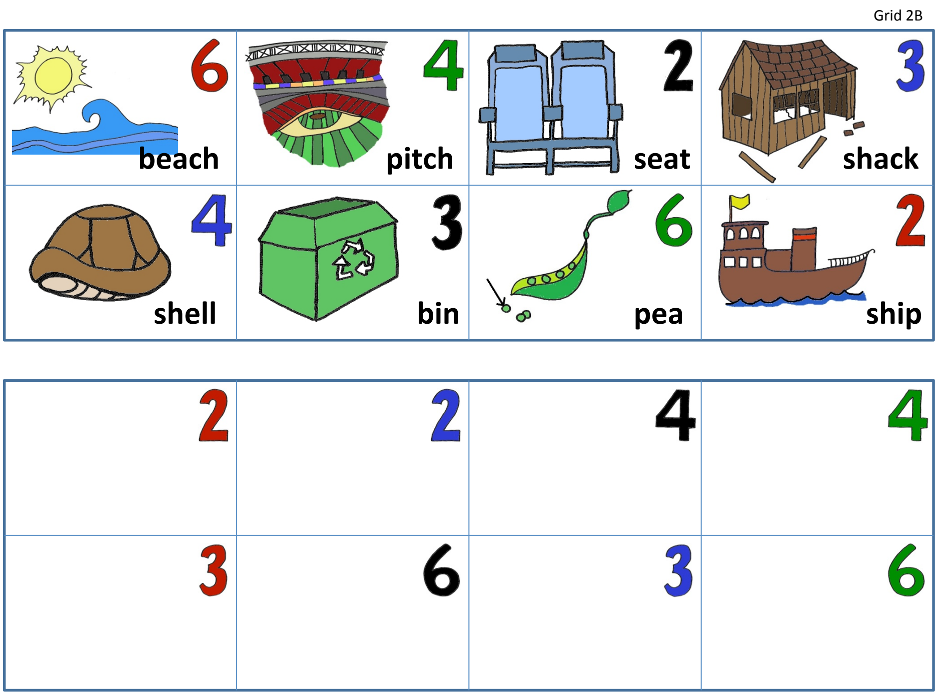




Figure A.2: Grids 1A-2A correct answers




(a) 1B




(b) 2B

Figure A.3: Grids 1B-2B

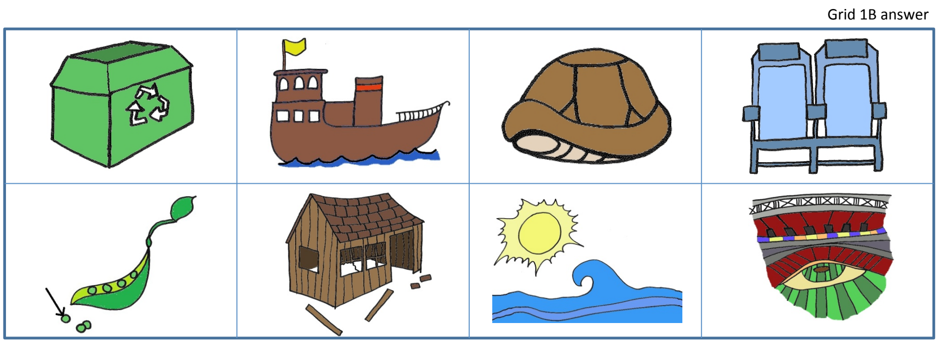
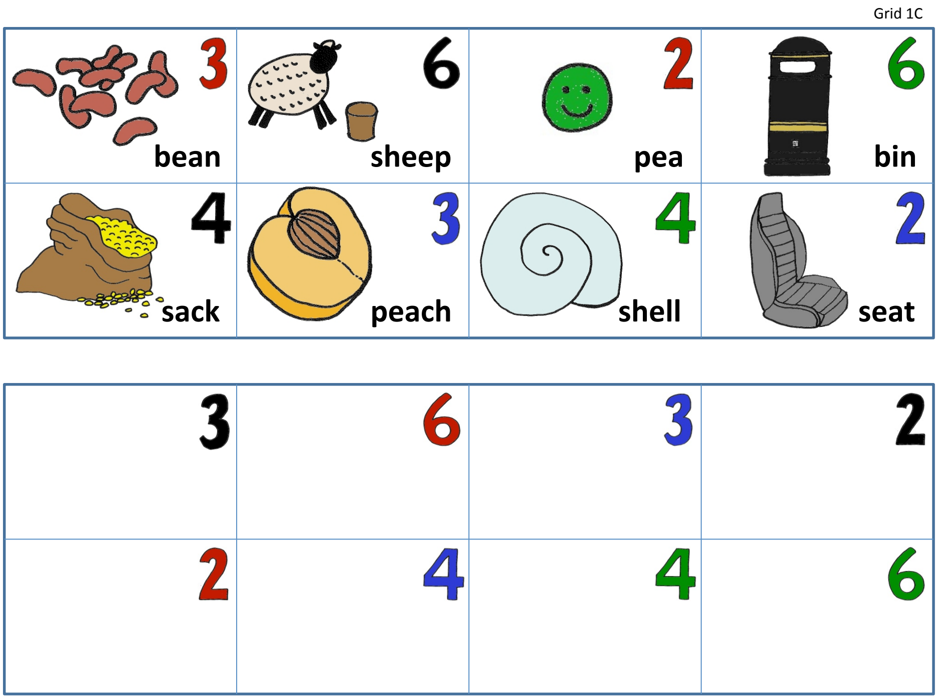
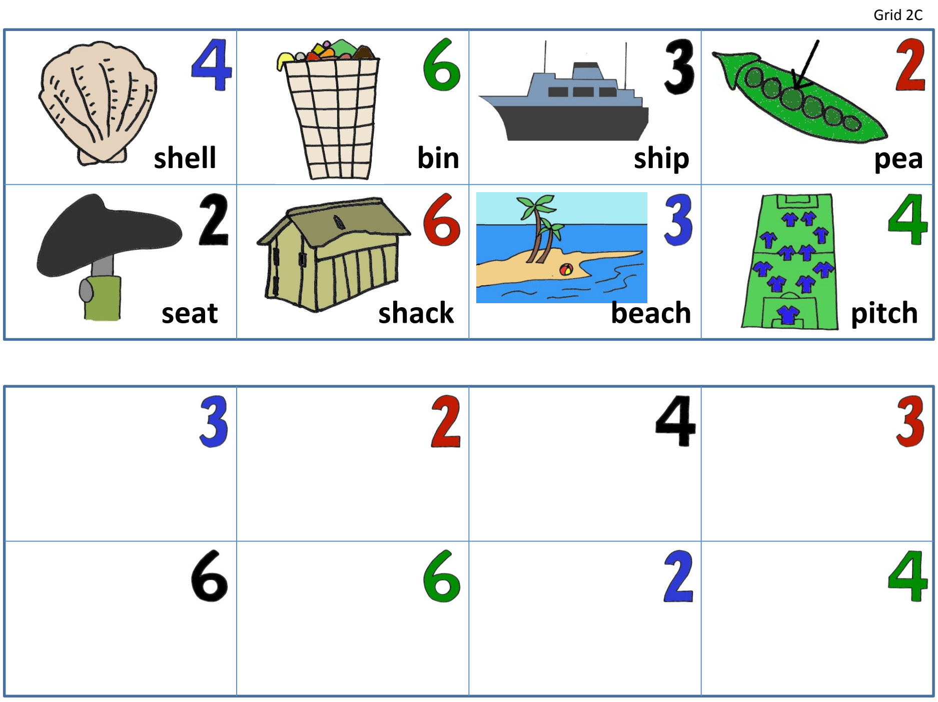




Figure A.4: Grids 1B-2B correct answers




(a) 1C




(b) 2C

Figure A.5: Grids 1C-2C

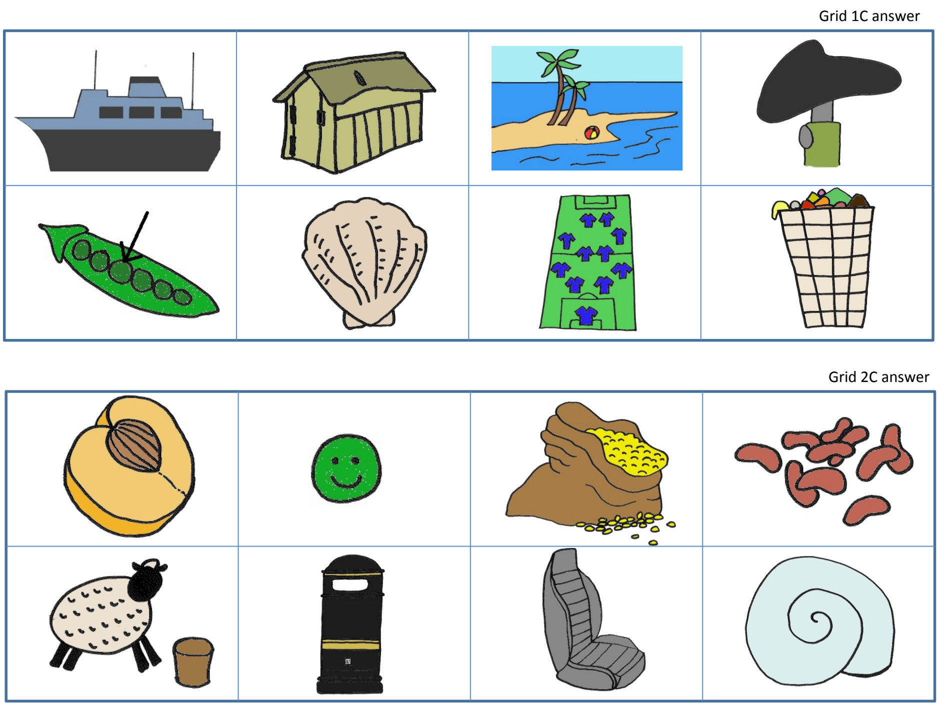
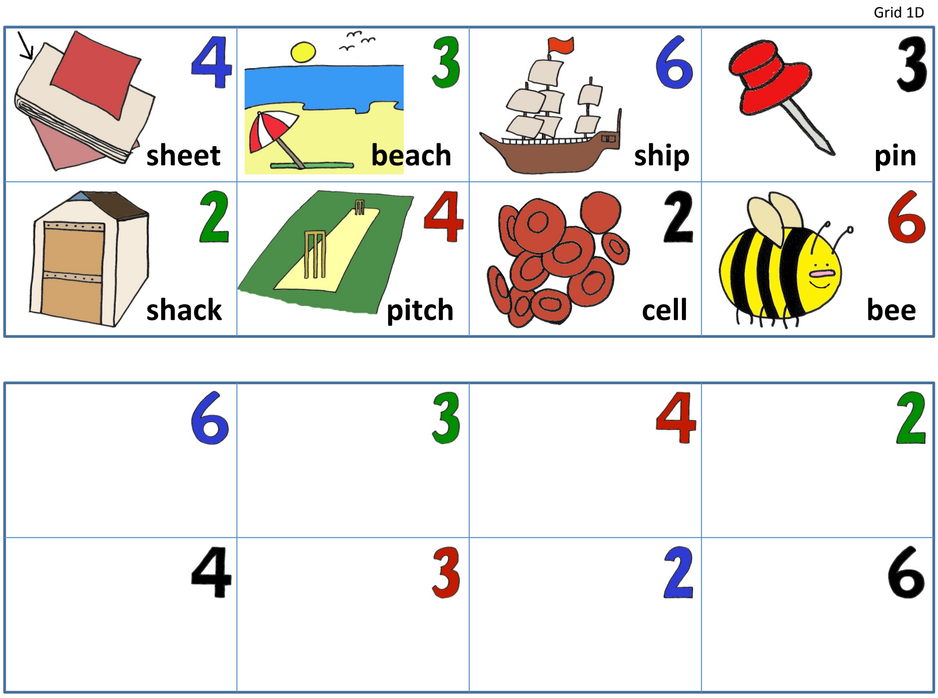
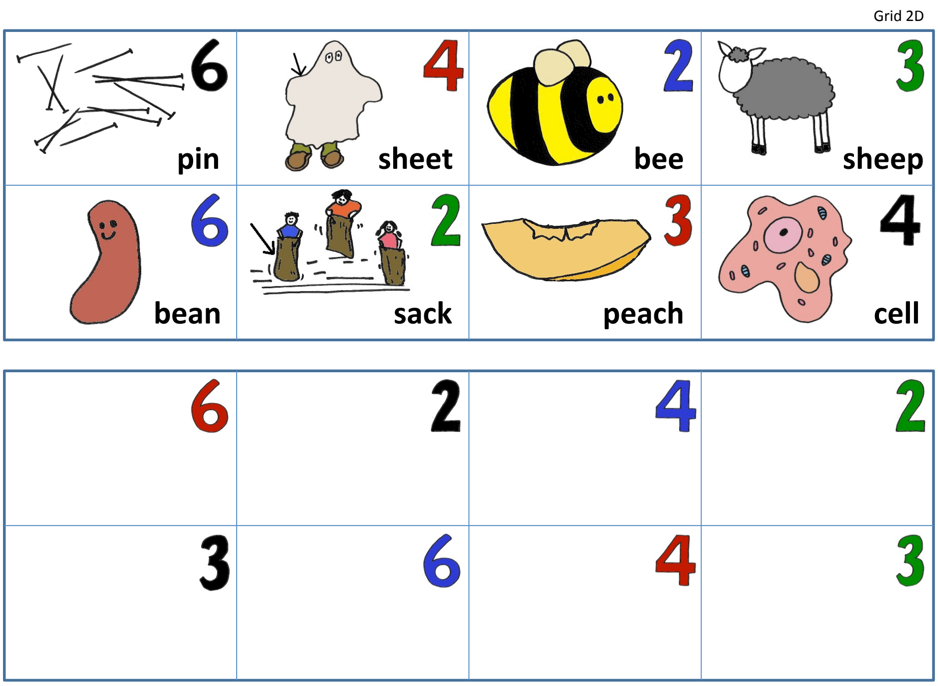



Figure A.6: Grids 1C-2C correct answers




(a) 1D




(b) 2D

Figure A.7: Grids 1D-2D

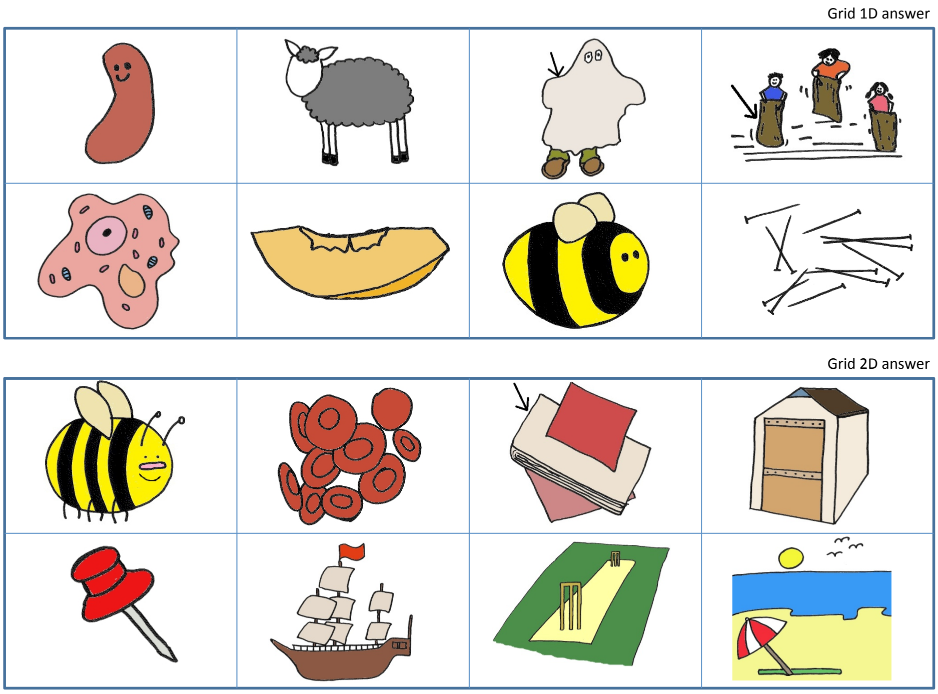
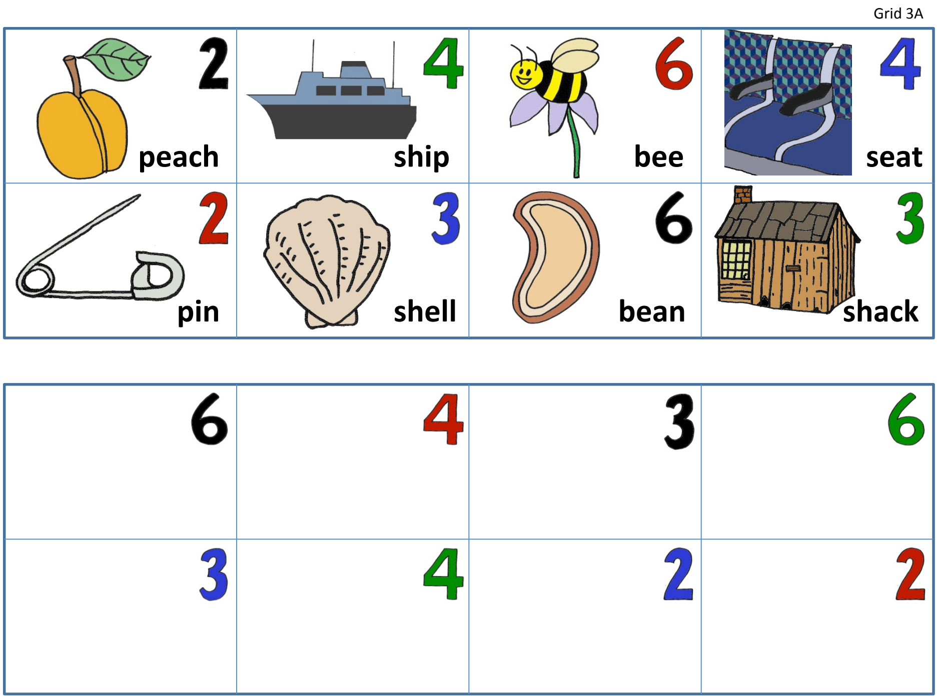
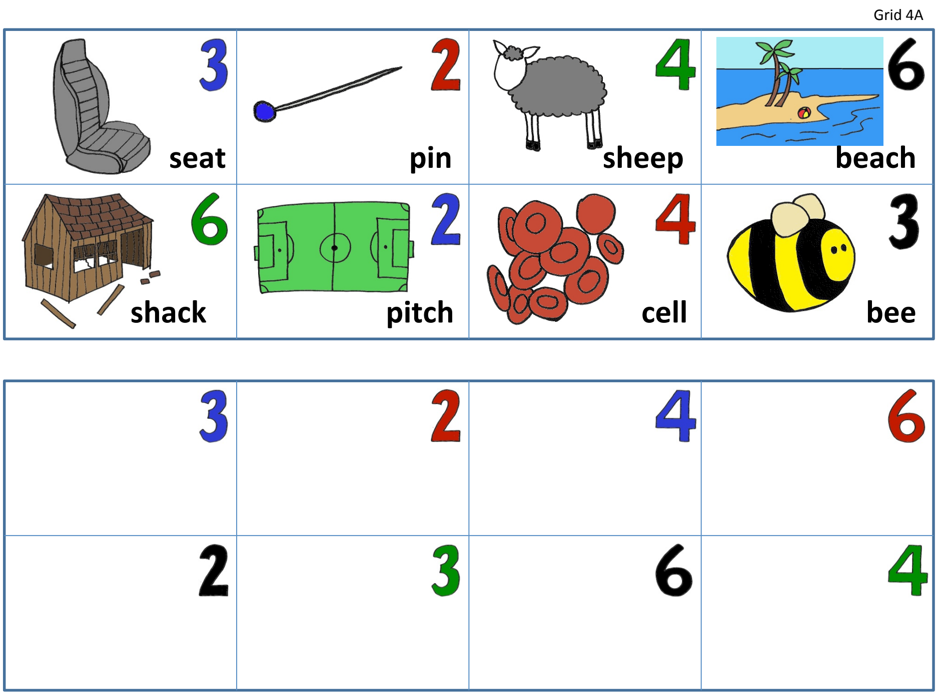



Figure A.8: Grids 1D-2D correct answers




(a) 3A




(b) 4A

Figure A.9: Grids 3A-4A

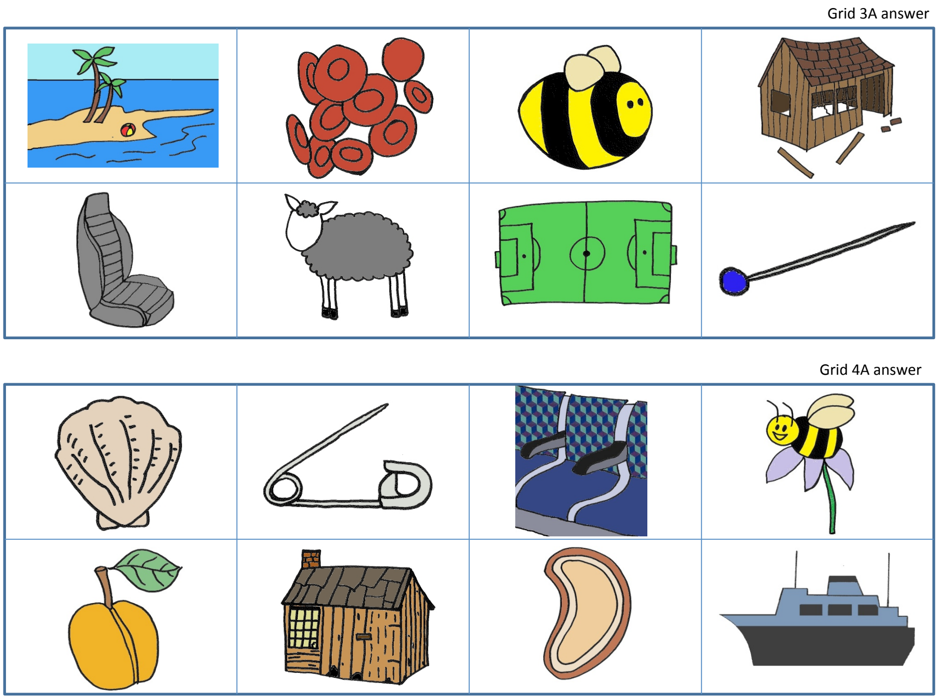
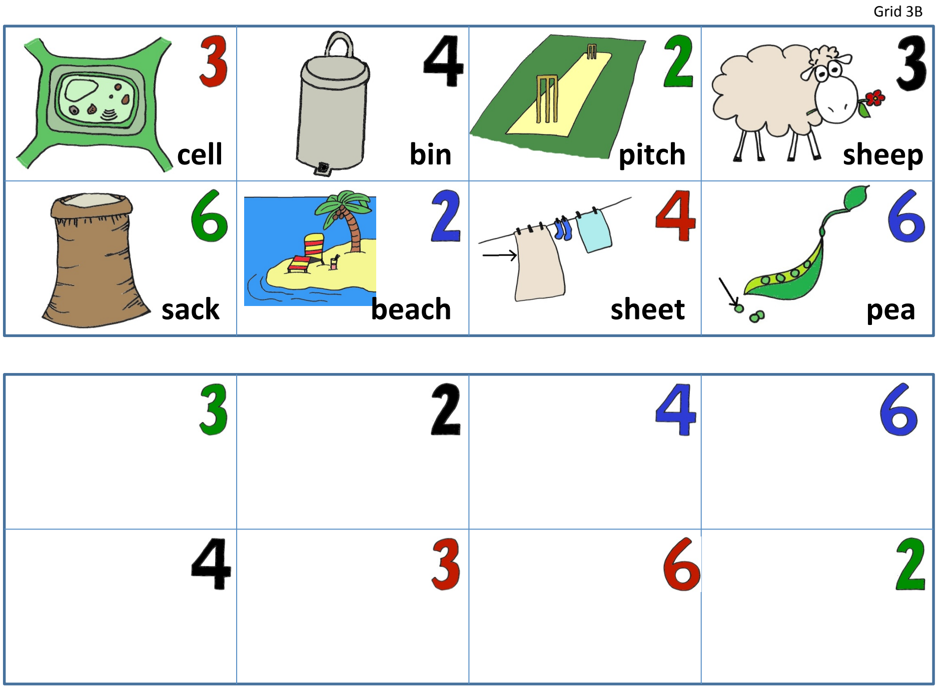
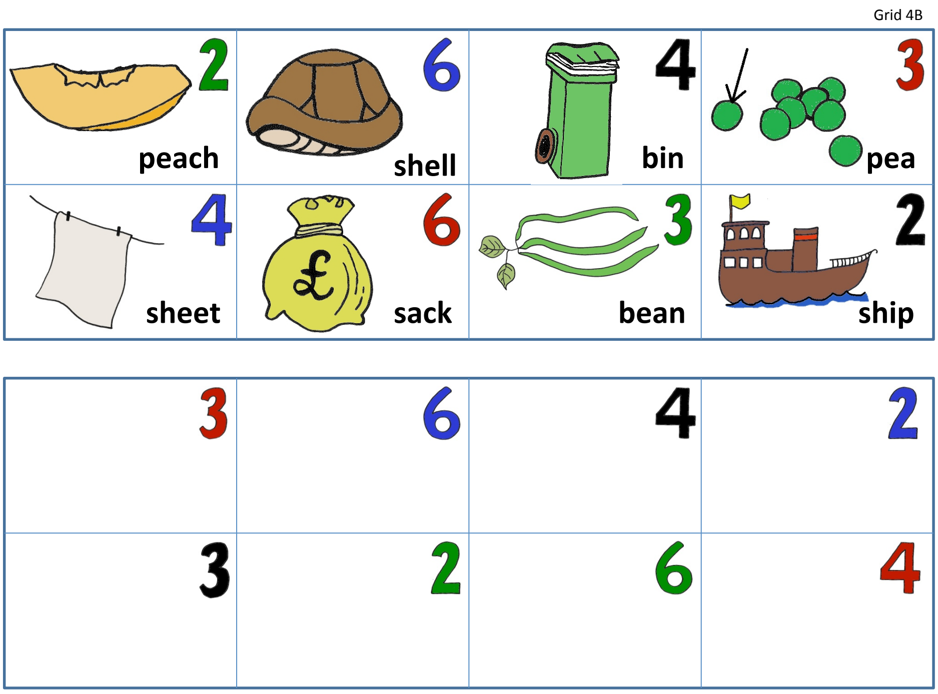



Figure A.10: Grids 3A-4A correct answers




(a) 3B




(b) 4B

Figure A.11: Grids 3B-4B

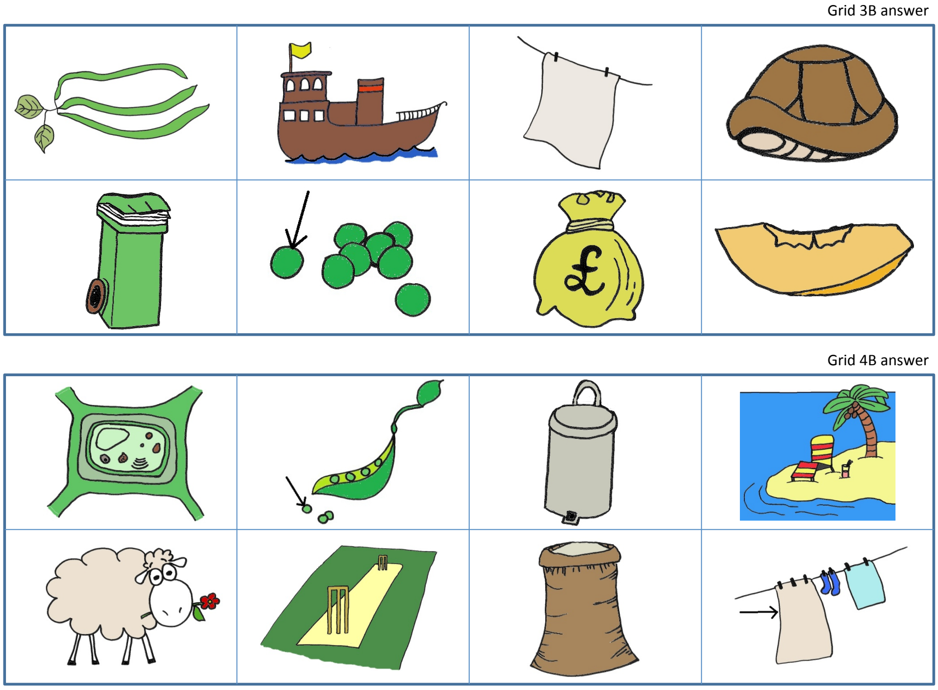
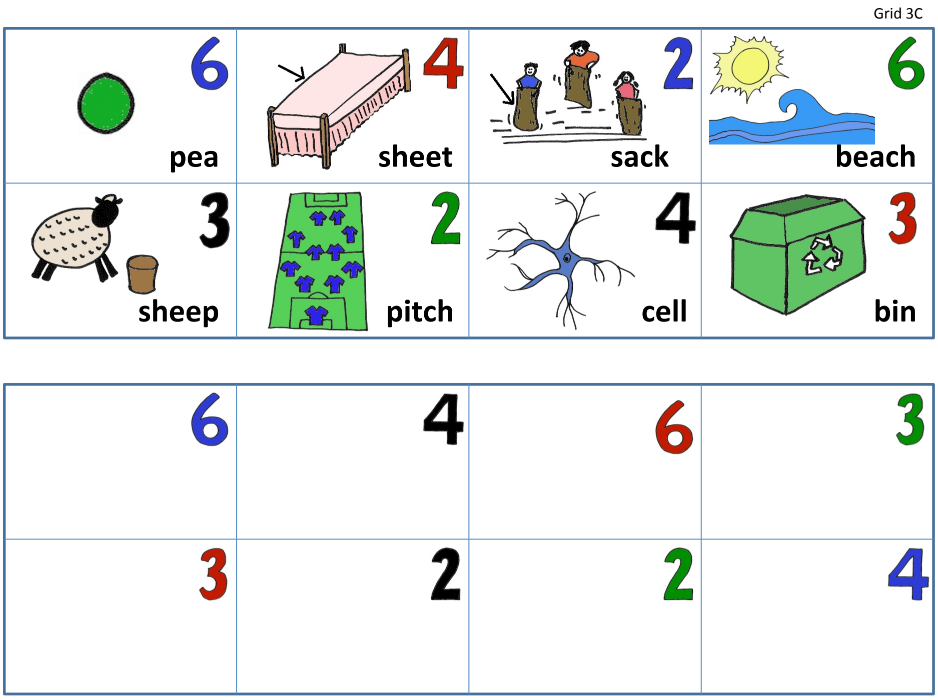
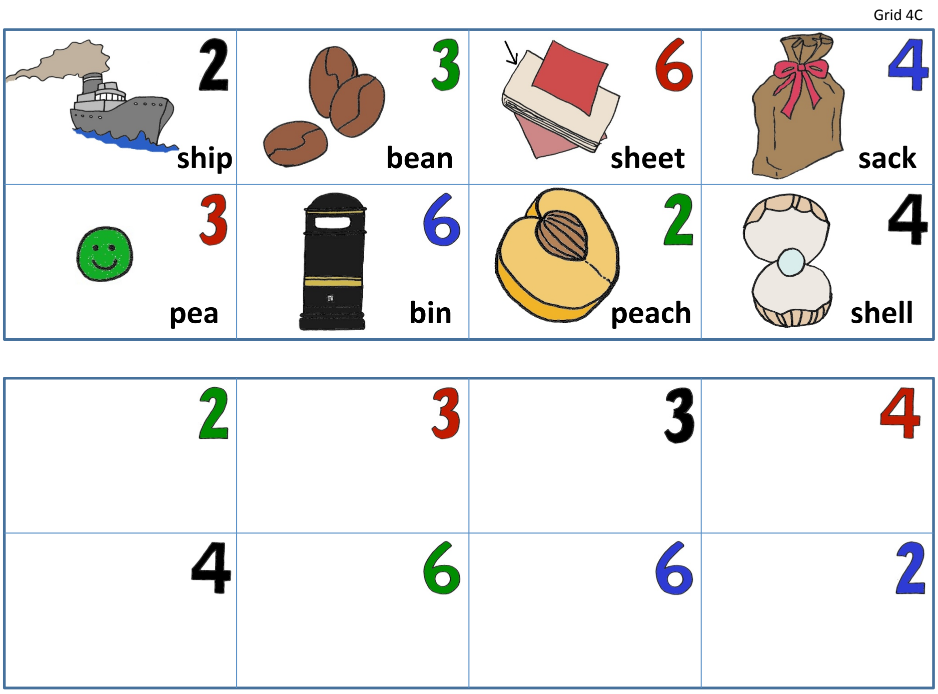



Figure A.12: Grids 3B-4B correct answers

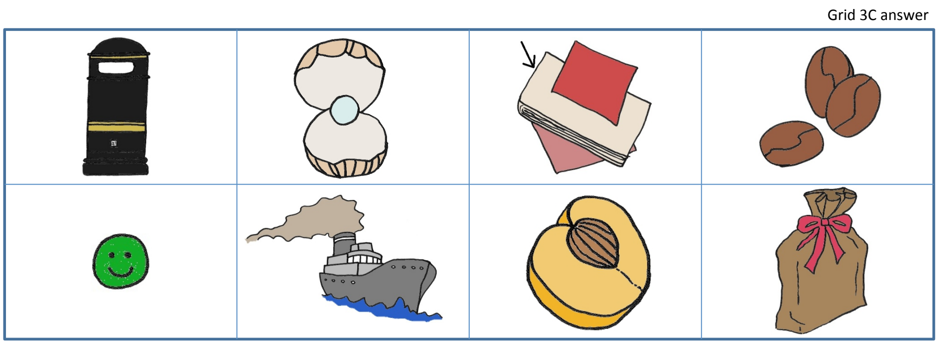



(a) 3C




(b) 4C

Figure A.13: Grids 3C-4C



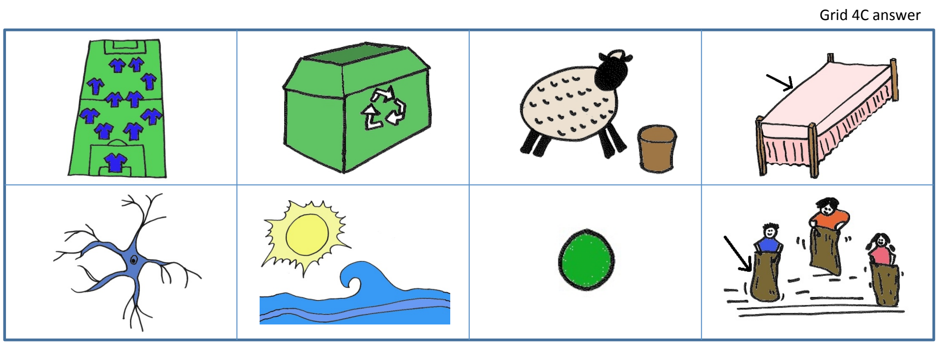
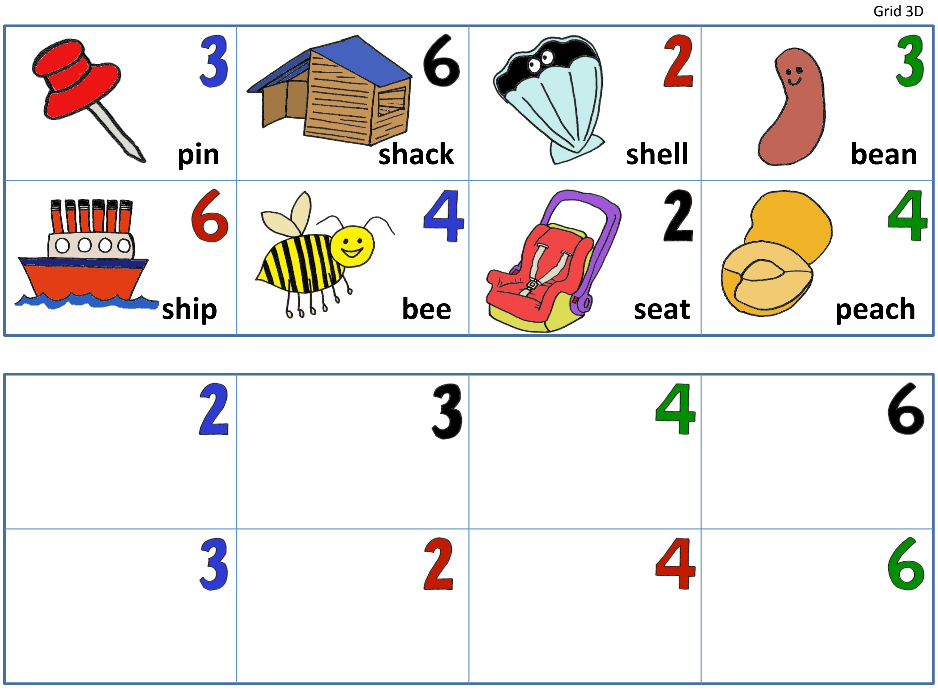
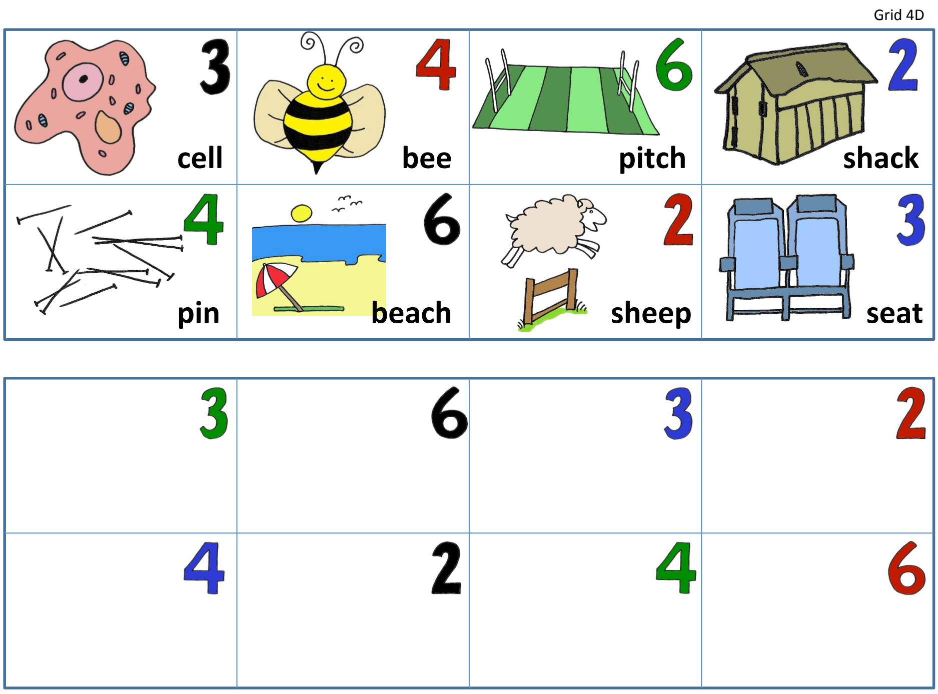




Figure A.14: Grids 3C-4C correct answers




(a) 3D




(b) 4D

Figure A.15: Grids 3D-4D

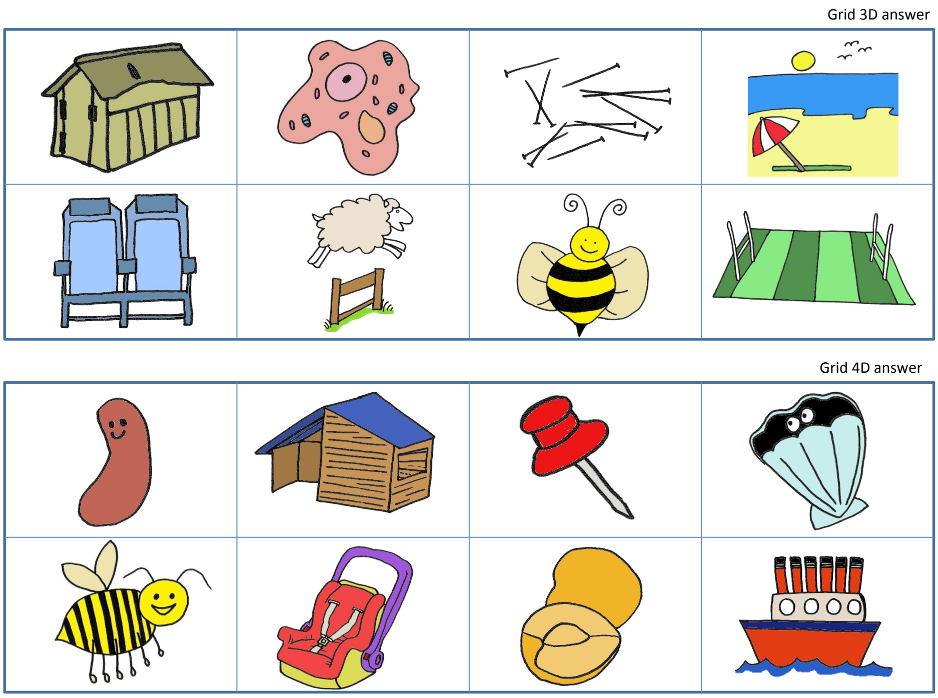
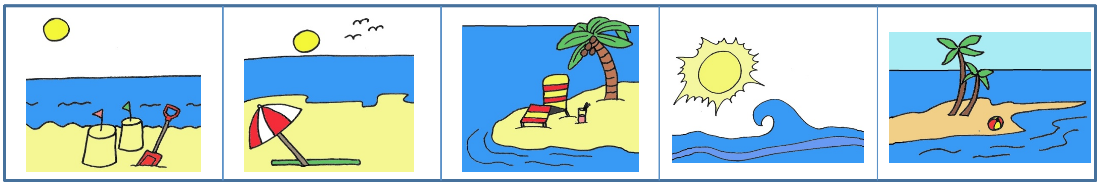
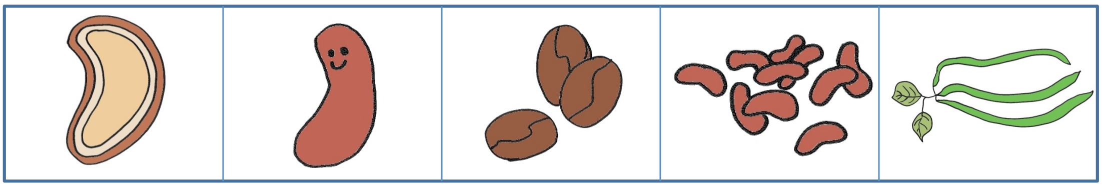


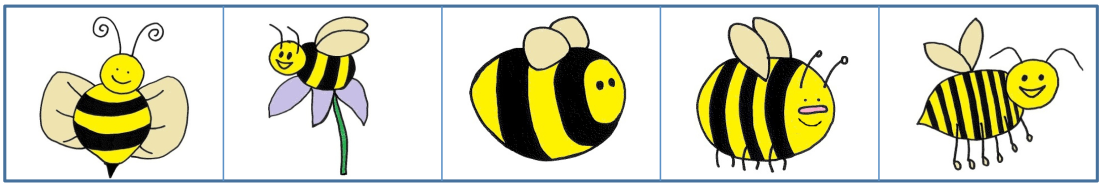

Figure A.16: Grids 3D-4D correct answers



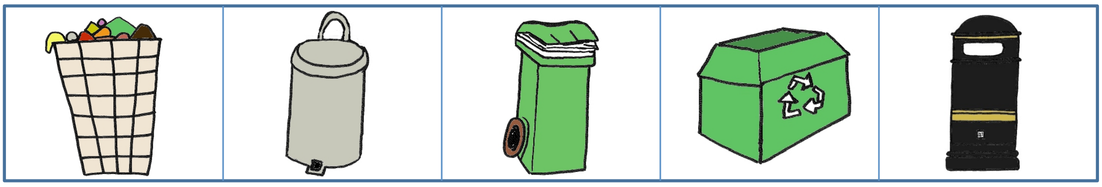
(a) beach



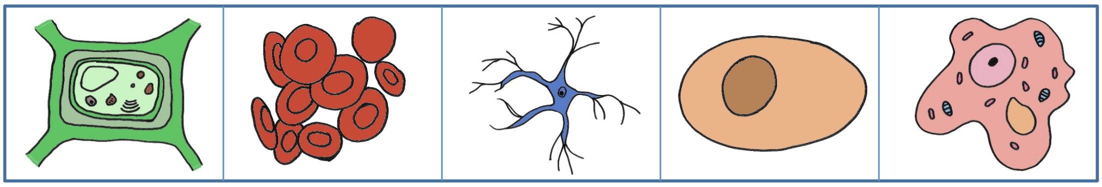
(b) bean



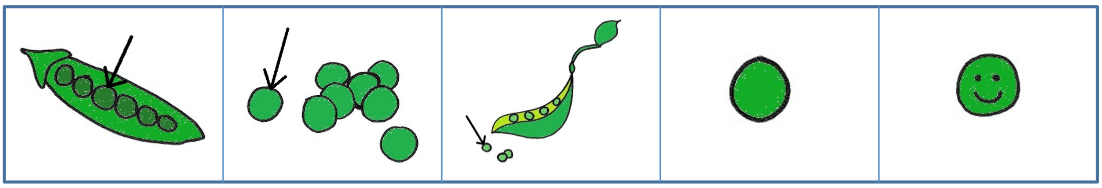
(c) bee



(d) bin

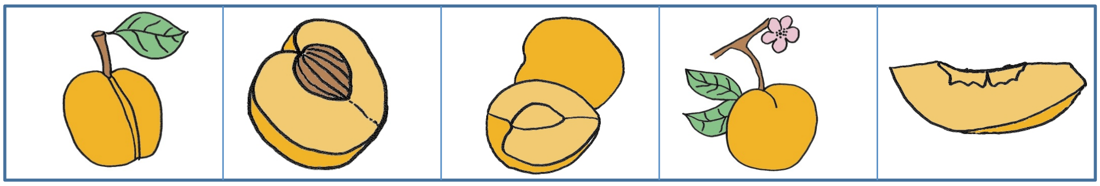


(e) cell

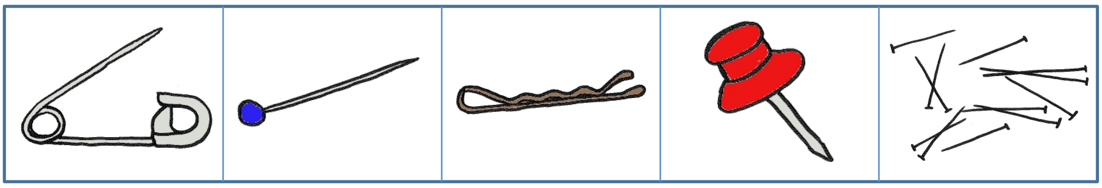


(f) pea

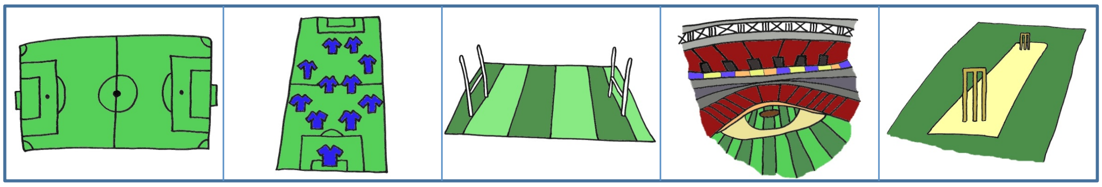
Figure A.17: Grid task keywords 1/3



(a) peach

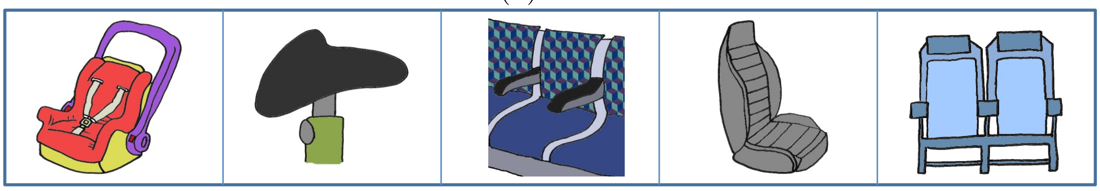


(b) pin

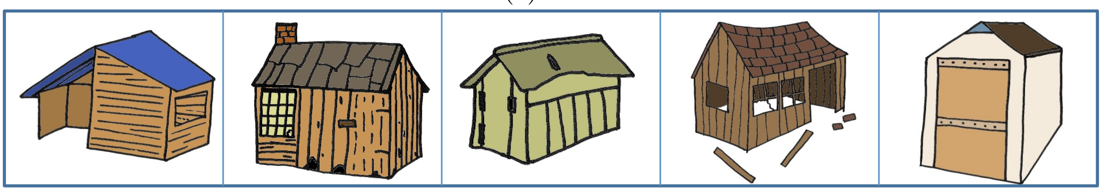


(c) pitch

(d) sack

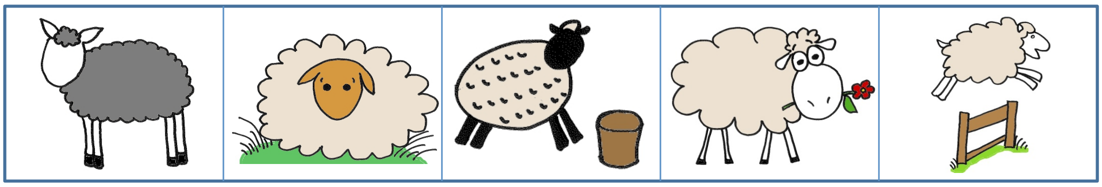


(e) seat

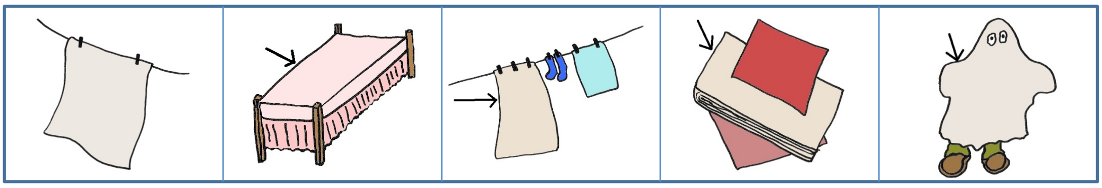


(f) shack

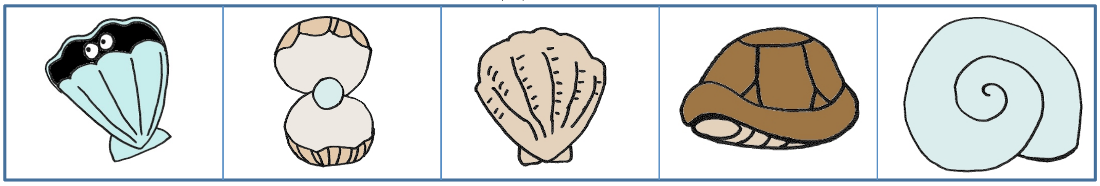
Figure A.18: Grid task keywords 2/3



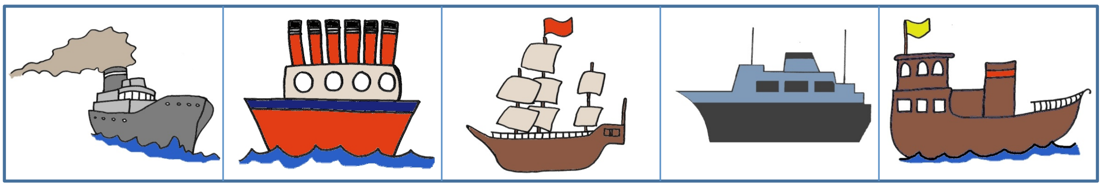
(a) sheep



(b) sheet



(c) shell



(d) ship

Figure A.19: Grid task keywords 3/3

# Appendix B

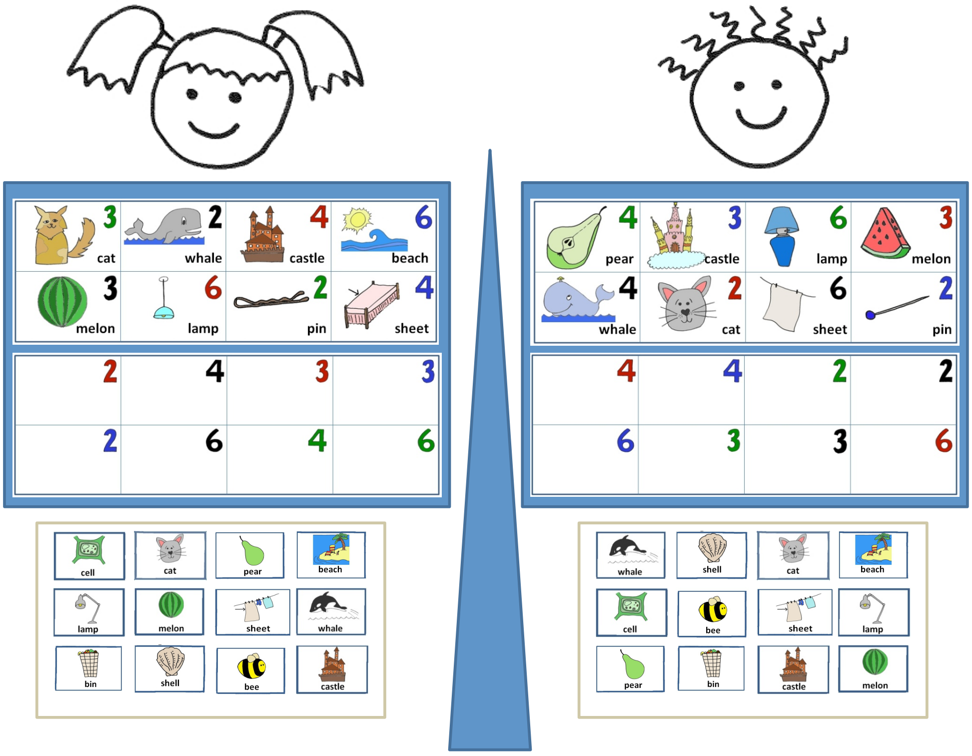
# Grid task instructions

Figures B.1 to B.9 display some of the slides used to illustrate to participants how the Grid task is played. The instructions given to participants by the researcher while showing them the slides were as follows.

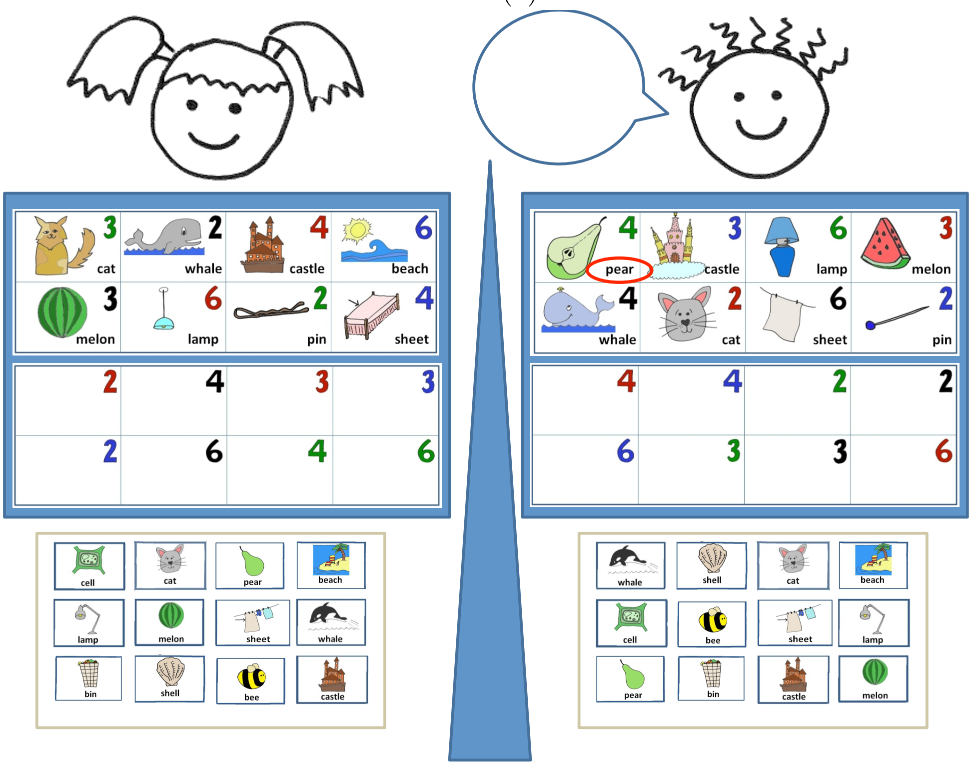
"[B.1a] Here we have a girl and a boy. The girl has a board with a picture-grid and an empty grid, and a tray full of cards. The boy also has a board with a picture-grid, an empty grid, and a tray of cards. The girl and boy can see each other, but they cannot see the front of each others' boards, or each others' trays. The aim of the game is for the girl to find the boy's pictures to put in her empty grid, and for the boy to find the girl's pictures to put in his empty grid. They also have to find the correct colour-number square to put the picture on.

[B.1b] Now the boy starts the game from the top left-hand corner of his picture-grid. First, he tells the girl what kind of picture he's got, so he might say: 'I've got a pear.'. [B.2a] Then, the girl looks in her tray, and finds the cards with pears on them. [B.2b] She takes the five cards out, and asks the boy: 'What does your pear look like?'. [B.3a] The boy then tells her about his pear. For example he could say: 'The pear is cut in half, and it has seeds in it.'. [B.3b] Then the girl knows that it must be this pear. But she does not yet know which square she should put the pear on. [B.4a] So the boy tells her: 'Put it on green four'. [B.4b] And the girl finds the green four, [B.5a] and puts the pear in that square in her empty grid.

[B.5b] Then, it's the girl's turn. She also starts from the top left-hand corner. She tells the boy: 'I've got a cat.'. [B.6a] So the boy finds the cat cards from his tray, [B.6b] takes out the five cards, [B.7a] and asks the girl: 'What does your cat look like?'. Then the girl tells the boy about her cat, for example she could say: 'It's colourful, and it's got a bushy tail.'. [B.7b] Then the boy knows which cat it is. But he doesn't yet know which square to put it in. [B.8a] So the girl tells him: 'Put it on green three.'. [B.8b] So the boy finds his green three, [B.9a] and puts the cat in the right square. [B.9b] After this it's the boy's turn again. He works his way from left to right on his picture-grid, and so the next picture he will talk about is the castle."

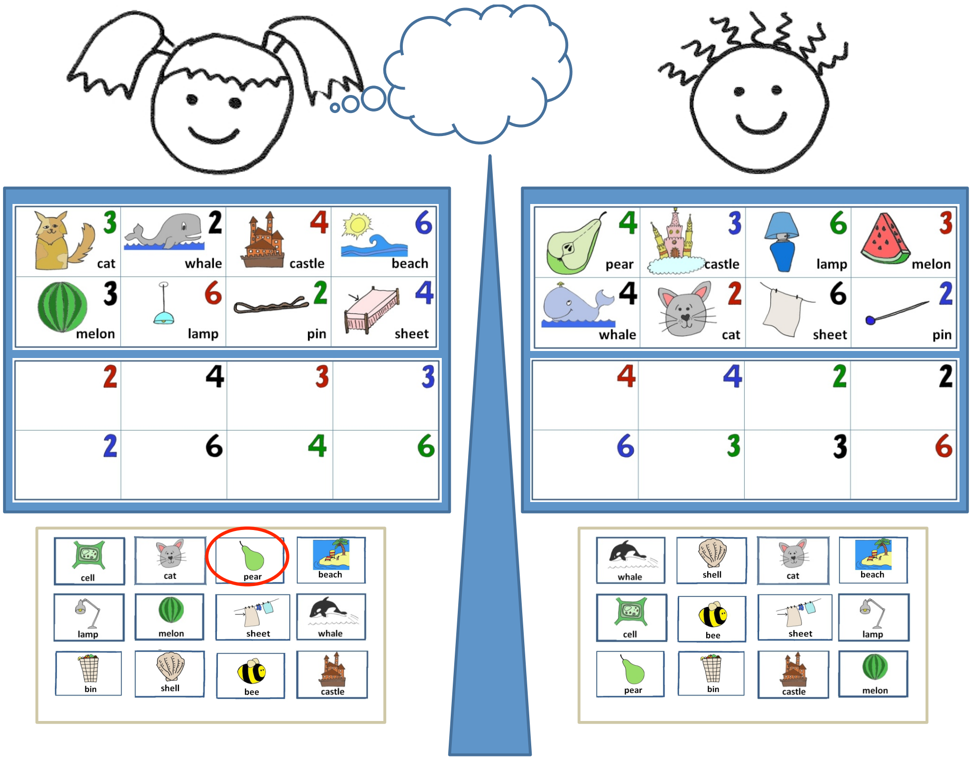


(a) 1

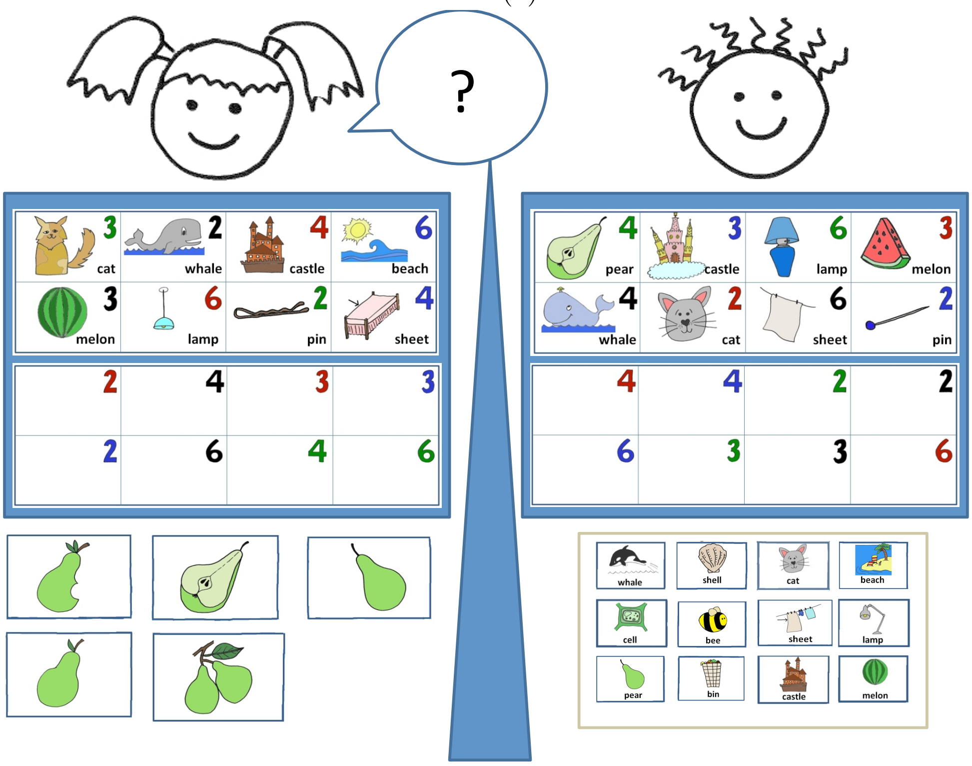


(b) 2

Figure B.1: Grid instructions: 1,2

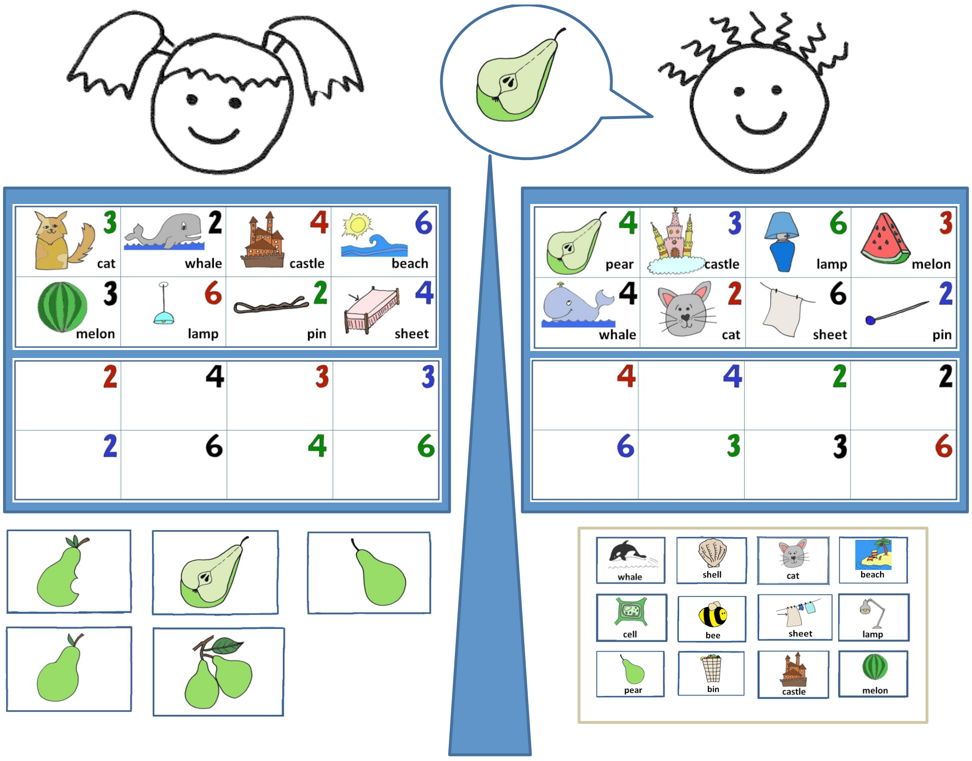



(a) 3

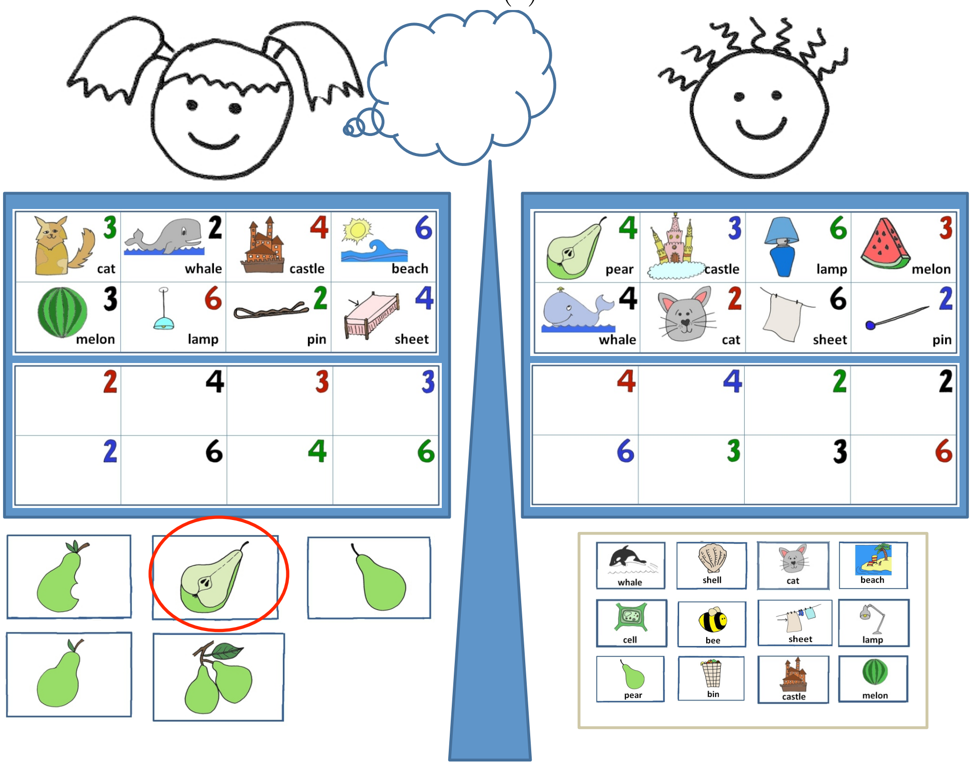



(b) 4

Figure B.2: Grid instructions: 3,4

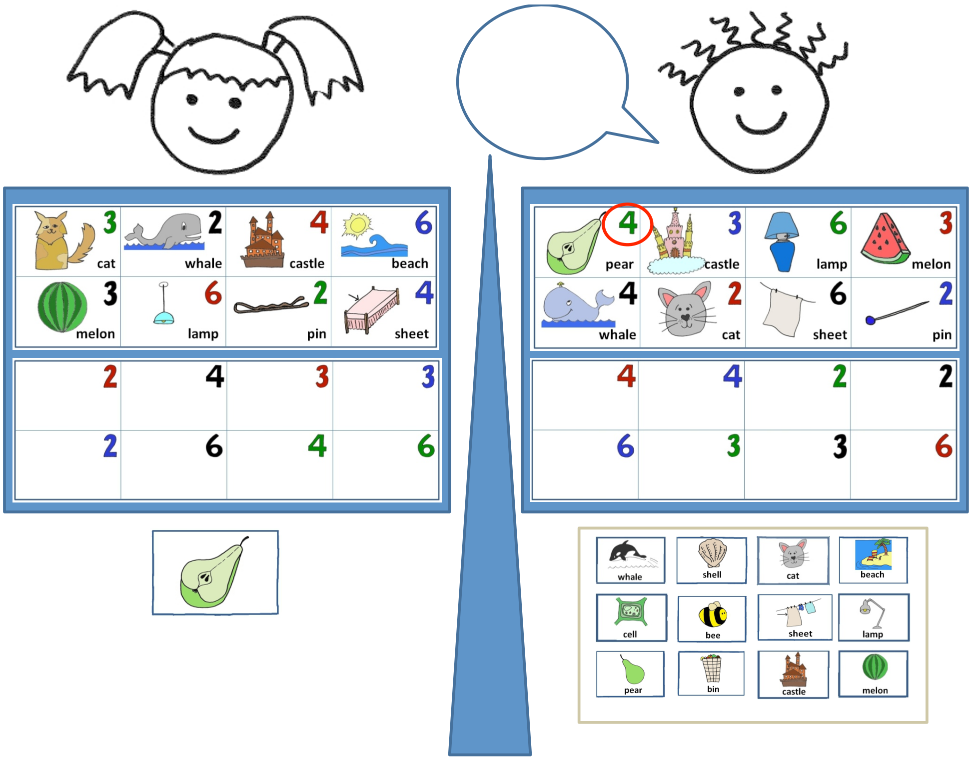



(a) 5

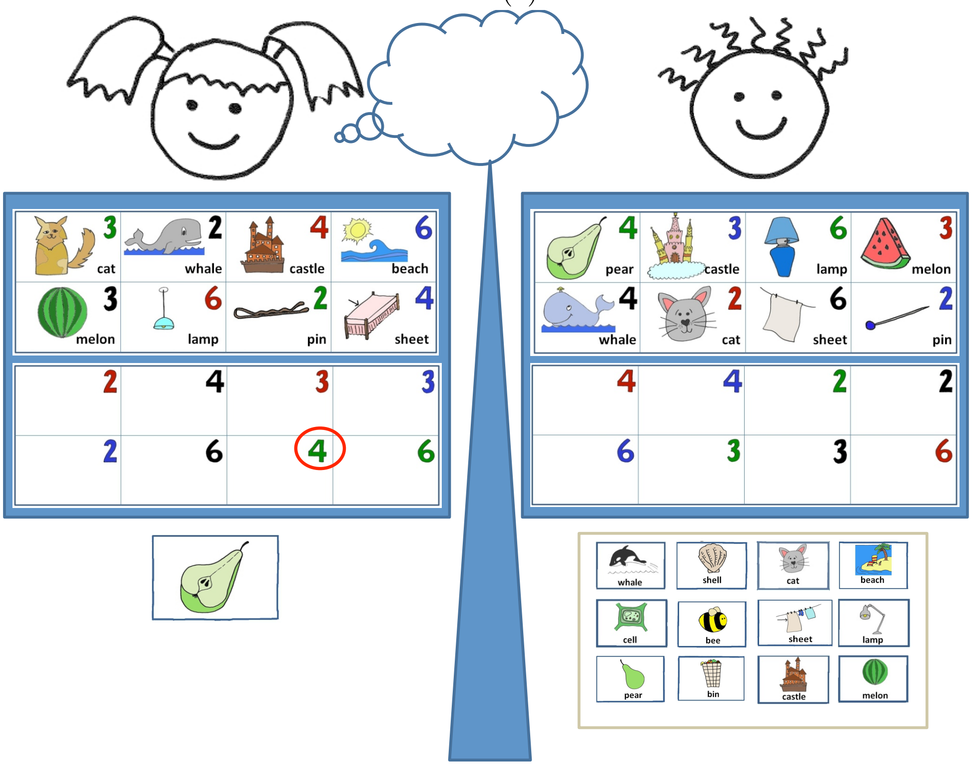



(b) 6

Figure B.3: Grid instructions: 5,6

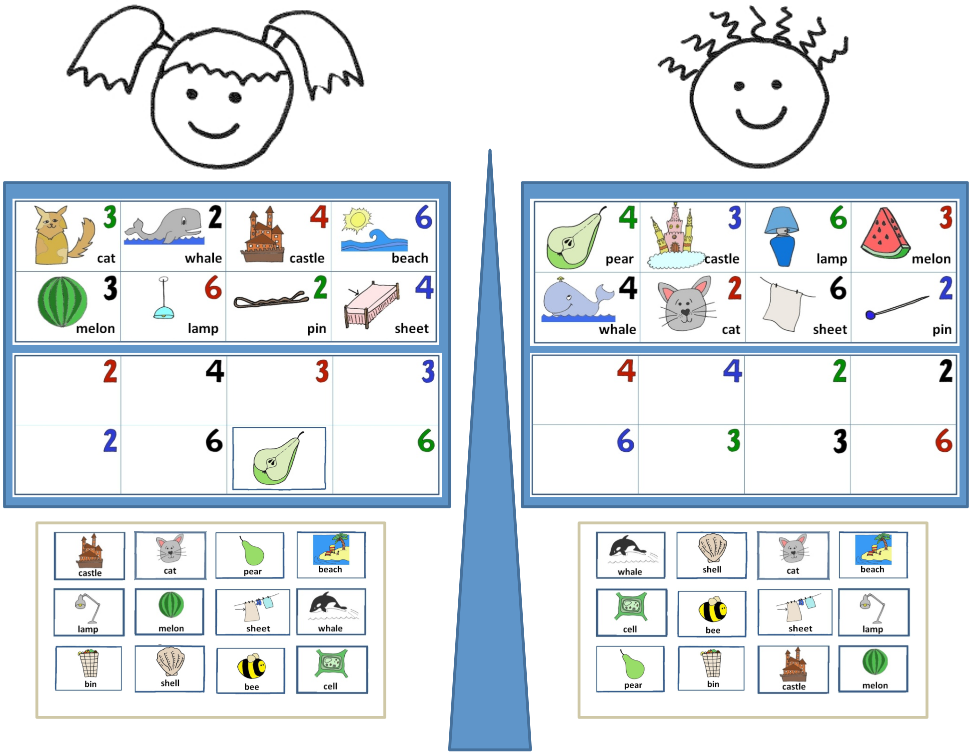



(a) 7

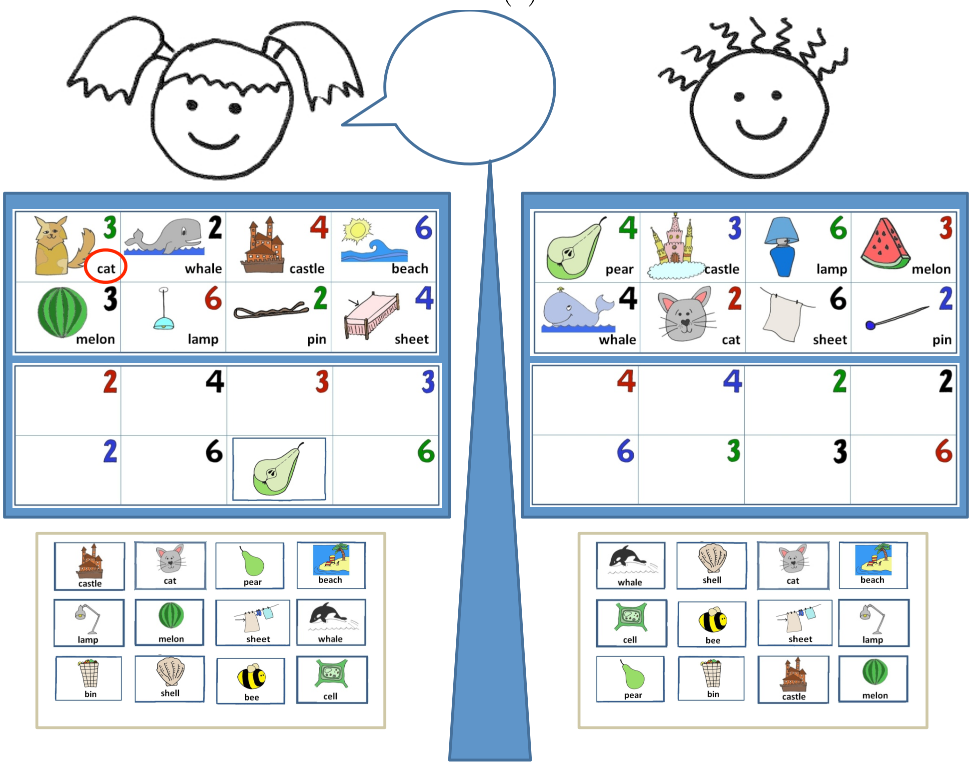



(b) 8

Figure B.4: Grid instructions: 7,8

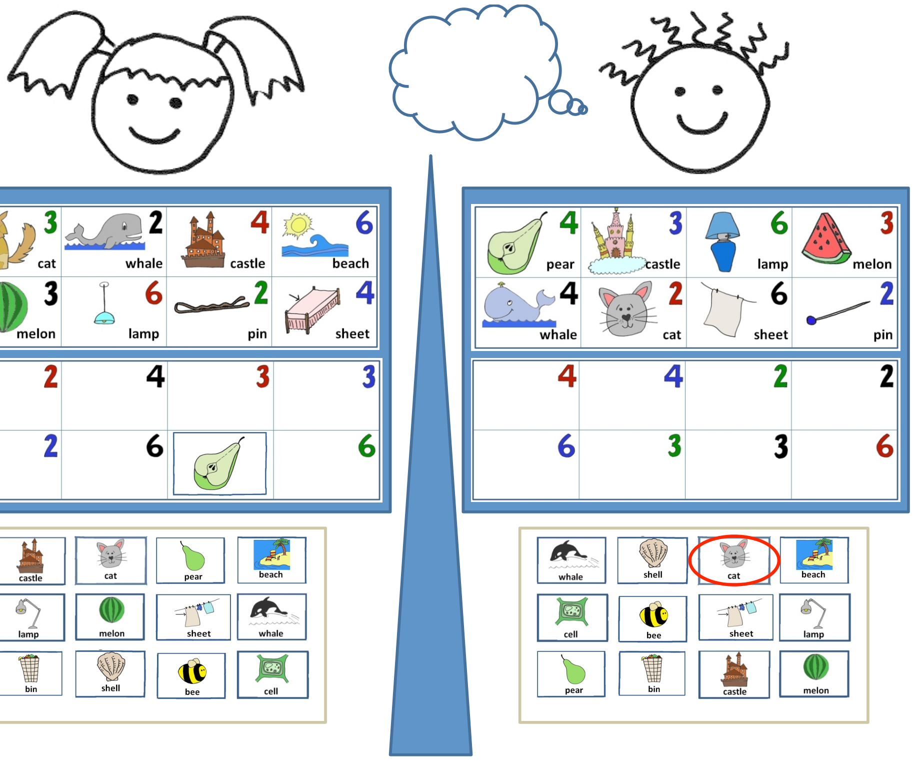



(a) 9

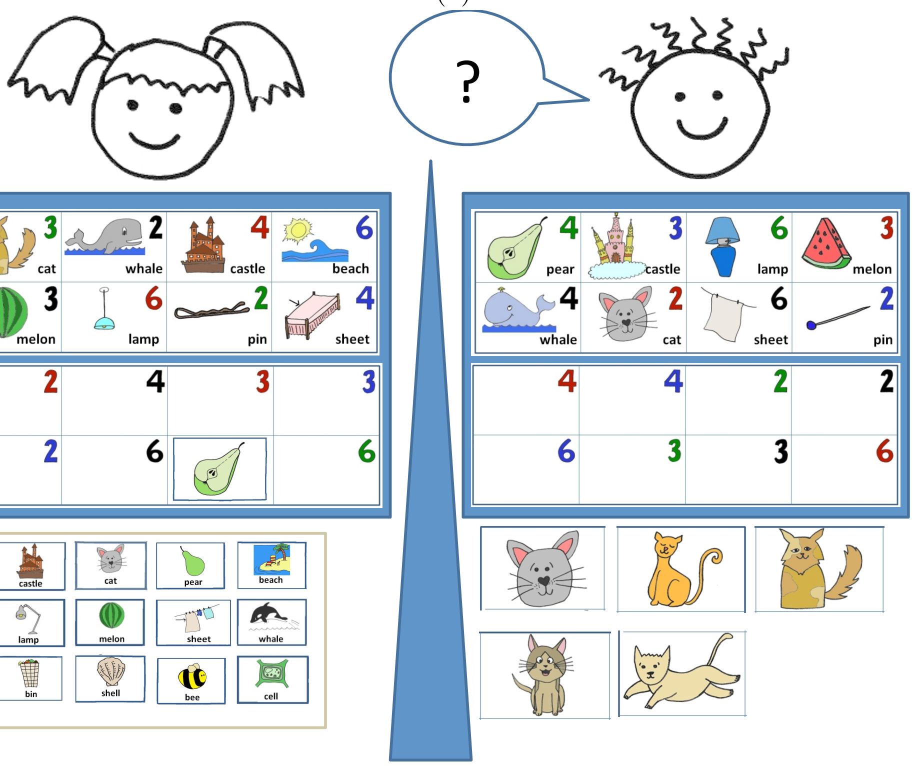



(b) 10

Figure B.5: Grid instructions: 9,10

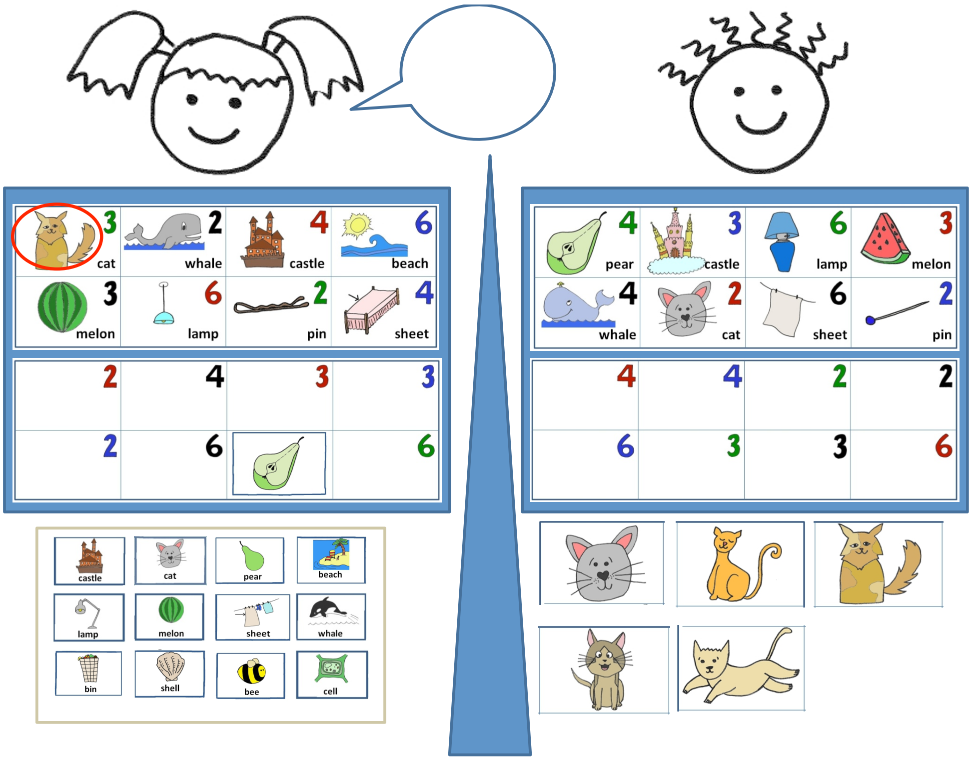



(a) 11

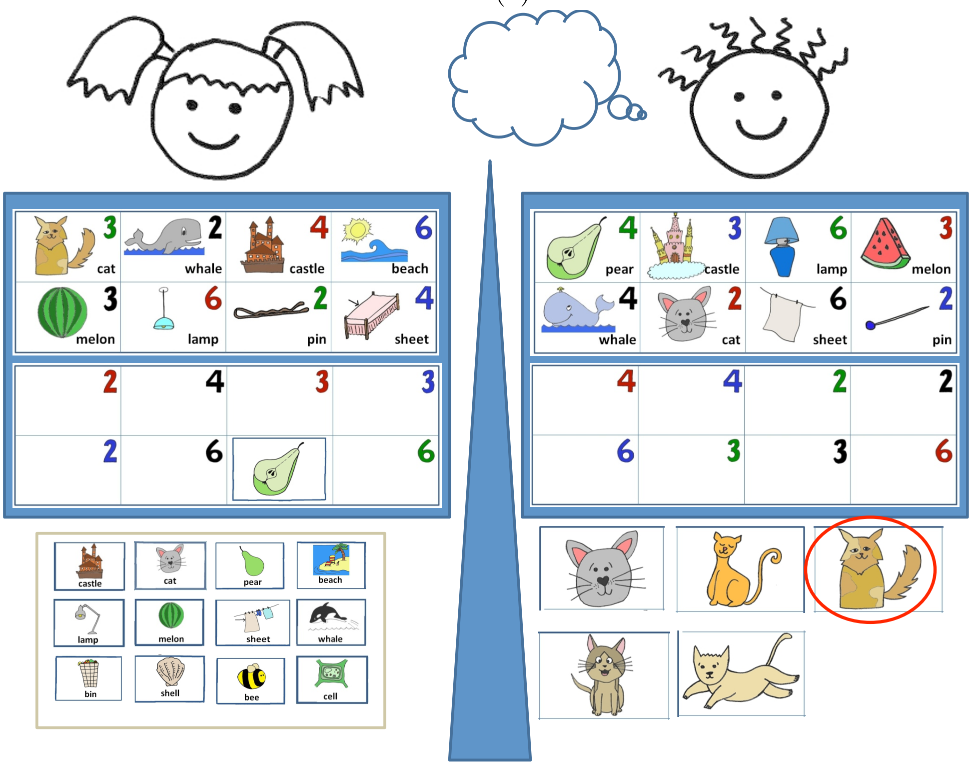



(b) 12

Figure B.6: Grid instructions: 11,12

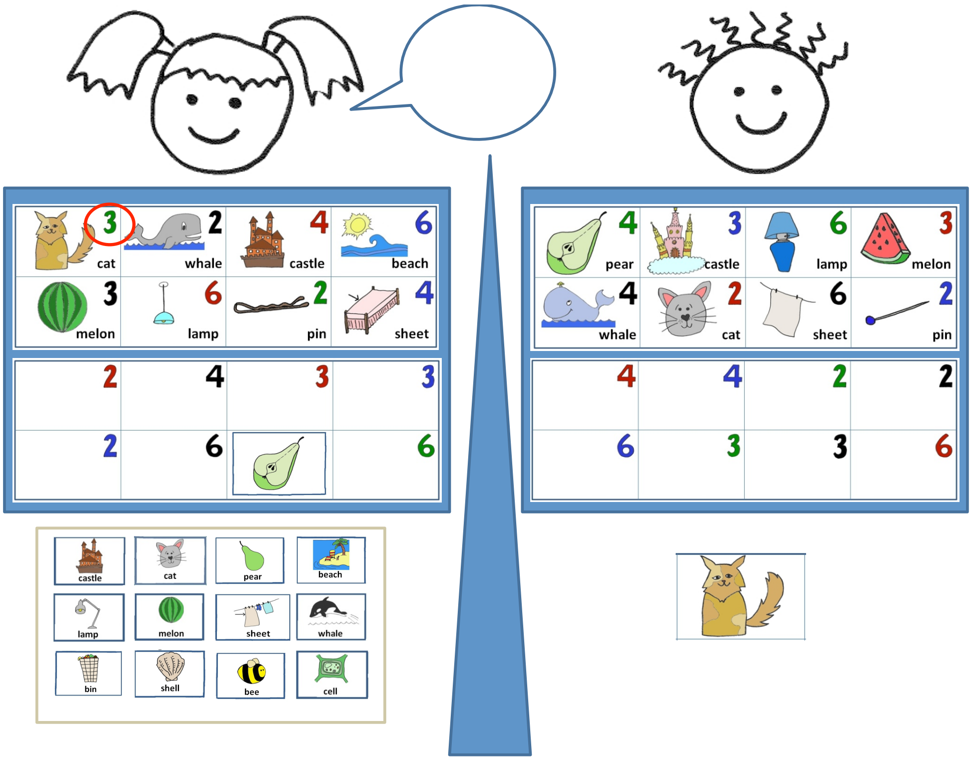



(a) 13

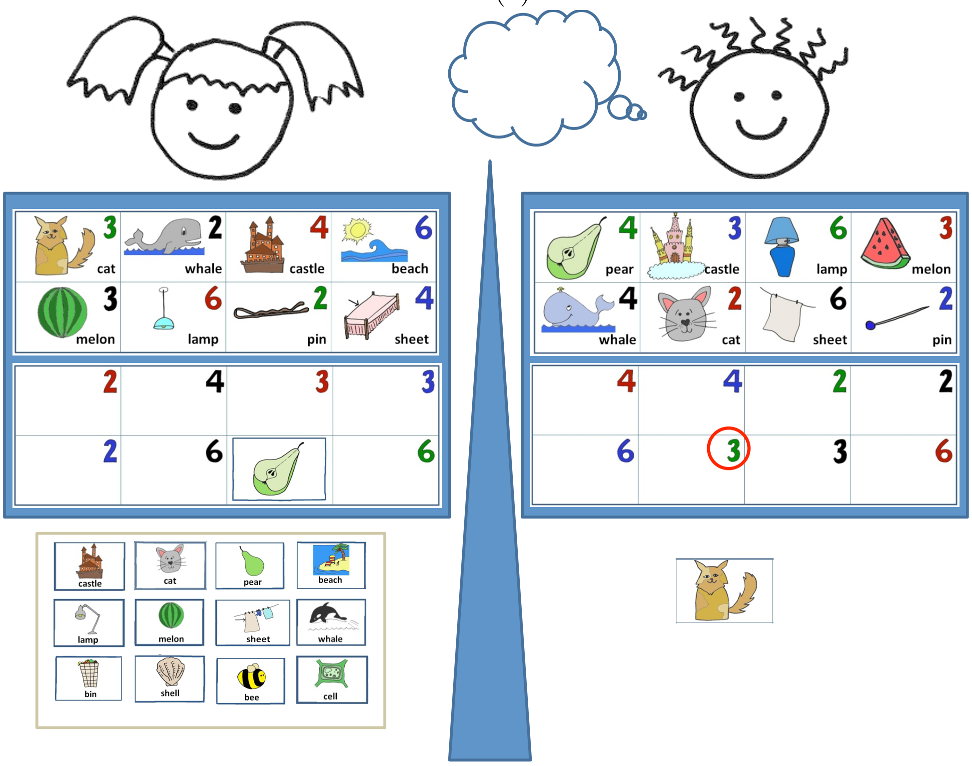



(b) 14

Figure B.7: Grid instructions: 13,14

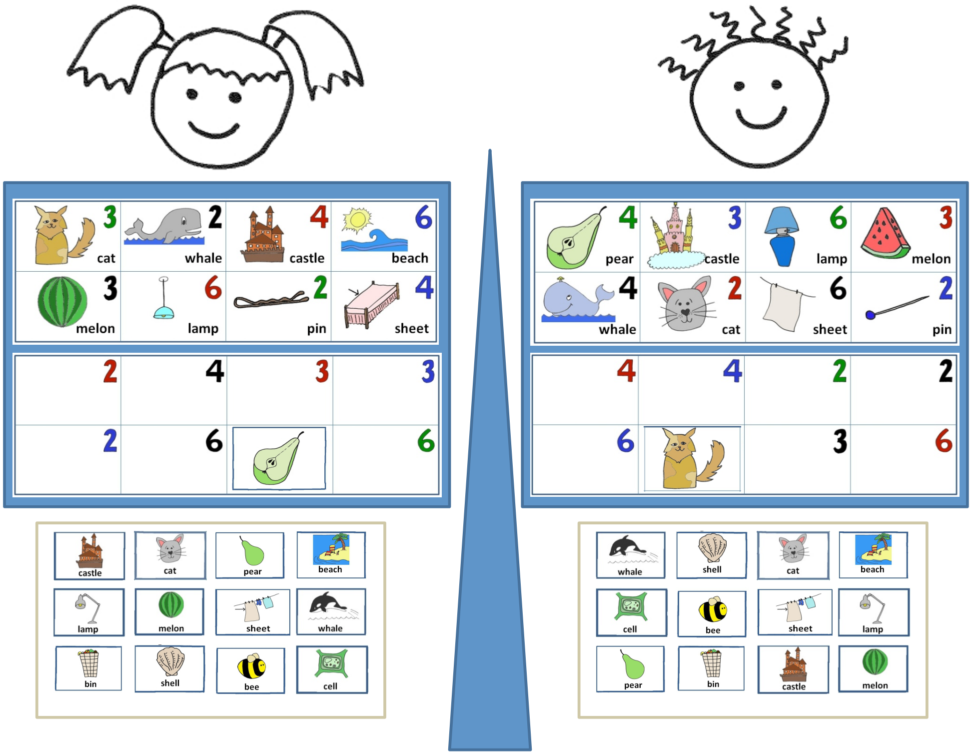



(a) 15

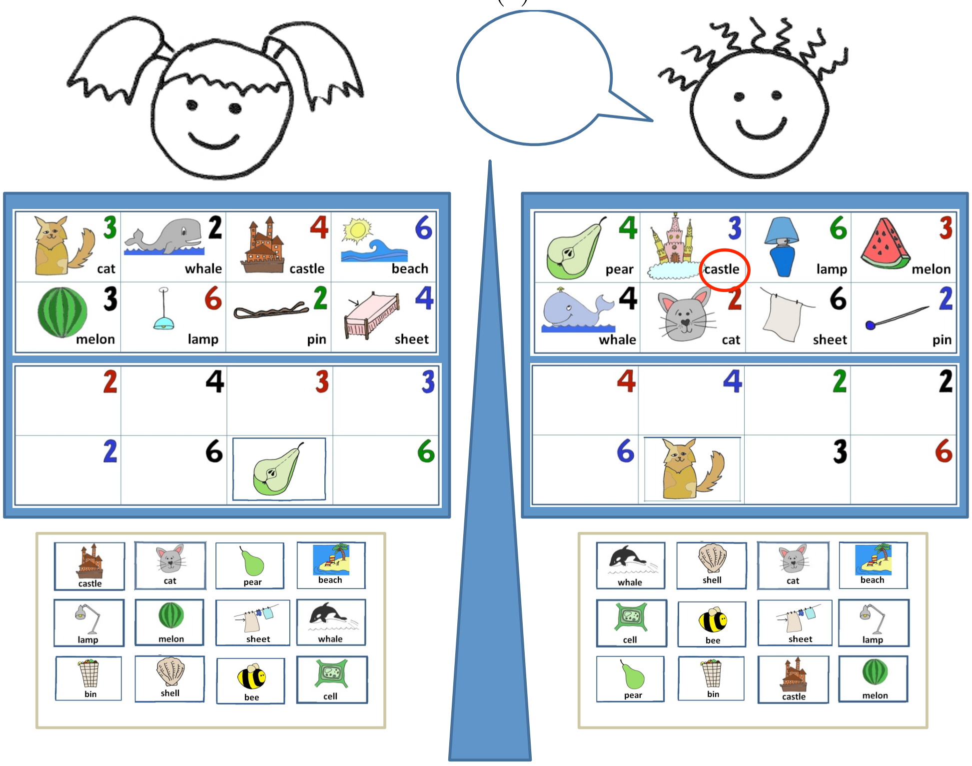



(b) 16

Figure B.8: Grid instructions: 15,16




(a) 17




(b) 18

Figure B.9: Grid instructions: 17,18

# Appendix C

# Parental questionnaire

RESEARCH PROJECT

Participant number:_______

**Title:** Speaker-controlled Variability in Children and Adolescents' Speech in Interaction

QUESTIONNAIRE FOR PARENTS/GUARDIANS

*In this study we will be doing recordings of your child's speech, and these recordings will then be analysed to see how speakers differ when talking in different communicative situations. To help with these analyses, we would like to ask you a few questions about your child.*

*This information will only be accessible to researchers working on the project and will only be used for the purposes of analysing the data. It will be stored in accordance with the Data Protection Act 1998.*

Your relation to the child: ___________

Your child's gender: male/female

His/her date of birth:___________

1. What is your child's first language? _______________________________

   Are any languages other than English used with your child? YES/NO

   If YES, please explain for each language who uses it with your child and how often

| **language** | **used with whom/in what situation** | **how often** |
|---|---|---|
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |

2. Does your child have hearing loss? YES/NO

   If NO, go to question 3.

   If YES, please answer questions a-d:

   a. How old was your child when s/he diagnosed with a hearing loss?

   ________________________________________________________

b. How old was your child when s/he was fitted with their first hearing aid?

___________________________________________________________

c. What type of hearing aid does your child have?

___________________________________________________________

d. Do any of your child's family members who live at home have a hearing impairment?

___________________________________________________________

3. Which schools has your child attended? Please name the schools and the years attended.

___________________________________________________________
___________________________________________________________
___________________________________________________________
___________________________________________________________
___________________________________________________________
___________________________________________________________

4. Has your child ever had any speech therapy?
   YES/NO
   If YES, please give brief details.

5. Does your child have any neurological, medical or learning difficulties? (e.g. attention deficit hyperactivity disorder, autism, cerebral palsy, speech/language difficulties, other)
   YES/NO
   If YES, please give brief details.

6. Does your child regularly interact with individuals who have a severe or profound hearing loss?
   YES/NO
   If YES, please give brief details.

7. How do you communicate with your child? (You can tick more than one box)

   - ☐ Spoken English
   - ☐ Spoken language other than English (Please specify) ____________
   - ☐ Speech and sign together
   - ☐ British Sign Language
   - ☐ Other_________________________

8. What is **your** main language?

   - ☐ Spoken English
   - ☐ Spoken language other than English (Please specify) ____________
   - ☐ Speech and sign together
   - ☐ British Sign Language
   - ☐ Other_________________________

9. How good is your child at understanding you?

☐ very good ☐ good ☐ fair ☐ poor

10. How does your child communicate with you mostly? (You can tick more than one box)

☐ Spoken English

☐ Spoken language other than English (Please specify) ____________

☐ Speech and sign together

☐ British Sign Language

☐ Other_________________________

11. Are there times when members of the family (including grandparents, aunts and uncles, cousins) cannot understand your child?

☐ very often ☐ quite often ☐ sometimes ☐ almost never

If 'very often' or 'quite often', please give examples if possible
_________________________________________________________________
_________________________________________________________________
_________________________________________________________________
_________________________________________________________________

12. Do you have any concern about your child's communication in general?

☐ none at all
☐ a little bit
☐ some
☐ a lot

Please comment if you wish.

# Appendix D

# Participant questionnaire

Participant number:_______

**Title:** Speaker-controlled Variability in Children and Adolescents' Speech in Interaction

## FEEDBACK QUESTIONNAIRE FOR PARTICIPANTS

*Thank you for taking part in this study. To help us with studying the recordings that we have made, we would like to ask you a few questions. This information will not have your name on it and it will be confidential.*

1. How well do you know your friend from **session 1**?

   On a scale of 'just know a little' to 'best friend'
   Please draw a line anywhere on the scale as your answer.

   |————————————————————|

   just know a little          friend          best friend

2. For how long have you known your friend from **session 1**?

   ☐ less than 6 months
   ☐ between 6 months and 1 year
   ☐ between 1 year and 3 years
   ☐ between 3 years and 5 years
   ☐ over 5 years

3. How well do you know your friend from **session 2**?

   On a scale of 'just know a little' to 'best friend'
   Please draw a line anywhere on the scale as your answer.

just know a little — friend — best friend

4. For how long have you known your friend from **session 2**?

   - ☐ less than 6 months
   - ☐ between 6 months and 1 year
   - ☐ between 1 year and 3 years
   - ☐ between 3 years and 5 years
   - ☐ over 5 years

5. How often during a typical week do you talk to people who are deaf?

   Please circle your answer.

   not at all    not very often    sometimes    often    all the time

6. How many of your friends are deaf?

   - ☐ all of them
   - ☐ most of them
   - ☐ some of them
   - ☐ a few of them
   - ☐ none of them

7. How do you usually communicate with your deaf friends?

   ☐ Spoken English

   ☐ Spoken language other than English

   ☐ Speech and sign together

   ☐ British Sign Language

   ☐ Other__________________________

8. Questions 8 a-c are for those participants who are deaf. Think about these situations, and answer the following questions.
   a. You are talking to someone in a quiet room.

      How much difficulty do you have in this situation?

      very difficult |————————————| easy

   b. You are talking to someone in a noisy room.

      How much difficulty do you have in this situation?

      very difficult |————————————| easy

   c. You are talking to a group of friends.

      How much difficulty do you have in this situation?

      very difficult |————————————| easy

# Appendix E

# Results from participant questionnaire

| pair type | familiarity | duration |
|---|---|---|
| HI-HI | 3.3 (1.3, 2-5) | 3.6 (1.1, 2-5) |
| HI-NH | 3.6 (1.3, 1-5) | 3.3 (1.1, 2-5) |
| NH-NH | 3.3 (1.5, 1-5) | 3.6 (1.7, 1-5) |
| mean | **3.4** | **3.5** |

Table E.1: Participant pairs' mean familiarity and duration of familiarity, with standard deviation and range in parentheses. In the individual session, each participant was asked 'How well do you know your friend?' [familiarity] (scale: 1-just a little, 3-friend, 5-best friend), and 'For how long have you known your friend?' [duration of familiarity] (scale: 1-less than 6mo, 2-between 6mo and 1yr, 3-between 1yr and 3yrs, 4-between 3yrs and 5yrs, 5-over 5yrs)

| part. | freq | prop | comm |
|---|---|---|---|
| NH1 | 4 | 2 | SP |
| NH2 | 5 | 3 | SP |
| NH3 | 4 | 3 | SP |
| NH4* | 4 | 3 | SP |
| NH5 | 4 | 2 | SP |
| NH6 | 4 | 2 | SP, G |
| NH7 | 5 | 2 | SP |
| NH8 | 5 | 2 | SP,SP+S |
| NH9 | 5 | 3 | SP+S |
| NH10 | 3 | 2 | SP, SP+S |
| NH11 | 5 | 5 | SP, G |
| NH12 | 5 | 2 | SP, BSL |
| NH13 | 3 | 2 | SP |
| NH14 | 2 | 3 | SP |
| NH15 | 5 | 2 | SP |
| NH16 | 4 | 3 | SP, SP+S |
| NH17 | 4 | 2 | SP, SP+S |
| NH18 | 5 | 2 | SP+S |
| **mean** | **4.2** | **2.5** | |

| part. | freq | prop | comm |
|---|---|---|---|
| HI1 | 1 | 1 | SP, BSL |
| HI2 | 2 | 2 | SP, BSL |
| HI3 | 5 | 4 | SP, SP+S |
| HI4* | 4 | 5 | SP |
| HI5 | 3 | 4 | SP, SP+S, BSL |
| HI6 | 5 | 3 | SP+S |
| HI7 | 5 | 3 | SP+S |
| HI8 | 5 | 4 | SP, SP+S |
| HI9 | 5 | 3 | SP, SP+S |
| HI10 | 5 | 3 | SP, BSL |
| HI11 | 5 | 3 | SP |
| HI12 | 3 | 1 | SP |
| HI13 | 2 | 3 | SP |
| HI14 | 3 | 4 | SP |
| HI15 | 3 | 3 | SP |
| HI16 | 5 | 4 | SP, SP+S |
| HI17 | 3 | 3 | SP |
| HI18 | 3 | 3 | SP, SP+S, BSL |
| **mean** | **3.7** | **3.1** | |

Table E.2: Results from participant questionnaire (see Appendix D). Part.=participant. Freq=frequency of contact with HI peers, from question: ‘How often do you talk to people who are deaf?’ (range: 1-not at all, 2-not very often, 3-sometimes, 4-often, 5-all the time). Prop=proportion of HI friends, from question: ‘How many of your friends are deaf?’ (range: 1-none of them, 2-a few of them, 3-some of them, 4-most of them, 5-all of them). Comm=usual communication mode with HI friends, from question: ‘How do you usually communicate with your deaf friends?’ (SP: spoken English, SP+S: speech and sign together, BSL: British Sign Language, G: gesture). The shaded areas between NH and HI participants indicate the group of four within which each child participated in the study.

# Appendix F

# Transcription protocol

This transcription protocol was taken from kidLUCID (2015), with slight modifications made for the needs of the current study. The guidelines are originally based on those used by Van Engen *et al.* (2010), with minor adaptations.

The speech was transcribed verbatim, and no punctuation was used, except for apostrophes for contractions and possessives. Numbers were written out. No hyphenation or abbreviation was used, and full dictionary spellings were used for all words except for those mentioned in below.

Collocations and fixed spellings of certain words were used. All hesitation sounds, filled pauses and agreement words were transcribed (such as ‘uh’, ‘err’, ‘yeah’, ‘mmhmm’). When a speaker said a sequence of letters or was spelling a word, the letter sequences are spelled out in capital letters and separate letters by spaces, such as ‘U C L’.

Other symbols used:

**SIL**
: Within-speaker pause of a minimum duration of 0.5 seconds

**SILP**
: Silence by current speaker when the interlocutor is talking

**word-**
: Word is spoken partially (even for unknown words)

**<LG>**
Laughter that is not part of any word

**<BR>**
Breath, sighs

**<LS>**
Lip smack

**<GA>**
Garbage: noise that is not from the speaker, such as microphone pops and background noise.

**<UN>**
Speaker produces a word which is unintelligible to the transcriber

**<LG_word>**
Laughter that a speaker produced while saying a word

**<GA_word>**
Noise that occurs while the speaker is saying a word

**<SIM_word>**
Speech that is spoken in overlap with the interlocutor

**<UN_word>**
The transcriber has attempted to transcribe a word, but is unsure of which word is intended

**<WH_word>**
The speaker whispers a word

# Appendix G

# Reported strategy use by participants

After each communication session, except for the sessions with participants NH13, NH14, HI13 and HI14, the researcher asked the two participants questions on task difficulty, as well as ‘What do you do if your friend can’t understand you?’. Tables G.1, G.2, G.3, G.4, G.5 and G.6 give the different answers provided by the participants, divided into sign language, acoustic-phonetic, linguistic, repetition, spelling, gestural/visual strategies.

| **speaker** | **condition** | **sign language strategy** |
|---|---|---|
| HI7 | HID | I would sign |
| HI6 | NHD, HID | for HI5 I’d use sign language |
| HI9 | NHD | just sign them |
| NH1 | HID | [...] or like sign it out like words or letters |
| NH10 | NHD | I’ll just sign or spell |
| NH9 | NHD, HID | [...] if they’re deaf I would sign [...] |

Table G.1: Sign language strategies reported by the participants.

| speaker | condition | acoustic-phonetic strategy |
|---|---|---|
| HI17 | NHD | umm I'd say it clearly |
| HI8 | HID | I'd do it really like slowly but loudly like FOUR RED |
| NH16 | NHD | repeat it more clearly than the first time |
| NH16 | HID | you have to- you repeat it for them so they understand it and say it a lot clearer this time |
| NH17 | HID | say it more clearly and do hand actions with it |
| NH18 | HID | if it was like shell I would say umm I would like sound it out so like [ʃ] [ɛ] [l], ask if that would make it easier |
| NH9 | HID | [...] I'd probably say it really slowly and clearly but if they're deaf I might sign |
| NH10 | HID | [...] sometimes we need to sound it out for each other so we can make it easier |

Table G.2: Acoustic-phonetic strategies reported by the participants.

| speaker | condition | linguistic strategy |
|---|---|---|
| HI2 | NHD | like you see shack and then sack, when you describe it we would say what do you mean a bag or like a shed |
| HI4 | NHD | we would try to umm tell them how does it look like and [...] |
| NH12 | HID | we helped them like [...] we gave them more information |
| NH15 | NHD | describe it |
| NH17 | NHD | explain it in more child language like words that they would use instead of bigger words |
| NH18 | NHD | try and explain it like more clearly |
| NH4 | NHD | yeah you would explain it but describe it in a more easier way so that they can understand |
| NH4 | HID | describe it more better, say more the colours exactly what it looks like |
| NH5 | NHD | describe it more, go into details |

Table G.3: Linguistic strategies reported by the participants.

| speaker | condition | repetition strategy |
|---|---|---|
| HI1 | NHD | you would repeat it wouldn't you |
| HI12 | NHD | we just kept on saying it and things like that and then they just knew it |
| HI12 | HID | I keep saying it |
| HI15 | HID | say it again |
| HI3 | HID | I'll just say it to them again or- [...] |
| HI3 | NHD | [...] but I do explain it to my friend so yeah so explain it over again |
| HI7 | HID | I would just go over and over and say it until he gets it |
| NH1 | HID | sometimes when she wouldn't understand me at all like [...] I can do it loads of times or [...] |
| NH1 | NHD | I would just like say it again |
| NH16 | NHD | repeat it more clearly than you did the first time |
| NH16 | HID | you have to- you repeat it for them so they understand it and say it a lot clearer this time |
| NH3 | NHD, HID | explain it again |
| NH8 | HID | I'd probably like say it two times and spell it if they couldn't understand me |

Table G.4: Repetition strategies reported by the participants.

| speaker | condition | spelling strategy |
|---|---|---|
| HI18 | NHD | [...] and spell it as well |
| HI7 | NHD | well, I would spell it out like shack, S H A C K |
| HI8 | NHD | spell it out |
| HI9 | NHD | [...] and I just show picture different *[points to tray]* and say S C A K, like that |
| NH10 | NHD | I'll just sign or spell |
| NH18 | HID | [...] I would spell it out or [...] |
| NH2 | HID | and we were spelling it out |
| NH9 | NHD | if they're not [deaf] then I would spell it out for them |
| NH9 | HID | I would spell it [...] |

Table G.5: Spelling strategies reported by the participants.

| **speaker** | **condition** | **gestural/visual strategy** |
|---|---|---|
| HI3 | HID | usually I would use sign language even though I don't know [sign language], so I'll go I *[points to self]* have *[gestures with both hands]* to go *[uses a 'away' gesture]* dinner *[gestures eating]* |
| HI4 | HID | if they don't understand what I said like sheet I'd go like *[gestures a sheet on a bed with hands flat down]* |
| HI4 | NHD | we would [...] tell the shape *[uses gestures]* how does it- yeah |
| NH10 | HID | we would like act it out or we could do like some pictures where we use our hands or sign language |
| NH17 | HID | say it more clearly and do hand actions with it [...] like if it was a bee or pea you could do that *[gestures a sign language 'P']* or with little wings on |
| NH2 | HID | I gave him a shape with my hands |
| NH2 | NHD | shaped it out with my fingers |
| NH5 | HID | well for like numbers we used our hands and like we just like with beans we'd done eating *[gestures eating]* things like that so it would be easier |
| NH5 | NHD | because we speak like that *[looks at floor]* sometimes and like they struggle looking at our lips |
| NH6 | HID | I have to s- to be looking at her to talk because she won't- she might not pay attention otherwise [...] I just have to make sure she's paying attention when you're talking [...] with HI5, I kind of said last week that we gesture and we do gesture with HI6 sometimes [...] |
| NH6 | NHD | we don't know exactly very well sign language so we just like [...] if we're saying walking do that *[mimes walking]* or like meet is like meet *[gestures 'meet']* me *[points at self]* at IT *[gestures]* or like art *[mimics drawing]* [...] |

Table G.6: Gestural/visual strategies reported by the participants.

# References

Action on Hearing Loss (**2011**). "Hearing Matters", URL `http://www.actiononhearingloss.org.uk/supporting-you/policy-research-and-influencing/research/hearing-matters.aspx`. 4

Action on Hearing Loss (**2012**). *Miscommunication? Tips when speaking to someone with a hearing loss.*, London, UK, leaflet. 30, 31, 32, 33, 136, 238, 239

Adank, P., Smits, R., and van Hout, R. (**2004**). "A comparison of vowel normalization procedures for language variation research", The Journal of the Acoustical Society of America **116**, 3099–3107. 144

Aitken, K. J. (**2013**). "It ain't what you do (it's the way that you do it).", Behavioral and Brain Sciences **36**, 363–364. 243

Allen, G. D. and Arndorfer, P. M. (**2000**). "Production of sentence-final intonation contours by hearing-impaired children", Journal of Speech Language and Hearing Research **43**, 441–455. 21, 22, 23, 66, 124

Allen, P. and Wightman, F. (**1992**). "Spectral pattern discrimination by children.", Journal of Speech and Hearing Research **35**, 222–233. 8

Anderson, A., Bader, M., Bard, E., Boyle, E., Doherty, G. M., Garrod, S. C., Isard, S. D., Kowtko, J. C., McAllister, J. M., Miller, J., Sotillo, C. F., Thompson, H. S., and Weinert, R. (**1991**). "The HCRC map task corpus", Language and Speech **34**, 351–366. 76, 137, 195

Anselmi, D., Tomasello, M., and Acunzo, M. (**1986**). "Young children's responses to neutral and specific contingent queries.", Journal of Child Language **13**, 135–144. 52

Antia, S. D. (**1982**). "Social Interaction of Partially Mainstreamed Hearing-Impaired Children", American Annals of the Deaf **127**, 18–25. 56

Antia, S. D., Reed, S., and Shaw, L. (**2011**). "Risk and Resilience for Social Competence: Deaf Students in General Education Classrooms", in *Resilience in Deaf Children: Adaptation Through Emerging Adulthood*, 139–167 (Springer New York, New York, NY). 2, 56, 246

Arnold, D. and Tremblay, A. (**1979**). "Interaction of deaf and hearing preschool children", Journal of Communication Disorders **12**, 245–251. 56, 57, 67

Arnold, P., Palmer, C., and Lloyd, J. (**1999**). "Hearing-impaired children's listening skills in a referential communication task: an exploratory study", Deafness & Education International **1**, 47–55. 65, 76

Asher, S. R. and Wigfield, A. (**1980**). "Training referential communication skills", in *Children's Oral Communication Skills*, edited by W. P. Dickson, 1–60 (Academic Press, New York). 47

Aslin, R. and Smith, L. (**1988**). "Perceptual development.", Annual Review of Psychology **39**, 435–473. 7

Aubanel, V. and Cooke, M. (**2013**). "Strategies adopted by talkers faced with fluctuating and competing-speech maskers", The Journal of the Acoustical Society of America **134**, 2884–2894. 37, 42, 241

Aubanel, V., Cooke, M., Foster, E., García Lecumberri, M. L., and Mayo, C. (**2012**). "Effects of the availability of visual information and presence of competing conversations on speech production", in *Interspeech 2012*, 2033–2036. 195

Bajaj, Y., Sirimanna, T., Albert, D. M., Qadir, P., Jenkins, L., Cortina-Borja, M., and Bitner-Glindzicz, M. (**2009**). "Causes of deafness in British Bangladeshi children: a prevalence twice that of the UK population cannot be accounted for by consanguinity alone.", Clinical Otolaryngology **34**, 113–119. 3

Baker, R. and Hazan, V. (**2011**). "DiapixUK: task materials for the elicitation of multiple spontaneous speech dialogs", Behavior Research Methods **43**, 761–770. 73, 77, 93, 237

Bamford, J., Fortnum, H., Bristow, K., Smith, J., Davies, L., Taylor, R., Watkin, P., Fonseca, S., Davis, A., and Hind, S. (**2007**). "Current practice, accuracy, effectiveness and cost-effectiveness of the school entry hearing screen.", Health Technology Assessment **11**, 1–168. 3

Bamford, J., Uus, K., and Davis, A. (**2005**). "Screening for hearing loss in childhood: issues, evidence and current approaches in the UK", Journal of Medical Screening **12**, 119–124. 3, 4

Bard, E., Anderson, A., Sotillo, C., Aylett, M., Doherty-Sneddon, G., and Newlands, A. (**2000**). "Controlling the intelligibility of referring expressions in dialogue.", Journal of Memory and Language **42**, 1–22. 43, 124, 141, 177, 243

Bard, E. G. and Aylett, M. P. (**2005**). "Referential form, word duration, and modeling the listener in spoken dialogue (chapter 7).", in *Approaches to Studying World-Situated Language Use*, edited by J. Trueswell and M. Tanenhaus, 173–191 (MIT Press, Cambridge, MA). 42, 43, 46, 71, 177, 178, 241, 243

Baron-Cohen, S., Leslie, A. M., and Frith, U. (**1985**). "Does the autistic child have a "theory of mind"?", Cognition **21**, 37–46. 48

Bat-Chava, Y. and Deignan, E. (**2001**). "Peer relationships of children with cochlear implants", Journal of Deaf Studies and Deaf Education **6**, 186–199. 6

Batten, G., Oakes, P. M., and Alexander, T. (**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**, 285–302. 2, 246

Baudonck, N. L. H., Buekers, R., Gillebert, S., and Van Lierde, K. M. (**2009**). "Speech Intelligibility of Flemish Children as Judged by Their Parents", Folia Phoniatrica et Logopaedica **61**, 288–295. 19

Beadle, E. A. R., McKinley, D. J., Nikolopoulos, T. P., Brough, J., O'Donoghue, G. M., and Archbold, S. M. (**2005**). "Long-term functional outcomes and academic-occupational status in implanted children after 10 to 14 years of cochlear implant use", Otology & neurotology **26**, 1152–1160. 11

Beer, J., Kronenberger, W. G., Castellanos, I., Colson, B. G., Henning, S. C., and Pisoni, D. B. (**2014**). "Executive Functioning Skills in Preschool-Age Children With Cochlear Implants", Journal of Speech Language and Hearing Research **57**, 1521–1534. 62

Beer, J., Pisoni, D. B., Kronenberger, W., and Geers, A. (**2010**). "New research findings: Executive functions of adolescents who use cochlear implants.", ASHA Leader **15**, 12–15. 62

Bench, J., Kowal, A., and Bamford, J. (**1979**). "The BKB (Bamford-Kowal-Bench) sentence list for partially-hearing children.", British Journal of Audiology **13**, 108–112. 98

Berger, K. W. (**1972**). *Speechreading – principles and methods.* (National Educational Press, Inc., Maryland, US). 137

Bergeson, T.R., M. R. and McCune, K. (**2006**). "Mothers' speech to hearing-impaired infants and children with cochlear implants.", Infancy **10**, 212–240. 41

Bergeson, T. R., Houston, D. M., and Miyamoto, R. T. (**2010**). "Effects of congenital hearing loss and cochlear implantation on audiovisual speech perception in infants and children", Restorative neurology and neuroscience **28**, 157–165. 10

Bernhardt, B., Gick, B., Bacsfalvi, P., and Ashdown, J. (**2003**). "Speech habilitation of hard of hearing adolescents using electropalatography and ultrasound as evaluated by trained listeners", Clinical Linguistics & Phonetics **17**, 199–216. 24, 74

Birnholz, J. C. and Benacerraf, B. R. (**1983**). "The development of human fetal hearing", Science **222**, 516–518. 7

Blamey, P. and Sarant, J. (**2011**). "Development of spoken language by deaf children (chapter 17).", in *The Oxford Handbook of Deaf Studies, Language, and Education*, edited by M. Marschark and P. Spencer, 241–257 (Oxford University Press, Oxford, UK). 9

Blamey, P. J., Sarant, J. Z., Paatsch, L. E., Barry, J. G., Bow, C. P., Wales, R. J., Wright, M., Psarros, C., Rattigan, K., and 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**, 264–285. 11, 14, 21, 75, 183, 195, 196

Blaylock, R. L., Scudder, R. R., and Wynne, M. K. (**1995**). "Repair behaviors used by children with hearing loss", Language, Speech, and Hearing Services in Schools **26**, 278–285. 64

Boersma, P. and Weenink, D. (**2015**). *Praat: doing phonetics by computer*, URL `http://www.praat.org`. 94

Bond, Z. S. and Moore, T. J. (**1994**). "A note on the acoustic-phonetic characteristics of inadvertently clear speech", Speech Communication **14**, 325–337. 31, 32, 33

Boothroyd, A. (**1984**). "Auditory perception of speech contrasts by subjects with sensorineural hearing loss.", Journal of Speech and Hearing Research **27**, 185–196. 12

Borg, E., Edquist, G., Reinholdson, A.-C., Risberg, A., and McAllister, B. (**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**, 1061–1077. 12, 13, 105, 179

Bouton, S., Serniclaes, W., Bertoncini, J., and Cole, P. (**2012**). "Perception of Speech Features by French-Speaking Children With Cochlear Implants", Journal of Speech Language and Hearing Research **55**, 139–153. 14

Boyle, E. A., Anderson, A. H., and Newlands, A. (**1994**). "The effects of visibility on dialogue and performance in a cooperative problem solving task", Language and Speech **37**, 1–20. 183

Bradlow, A. R. (**2002**). "Confluent talker-and listener-oriented forces in clear speech production", Laboratory Phonology **7**, 241–273. 35, 44, 142, 176

Bradlow, A. R. and Alexander, J. A. (**2007**). "Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners", The Journal of the Acoustical Society of America **121**, 2339–2349. 35

Bradlow, A. R. and Bent, T. (**2002**). "The clear speech effect for non-native listeners", The Journal of the Acoustical Society of America **112**, 272–284. 35

Bradlow, A. R., Kraus, N., and Hayes, E. (**2003**). "Speaking Clearly for Children With Learning Disabilities: Sentence Perception in Noise", Journal of Speech Language and Hearing Research **46**, 80–97. 35, 44

Bradlow, A. R., Torretta, G. M., and Pisoni, D. B. (**1996**). "Intelligibility of normal speech I: global and fine-grained acousticphonetic talker characteristics.", Speech Communication **20**, 255–272. 31, 32

Breland, H. M. (**1996**). "Word frequency and word difficulty: a comparison of counts in four corpora", Psychological Science **7**, 96–99. 182

Brinton, B., Fujiki, M., Loeb, D. F., and Winkler, E. (**1986**). "Development of conversational repair strategies in response to requests for clarification", Journal of Speech Language and Hearing Research **29**, 75–81. 51, 53, 56, 63, 194

Briscoe, J., Bishop, D., and Norbury, C. (**2001**). "Phonological processing, language, and literacy: a comparison of children with mild-to-moderate sensorineural hearing loss and those with specific language impairment.", Journal of Child Psychology and Psychiatry **3**, 329–340. 13, 14, 15

Brown, G., Anderson, A., Yule, G., and Shillcock, R. (**1983**). *Teaching Talk* (Cambridge University Press, Cambridge, UK). 76

Brown, P. and Dell, G. (**1987**). "Adapting production to comprehension – the explicit mention of instruments.", Cognitive Psychology **19**, 441–472. 43, 217, 241

Brysbaert, M. and New, B. (**2009**). "Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English", Behavior Research Methods **41**, 977–990. 184

Burkholder, R. A. and Pisoni, D. B. (**2003**). "Speech timing and working memory in profoundly deaf children after cochlear implantation", Journal of Experimental Child Psychology **85**, 63–88. 23, 28, 139, 214, 237

Burnham, D., Kitamura, C., and Vollmer-Conna, U. (**2002**). "What's new, pussycat? On talking to babies and animals", Science **296**, 1435–1435. 39, 40

Caissie, R., Dawe, A. L., Donovan, C., Brooks, H., and MacDonald, S. (**1998**). "Conversational performance of adults with hearing loss", Journal of the Academy of Rehabilitative Audiology **31**, 45–67. 196

Caissie, R. and Gibson, C. L. (**1997**). "The effectiveness of repair strategies used by people with hearing losses and their conversational partners.", Volta Review **99**, 203–218. 42

Caissie, R. and Tranquilla, M. (**2010**). "Enhancing Conversational Fluency: Training Conversation Partners in the Use of Clear Speech and Other Strategies", Seminars in Hearing **31**, 95–103. 30, 31, 32, 238

Caissie, R. and Wilson, E. (**1995**). "Communication breakdown management during cooperative learning activities by mainstreamed students with hearing losses.", Volta Review **97**, 105–121. 62

Campisi, P., Low, A., Papsin, B., Mount, R., Cohen-Kerem, R., and Harrison, R. (**2005**). "Acoustic Analysis of the Voice in Pediatric Cochlear Implant Recipients: A Longitudinal Study", The Laryngoscope **115**, 1046–1050. 22, 66, 138

Campisi, P., Low, A. J., Papsin, B. C., Mount, R. J., and Harrison, R. V. (**2006**). "Multidimensional voice program analysis in profoundly deaf children: Quantifying frequency and amplitude control", Perceptual and motor skills **103**, 40–50. 138

Charles-Luce, J. (**1997**). "Cognitive factors involved in preserving a phonemic contrast.", Language and Speech **40**, 229–248. 37

Chatterjee, M. and Peng, S.-C. (**2008**). "Processing F0 with cochlear implants: Modulation frequency discrimination and speech intonation recognition", Hearing Research **235**, 143–156. 12

Cheng, H. Y., Murdoch, B. E., and Goozée, J. V. (**2007**a). "Temporal features of articulation from childhood to adolescence: An electropalatographic investigation", Clinical Linguistics & Phonetics **21**, 481–499. 17

Cheng, H. Y., Murdoch, B. E., Goozée, J. V., and Scott, D. (**2007**b). "Electropalatographic assessment of tongue-to-palate contact patterns and variability in children, adolescents, and adults", Journal of Speech Language and Hearing Research **50**, 375–392. 17

Cheng, H. Y., Murdoch, B. E., Goozée, J. V., and Scott, D. (**2007**c). "Physiologic development of tongue–jaw coordination from childhood to adulthood", Journal of Speech Language and Hearing Research **50**, 352–360. 17

Chin, S. B., Bergeson, T. R., and Phan, J. (**2012**). "Speech intelligibility and prosody production in children with cochlear implants", Journal of Communication Disorders **45**, 355–366. 22

Chu, J. X. H. (**2015**). "Breakdown and repair in conversations between pairs of hearing-impaired and normally-hearing children aged 11 to 13 years.", Undergraduate thesis, University College London (UCL). 90, 196, 246

Chuang, H.-F., Yang, C.-C., Chi, L.-Y., Weismer, G., and Wang, Y.-T. (**2012**). "Speech intelligibility, speaking rate, and vowel formant characteristics in Mandarin-speaking children with cochlear implant", International Journal of Speech-Language Pathology **14**, 119–129. 23, 26, 74, 139, 179

Ciocca, V., Francis, A. L., Aisha, R., and Wong, L. (**2002**). "The perception of Cantonese lexical tones by early-deafened cochlear implantees", The Journal of the Acoustical Society of America **111**, 2250–2256. 12

Ciocci, S. R. and Baran, J. A. (**1998**). "The use of conversational repair strategies by children who are deaf", American Annals of the Deaf **143**, 235–245. 63, 196, 242

Clark, H. H. and Wilkes-Gibbs, D. (**1986**). "Referring as a Collaborative Process.", Cognition **22**, 1–39. 245

Connor, C. M., Craig, H. K., Raudenbush, S. W., Heavner, K., and Zwolan, T. A. (**2006**). "The age at which young deaf children receive cochlear implants and their vocabulary and speech-production growth: is there an added value for early implantation?", Ear and Hearing **27**, 628–644. 20, 21

Conway, C. M., Deocampo, J. A., Walk, A. M., Anaya, E. M., and Pisoni, D. B. (**2014**). "Deaf Children With Cochlear Implants Do Not Appear to Use Sentence Context to Help Recognize Spoken Words", Journal of Speech Language and Hearing Research **57**, 2174–2190. 15

Cooke, M., King, S., Garnier, M., and Aubanel, V. (**2014**a). "The listening talker: A review of human and algorithmic context-induced modifications of speech", Computer Speech & Language **28**, 543–571. 30, 31, 34, 37, 38, 123

Cooke, M. and Lu, Y. (**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**, 2059–2069. 37, 42, 236

Cooke, M., Mayo, C., and Villegas, J. (**2014**b). "The contribution of durational and spectral changes to the Lombard speech intelligibility benefit", The Journal of the Acoustical Society of America **135**, 874–883. 137

Costa, A., Pickering, M. J., and Sorace, A. (**2008**). "Alignment in second language dialogue", Language and Cognitive Processes **23**, 528–556. 38, 196

Courtin, C., Melot, A.-M., and Corroyer, D. (**2008**). "Achieving efficient learning : Why understanding theory of mind is essential for deaf children... and their teachers.", in *Deaf cognition, foundations and outcomes.*, edited by M. Marschark and P. Hauser (Oxford University Press, New York). 48

Crawley, M. J. (**2007**). *The R Book* (Wiley, Chichester). 126

CRIDE (**2013**). "CRIDE report on 2013 survey on educational provision for deaf children in England", URL `http://www.ndcs.org.uk/document.rm?id=9499`. 3, 4, 5, 6

CRIDE (**2014**). "CRIDE report on 2014 survey on educational provision for deaf children in England", URL `http://www.ndcs.org.uk/document.rm?id=9796`. 1, 3, 6, 236

Cristia, A. (**2010**). "Phonetic enhancement of sibilants in infant-directed speech.", Journal of the Acoustical Society of America **128**, 424–434. 39, 45

Cristia, A. and Seidl, A. (**2014**). "The hyperarticulation hypothesis of infant-directed speech* – CORRIGENDUM", Journal of Child Language **41**, 913–934. 39, 45, 176

Dammeyer, J. (**2012**). "A Longitudinal Study of Pragmatic Language Development in Three Children with Cochlear Implants", Deafness & Education International **14**, 217–232. 62

Davis, A., Bamford, J., Wilson, I., Ramkalawan, T., Forshaw, M., and Wright, S. (**1997**). "A critical review of the role of neonatal hearing screening in the detection of congenital hearing impairment", Health Technology Assessment **1**, 1–76. 4

Davis, A., Davis, K., and Mencher, G. (**2009**). "Epidemiology of permanent childhood hearing impairment", in *Paediatric Audiological Medicine.*, edited by V. E. Newton (Wiley-Blackwell). 3, 4 de Boysson-Bardies, B. and Vihman, M. M. (**1991**). “Adaptation to language: Evidence from babbling and first words in four languages”, Language 297–319. 16

Dell, G. and Brown, P. (**1991**). “Mechanisms for listener-adaptation in language production: Limiting the role of the ”model of the listener”.”, in *Bridges between Psychology and Linguistics*, edited by D. Napoli and J. Kegl (Erlbaum, Hillsdale, NJ). 43

Deutsch, W. and Pechmann, T. (**1982**). “Social interaction and the development of definite descriptions”, Cognition **11**, 159–184. 54

Diamond, K. E. and Hestenes, L. L. (**1996**). “Preschool Children’s Conceptions of Disabilities The Salience of Disability in Children’s Ideas About Others”, Topics in Early Childhood Special Education **16**, 458–475. 50

Dodd, B., Hua, Z., Crosbie, S., Holm, A., and Ozanne, A. (**2002**). *Diagnostic Evaluation of Articulation and Phonology.* (Harcourt Assessment, London). 16

Dorman, M. F., Loizou, P. C., Spahr, A. J., and Maloff, E. (**2002**). “Factors that allow a high level of speech understanding by patients fit with cochlear implants”, American Journal of Audiology **11**, 119–123. 5

Dorval, B., Eckerman, C. O., and Ervin-Tripp, S. (**1984**). “Developmental trends in the quality of conversation achieved by small groups of acquainted peers”, Monographs of the society for research in child development 1–91. 183

Dowker, A., Hart, A., Heal, H., Phillips, C., and Wilson, A. (**1994**). “‘developmental psychologists’ in the reception class: Young childrens ideas about age differences in abilities.”, Poster presentation at the British Psychological Society Developmental Section Conference, Portsmouth. 49

Doyle, M. and 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`. 30, 33

Dumontheil, I., Apperly, I. A., and Blakemore, S.-J. (**2010**). "Online usage of theory of mind continues to develop in late adolescence", Developmental Science **13**, 331–338. 49, 62

Duncan, J., Rhoades, E., and Fitzpatrick, E. (**2014**). *Auditory (Re)Habilitation for Adolescents with Hearing Loss: Theory and Practice.* (Oxford University Press, New York). 9, 48, 183

Dunn, E. (**2014**). "Breakdown and repair in hearing-impaired and normally-hearing 11-year-olds.", Master's thesis, University College London (UCL). 90, 196, 245, 246

Dunn, J. and Kendrick, C. (**1982**). "The speech of two- and three-year-olds to infant siblings: 'baby talk' and the context of communication", Journal of Child Language **9**, 579–595. 57, 193, 194

Ebrahim, R. (**2015**). "Features of talk that enable or hinder the successful completion of a referential communication task by hearing-impaired and normally-hearing children: A conversation analytical examination.", Undergraduate thesis, University College London (UCL). 90, 113, 246

Eimas, P., Siqueland, E., Jusczyk, P., and Vigorito, J. (**1971**). "Speech perception in infants.", Science **968**, 303–306. 7

Eisenberg, L. S. (**2007**). "Current state of knowledge: speech recognition and production in children with hearing impairment", Ear and Hearing **28**, 766–772. 10, 12, 19, 80, 100, 104

Eisenberg, L. S., Dirks, D. D., Takayanagi, S., and Martinez, A. S. (**1998**). "Subjective judgments of clarity and intelligibility for filtered stimuli with equivalent speech intelligibility index predictions", Journal of Speech Language and Hearing Research **41**, 327–339. 199

Eisenberg, L. S., Martinez, A. S., Holowecky, S. R., and Pogorelsky, S. (**2002**). "Recognition of lexically controlled words and sentences by children with normal hearing and children with cochlear implants", Ear and Hearing **23**, 450–462. 15

Eisenberg, L. S., Shannon, R. V., Martinez, A. S., Wygonski, J., and Boothroyd, A. (**2000**). "Speech recognition with reduced spectral cues as a function of age", The Journal of the Acoustical Society of America **107**, 2704–2710. 9

Elfenbein, J., Small, A., and Davis, J. (**1993**). "Developmental patterns of duration discrimination.", Journal of Speech and Hearing Research **36**, 842–849. 8

Elfenbein, J. L., Hardin-Jones, M. A., and Davis, J. M. (**1994**). "Oral communication skills of children who are hard of hearing", Journal of Speech Language and Hearing Research **37**, 216–226. 21, 23, 24, 25, 27, 64, 80, 179, 199, 237

Elliott, L. (**1979**). "Performance of children aged 9 to 17 years on a test of speech intelligibility in noise using sentence materials with controlled word predictability.", Journal of the Acoustical Society of America **66**, 651–653. 9

Elliott, L., Clifton, L., and Servi, D. (**1983**). "Word frequency effects for a closed-set word identification task.", Audiology **22**, 229–240. 9

Ellis, R. J. (**2012**). "Benefit and predictors of outcome from frequency compression hearing aid use", Ph.D. thesis, University of Manchester. 5, 12

Eriks-Brophy, A., Gibson, S., and Tucker, S.-K. (**2013**). "Articulatory Error Patterns and Phonological Process Use of Preschool Children with and without Hearing Loss.", Volta Review **113**, 87–125. 23, 25, 74

Ernestus, M., Hanique, I., and Verboom, E. (**2015**). "The effect of speech situation on the occurrence of reduced word pronunciation variants.", Journal of Phonetics **48**, 60–75. 35

Ertmer, D. J. and Goffman, L. A. (**2011**). "Speech production accuracy and variability in young cochlear implant recipients: comparisons with typically developing age-peers", Journal of Speech Language and Hearing Research **54**, 177–189. 18, 21, 25

Ertmer, D. J., Jung, J., and Kloiber, D. T. (**2013**). "Beginning to Talk Like an Adult: Increases in Speech-Like Utterances in Young Cochlear Implant

Recipients and Typically Developing Children", American Journal of Speech-Language Pathology **22**, 591–603. 20

Fagan, M. K., Pisoni, D. B., Horn, D. L., and Dillon, C. M. (**2007**). "Neuropsychological Correlates of Vocabulary, Reading, and Working Memory in Deaf Children With Cochlear Implants", Journal of Deaf Studies and Deaf Education **12**, 461–471. 62

Farantouri, V., Potamianos, A., and Narayanan, S. (**2008**). "Linguistic analysis of spontaneous children speech", in *Workshop of Child, Computer and Interaction*, 1–8 (Chania, Greece). 18

Ferguson, S. H. (**2004**). "Talker differences in clear and conversational speech: Vowel intelligibility for normal-hearing listeners", The Journal of the Acoustical Society of America **116**, 2365–2373. 35, 178, 199, 212, 247

Ferguson, S. H. (**2012**). "Talker Differences in Clear and Conversational Speech: Vowel Intelligibility for Older Adults With Hearing Loss", Journal of Speech Language and Hearing Research **55**, 779–790. 35, 212

Ferguson, S. H. and Kerr, E. E. (**2008**). "Subjective ratings of sentences in clear versus conversational speech.", Journal of the Acoustical Society of America **124**, 2457. 199

Ferguson, S. H. and Kewley-Port, D. (**2002**). "Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners", The Journal of the Acoustical Society of America **112**, 259–271. 34, 35, 44, 142, 176, 238

Ferguson, S. H. and Kewley-Port, D. (**2007**). "Talker differences in clear and conversational speech: Acoustic characteristics of vowels", Journal of Speech Language and Hearing Research **50**, 1241–1255. 32, 34, 35, 44, 176, 216, 233, 237

Ferguson, S. H. and Lee, J. (**2006**). "Vowel intelligibility in clear and conversational speech for cochlear implant users: A preliminary study", J. Acad. Reh. Audiol **39**, 1–16. 35

Ferguson, S. H., Poore, M. A., and Shrivastav, R. (**2010**). "Acoustic correlates of reported clear speech strategies", Journal of the Academy of Rehabilitative Audiology **43**, 45–64. 34

Ferguson, S. H. and Quené, H. (**2014**). "Acoustic correlates of vowel intelligibility in clear and conversational speech for young normal-hearing and elderly hearing-impaired listenersa)", The Journal of the Acoustical Society of America **135**, 3570–3584. 32, 33, 34, 176

Fernald, A. and Kuhl, P. (**1987**). "Acoustic determinants of infant preference for motherese speech", Infant behavior and Development **10**, 279–293. 39

Ferrier, S., Dunham, P., and Dunham, F. (**2000**). "The Confused Robot: TwoYearOlds' Responses to Breakdowns in Conversation", Social Development **9**, 337–347. 52

Field, A., Miles, J., and Field, Z. (**2012**). *Discovering Statistics Using R* (Sage, London). 127

Finnegan, E. M., Luschei, E. S., and Hoffman, H. T. (**2000**). "Modulations in respiratory and laryngeal activity associated with changes in vocal intensity during speech", Journal of Speech Language and Hearing Research **43**, 934–950. 17

Flavell, J. (**1974**). "The development of inferences about others.", in *Understanding other persons.*, edited by T. Mischel (Rowan and Littlefield, Totowa, New Jersey). 47

Flavell, J., Botkin, P., Fry, C., Wright, J., and Jarvis, P. (**1968**). *The development of role-taking and communication skills in children.* (Wiley, New York). 59

Fletcher, S. G. (**1989**). "Palatometric Specification of Stop, Affricate, and Sibilant Sounds", Journal of Speech Language and Hearing Research **32**, 736–748. 18

Fletcher, S. G., Mahfuzh, F., and Hendarmin, H. (**2001**). "Nasalance in the speech of children with normal hearing and children with hearing loss", American Journal of Speech-Language Pathology **8**, 241–248. 22

Flipsen, P. (**2008**). "Intelligibility of spontaneous conversational speech produced by children with cochlear implants: A review", International Journal of Pediatric Otorhinolaryngology **72**, 559–564. 27, 199

Flipsen, P. and Kangas, K. (**2014**). "Mean length of utterance (MLU) in children with cochlear implants", The Volta Review **114**, 135–155. 26, 27, 182, 183, 195

Flipsen, P. and Lee, S. (**2012**). "Reference data for the American English acoustic vowel space", Clinical Linguistics & Phonetics **26**, 926–933. 18

Forster, K. and Forster, J. (**2003**). "DMDX: A Windows display program with millisecond accuracy.", Behavior Research Methods **35**, 116–124. 96

Fortnum, H. and Davis, A. (**1997**). "Epidemiology of permanent childhood hearing impairment in Trent Region, 1985-1993", British journal of audiology **31**, 409–446. 3

Fortnum, H., Summerfield, A., Marshall, D., Davis, A., and Bamford, J. (**2001**). "Prevalence of permanent childhood hearing impairment in the United Kingdom and implications for universal neonatal hearing screening: questionnaire based ascertainment study.", British Medical Journal **323**, 536–540. 3

Foster, S. H. (**1990**). *The communicative competence of young children: A modular approach* (Longman, London). 1

Franks, J. R. (**1979**). "The influence of exaggerated mouth movement on lip reading.", Audiology and Hearing Education **5**, 12–16. 137

Fussell, S. and Krauss, R. (**1992**). "Coordination of knowledge in communication: Effects of speakers' assumptions about what others know.", Journal of Personality and Social Psychology **62**, 378–391. 43, 217, 241

Gallagher, T. (**1981**). "Contingent query sentences with adult-child discourse.", Journal of Child Language **8**, 51–62. 51, 52

Gallaway, C., Hostler, M., and Reeves, D. (**1990**). "Speech addressed to hearing impaired children by their mothers.", Clinical Linguistics and Phonetics **4**, 221–237. 40

Garnier, M., Henrich, N., and Dubois, D. (**2010**). "Influence of sound immersion and communicative interaction on the Lombard effect", Journal of Speech Language and Hearing Research **53**, 588–608. 37, 45, 78, 111

Geers, A., Brenner, C., and Davidson, L. (**2003**). "Factors Associated with Development of Speech Perception Skills in Children Implanted by Age Five", Ear and Hearing **24**, 24S–35S. 13, 14, 27, 195

Geers, A., Tobey, E., Moog, J., and Brenner, C. (**2008**). "Long-term outcomes of cochlear implantation in the preschool years: From elementary grades to high school", International Journal of Audiology **47**, S21–S30. 11

Geers, A. E. (**2002**). "Factors affecting the development of speech, language, and literacy in children with early cochlear implantation", Language, Speech, and Hearing Services in Schools **33**, 172–183. 11

Geers, A. E. and Sedey, A. L. (**2011**). "Language and verbal reasoning skills in adolescents with 10 or more years of cochlear implant experience", Ear and Hearing **32**, 39S–48S. 11, 15, 183

Giezen, M. R., Escudero, P., and Baker, A. (**2010**). "Use of acoustic cues by children with cochlear implants", Journal of Speech Language and Hearing Research **53**, 1440–1457. 13, 80, 238

Gilbert, H. R. and Campbell, M. I. (**1980**). "Speaking fundamental frequency in three groups of hearing-impaired individuals", Journal of Communication Disorders **13**, 195–205. 22

Gilbert, R. C., Chandrasekaran, B., and Smiljanić, R. (**2014**). "Recognition memory in noise for speech of varying intelligibility", The Journal of the Acoustical Society of America **135**, 389–399. 35

Gilbertson, M. and Kamhi, A. (**1995**). "Novel word learning in children with hearing impairment.", Journal of Speech and Hearing Research **38**, 630–642. 13, 14, 15

Gildersleeve-Neumann, C. E., Davis, B. L., and MacNeilage, P. F. (**2011**). "Syllabic patterns in the early vocalizations of Quichua children", Applied Psycholinguistics **34**, 111–134. 16

Glista, D., Scollie, S., Bagatto, M., Seewald, R., Parsa, V., and Johnson, A. (**2009**). "Evaluation of nonlinear frequency compression: Clinical outcomes.", International Journal of Audiology **48**, 632–644. 12

Goberis, D., Beams, D., Dalpes, M., Abrisch, A., Baca, R., and Yoshinaga-Itano, C. (**2012**). "The Missing Link in Language Development of Deaf and Hard of Hearing Children: Pragmatic Language Development", Seminars in Speech and Language **33**, 297–309. 62, 65, 184

Goberman, A. M. and Elmer, L. W. (**2005**). "Acoustic analysis of clear versus conversational speech in individuals with Parkinson disease", Journal of Communication Disorders **38**, 215–230. 68, 140, 235

Gold, T. (**1980**). "Speech production in hearing-impaired children", Journal of Communication Disorders **13**, 397–418. 27

Granlund, S., Hazan, V., and Baker, R. (**2012**). "An acoustic–phonetic comparison of the clear speaking styles of Finnish–English late bilinguals", Journal of Phonetics **40**, 509–520. 35, 44, 45, 77, 140, 176, 177

Grant, L., Bow, C., Paatsch, L., and Blamey, P. (**2002**). "Comparison of production of /s/ and /z/ between children using cochlear implants and children using hearing aids.", in *Poster presented at: Ninth Australian International Conference on Speech Science and Technology, December 2-5, 2002.* (Melbourne, Australia). 26

Grigos, M. I. (**2009**). "Changes in articulator movement variability during phonemic development: a longitudinal study", Journal of Speech Language and Hearing Research **52**, 164–177. 17

Grynpas, J., Baker, R., and Hazan, V. (**2011**). "Clear speech strategies and speech perception in adverse listening conditions", in *Proceedings of the In-*

*ternational Conference of the Phonetic Sciences (ICPhS) XVI*, 779–782 (Hong Kong). 38, 214

Guralnick, M. J. and Paul-Brown, D. (**1986**). "Communicative Interactions of Mildly Delayed and Normally Developing Preschool Children", Journal of Speech Language and Hearing Research **29**, 2–10. 57

Guralnick, M. J. and Paul-Brown, D. (**1989**). "Peer-Related Communicative Competence of Preschool Children", Journal of Speech Language and Hearing Research **32**, 930–943. 68

Haley, K. L., Seelinger, E., Mandulak, K. C., and Zajac, D. J. (**2010**). "Journal of Phonetics", Journal of Phonetics **38**, 548–554. 32, 180

Halliday, L. and Moore, D. R. (**2010**). "Auditory basis of language and learning disorders (chapter 15).", in *Oxford Handbook of Auditory Science: Hearing.*, edited by D. Moore, P. Fuchs, C. Plack, A. Rees, and A. Palmer (Oxford University Press, Oxford). 4, 12, 13, 20, 79

Harris, H. (**2014**). "Investigating repair in conversations between hearing and hearing-impaired adolescents using conversation analysis.", Master's thesis, University College London (UCL). 90, 196, 246

Harris, K. S., Rubin-Spitz, J., and McGarr, N. S. (**1985**). "The role of production variability in normal and deviant developing speech", in *Conference on the Planning and Production of Speech in Normal and Hearing-impaired Individuals (ASHA Reports 15)*, edited by J. L. Lauter, 27–43 (Rockville, MD). 24, 74

Hazan, V. and 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**, 2139–2152. 36, 37, 38, 77, 94, 107, 125, 136, 199, 200, 201, 212, 213, 236

Hazan, V. and Barrett, S. (**2000**). "The development of phonemic categorization in children aged 6–12", Journal of Phonetics **28**, 377–396. 8, 9, 55, 177

Hazan, V., Fourcin, A., and Abberton, E. (**1991**). "Development of phonetic labeling in hearing-impaired children", Ear and Hearing **12**, 71–84. 12

Hazan, V. and Kim, J. (**2013**). "Acoustic and visual adaptations in speech produced to counter adverse listening conditions", in *Proceedings of Auditory-Visual Speech Processing 2013*, 93–98 (Inria, Rocquencourt, France). 78

Hazan, V. and Markham, D. (**2004**). "Acoustic-phonetic correlates of talker intelligibility for adults and children", The Journal of the Acoustical Society of America **116**, 3108–3118. 31, 32, 204

Hazan, V., Messaoud-Galusi, S., Rosen, S., Nouwens, S., and Shakespeare, B. (**2009**). "Speech Perception Abilities of Adults With Dyslexia: Is There Any Evidence for a True Deficit?", Journal of Speech Language and Hearing Research **52**, 1510–1529. 218

Hazan, V., Romeo, R., and Pettinato, M. (**2013**). "The impact of variation in phoneme category structure on consonant intelligibility", in *International Congress on Acoustics, 2013 Montreal*, 1–6 (Proceedings of Meetings on Acoustics). 32, 74, 142

Hazan, V., Tuomainen, O., and Pettinato, M. (**submitted**). "Suprasegmental characteristics of spontaneous speech produced in good and challenging communicative conditions by talkers aged 9 to 14 years old.", Journal of Speech, Language and Hearing Research . 17, 18, 29, 36, 55, 58, 59, 77, 94, 107, 123, 124, 135, 140, 200, 217, 233, 239

Hazan, V., Uther, M., and Granlund, S. (**2015**). "How does foreigner-directed speech differ from other forms of listener-directed clear speaking styles?", in *International Conference of the Phonetic Sciences (ICPhS) XVIII*. 38

Heldner, M. and Edlund, J. (**2010**). "Pauses, gaps and overlaps in conversations", Journal of Phonetics **38**, 555–568. 94, 137, 183, 195

Helfer, K. S. (**1997**). "Auditory and Auditory-Visual Perception of Clear and Conversational Speech", Journal of Speech Language and Hearing Research **40**, 432–443. 35, 212

Helfer, K. S. (**1998**). “Auditory and Auditory-Visual Recognition of Clear and Conversational Speech by Older Adults”, Journal of the American Academy of Audiology **9**, 234–242. 35

Hennies, J., Penke, M., Rothweiler, M., Wimmer, E., and Hess, M. (**2012**). “Testing the phonemes relevant for German verb morphology in hard-of-hearing children: The FinKon-test”, Logopedics Phoniatrics Vocology **37**, 83–93. 10, 12, 105

Henning, A., Striano, T., and Lieven, E. V. M. (**2005**). “Maternal speech to infants at 1 and 3 months of age”, Infant Behavior and Development **28**, 519–536. 39

Higgins, M. B., McCleary, E. A., Carney, A. E., and Schulte, L. (**2003**). “Longitudinal Changes in Children’s Speech and Voice Physiology after Cochlear Implantation”, Ear and Hearing **24**, 48–70. 21, 22, 66

Hocevar-Boltezar, I., Radsel, Z., Vatovec, J., Geczy, B., Cernelc, S., Gros, A., Zupancic, J., Battelino, S., Lavrencak, B., and Zargi, M. (**2006**). “Change of phonation control after cochlear implantation”, Otology & neurotology **27**, 499–503. 22, 138

Hoff-Ginsberg, E. (**1986**). “Function and structure in maternal speech: Their relation to the childs development of syntax.”, Developmental Psychology **22**, 155–163. 39

Holden-Pitt, L. D., Hazan, V., Revoile, S. G., Edward, D. A., and Droge, J. (**1995**). “Temporal and spectral use for initial plosive voicing perception by hearing-impaired children and normal-hearing children and adults”, European Journal of Disorders of Communication **30**, 417–434. 13

Hollien, H., Green, R., and Massey, K. (**1994**). “Longitudinal research on adolescent voice change in males.”, Journal of the Acoustical Society of America **96**, 2646–2654. 125

Hong, S.-Y., Kwon, K.-A., and Jeon, H.-J. (**2013**). “Children’s Attitudes towards Peers with Disabilities: Associations with Personal and Parental Factors”, Infant and Child Development **23**, 170–193. 50

Hoonhorst, I., Medina, V., Colin, C., Markessis, E., Radeau, M., Deltenre, P., and Serniclaes, W. (**2011**). “Categorical perception of voicing, colors and facial expressions: A developmental study”, Speech Communication **53**, 417–430. 8, 177

Horga, D. and Liker, M. (**2006**). “Voice and pronunciation of cochlear implant speakers”, Clinical Linguistics & Phonetics **20**, 211–217. 24, 74, 179

Horton, W. and Keysar, B. (**1996**). “When do speakers take into account common ground?”, Cognition **59**, 91–117. 43

Houston, D. M. (**2011**). “Infant Speech Perception”, in *Comprehensive Handbook of Pediatric Audiology*, edited by R. Seewald and A. M. Tharpe, 1–16 (Plural Publishing, San Diego, CA). 7

Houston, D. M., Carter, A. K., Pisoni, D. B., Kirk, K. I., and Ying, E. A. (**2005**). “Word learning in children following cochlear implantation”, The Volta Review **105**, 39–70. 11

Houston, D. M., Pisoni, D. B., Kirk, K. I., Ying, E. A., and Miyamoto, R. T. (**2003**). “Speech perception skills of deaf infants following cochlear implantation: a first report”, International Journal of Pediatric Otorhinolaryngology **67**, 479–495. 9, 10

HTK Team (**2012**). “The Hidden Markov Model Toolkit.”, URL `http://htk.eng.cam.ac.uk/`. 95

Hurst, C., Corning, K., and Ferrante, R. (**2012**). “Children’s Acceptance of Others with Disability: The Influence of a Disability-Simulation Program”, Journal of Genetic Counseling **21**, 873–883. 50

Ibertsson, T., Hansson, K., Mäki Torkko, E., Willstedt Svensson, U., and Sahlén, B. (**2009**). “Deaf teenagers with cochlear implants in conversation with hearing peers", International Journal of Language & Communication Disorders **44**, 319–337. 76

Imaizumi, S., Hayashi, A., and Deguchi, T. (**1993**). "Planning in Speech Production: Listener Adaptive Characteristics", The Japan Journal of Logopedics and Phoniatrics **34**, 394–401. 41

Imaizumi, S., Hayashi, A., and Deguchi, T. (**1995**). "Listener adaptive characteristics of vowel devoicing in Japanese dialogue", The Journal of the Acoustical Society of America **98**, 768–778. 41

Ito, K. and Speer, S. R. (**2006**). "Using interactive tasks to elicit natural dialogue", in *Methods in Empirical Prosody Research*, edited by S. Sudhoff, 229–257 (Walter de Gruyter, Berlin). 75, 76

Iyer, S. N. and Oller, D. K. (**2008**). "Prelinguistic vocal development in infants with typical hearing and infants with severe-to-profound hearing loss", The Volta Review **108**, 115–138. 20

Jacewicz, E. and Fox, R. A. (**2014**). "The Effects of Indexical and Phonetic Variation on Vowel Perception in Typically Developing 9- to 12-Year-Old Children", Journal of Speech Language and Hearing Research **57**, 389–405. 9

Jamieson, J. R. (**2010**). "Children and youth who are hard of hearing: hearing accessibility, acoustical context, and development.", in *The Oxford Handbook of Deaf Studies, Language, and Education, volume 2*, edited by M. Marschark and P. Spencer, 376–389 (Oxford University Press, Oxford, UK). 10

Jeanes, R. C., Nienhuys, T. G., and Rickards, F. W. (**2000**). "The pragmatic skills of profoundly deaf children", Journal of Deaf Studies and Deaf Education **5**, 237–247. 63, 76

Jensen, J. and Neff, D. (**1993**). "Development of basic auditory discrimination in preschool children.", Psychological Science **4**, 104–107. 8

Jerger, S. (**2007**). "Current state of knowledge: Perceptual processing by children with hearing impairment", Ear and Hearing **28**, 754–765. 4, 10

Jerger, S., Martin, R., and Damian, M. (**2002**). "Semantic and phonological influences on picture naming by children and teenagers.", Journal of Memory and Language **47**, 229–249. 13

Johansson, V. (**2008**). "Lexical diversity and lexical density in speech and writing: a developmental perspective", Lund University, Dept. of Linguistics and Phonetics: Working Papers 61–79. 183

Johnson, D., Whaley, P., and Dorman, M. (**1984**). "Processing of cues for stop consonant voicing in young hearing-impaired listeners.", Journal of Speech and Hearing Research **27**, 112–118. 12

Jongman, A., Wayland, R., and Wong, S. (**2000**). "Acoustic characteristics of English fricatives.", Journal of the Acoustical Society of America **108**, 1252–1263. 146

Junqua, J. C. (**1993**). "The Lombard reflex and its role on human listener and automatic speech recognizers.", Journal of the Acoustical Society of America **93**, 510–524. 31

Jusczyk, P. and Aslin, R. (**1995**). "Infants detection of the sound patterns of words in fluent speech.", Cognitive Psychology **29**, 1–23. 8

Jusczyk, P., Cutler, A., and Redanz, N. (**1993**). "Infants preference for the predominant stress patterns of English words.", Child Development **64**, 675–687. 8

Jusczyk, P., Hohne, E., and Bauman, A. (**1999**). "Infants sensitivity to allophonic cues for word segmentation.", Perception and Psychophysics **61**, 1465–1476. 8

Kang, K.-H. and Guion, S. G. (**2008**). "Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies", The Journal of the Acoustical Society of America **124**, 3909–3917. 44, 142

Keating, E. and Mirus, G. (**2003**). "Examining interactions across language modalities. Deaf children and hearing peers at school", Anthropology and Education Quarterly **34**, 115–135. 6

Kempe, V. (**2009**). "Personality and Individual Differences", Personality and Individual Differences **47**, 610–615. 58

Kennedy, C. R., McCann, D. C., Campbell, M. J., Law, C. M., Mullee, M., Petrou, S., Watkin, P., Worsfold, S., Yuen, H. M., and Stevenson, J. (**2006**). "Language ability after early detection of permanent childhood hearing impairment.", New England Journal of Medicine **354**, 2131–2142. 4

Keysar, B., Lin, S., and Barr, D. J. (**2003**). "Limits on theory of mind use in adults", Cognition **89**, 25–41. 49

kidLUCID (**2015**). "The kidLUCID website.", URL `http://www.phon.ucl.ac.uk/project/kidLUCID/transcriptions.php`. 293

Kim, H., Hasegawa-Johnson, M., and Perlman, A. (**2011**). "Vowel contrast and speech intelligibility in dysarthria.", Folia Phoniatrica et Logopaedica **63**, 187–194. 32, 180

Kishon-Rabin, L., Taitelbaum, R., Muchnik, C., Gehtler, I., Kronenberg, J., and Hildesheimer, M. (**2002**). "Development of speech perception and production in children with cochlear implants", Annals of Otology, Rhinology and Laryngology **111**, 85–90. 13, 14, 79, 80, 179

Kisilevsky, B. S., Hains, S. M. J., Brown, C. A., Lee, C. T., Cowperthwaite, B., Stutzman, S. S., Swansburg, M. L., Lee, K., Xie, X., Huang, H., Ye, H. H., Zhang, K., and Wang, Z. (**2009**). "Fetal sensitivity to properties of maternal speech and language", Infant behavior and Development **32**, 59–71. 7

Kitamura, C. and Burnham, D. (**2003**). "Pitch and Communicative Intent in Mother's Speech: Adjustments for Age and Sex in the First Year", Infancy **4**, 85–110. 39

Knudsen, E. (**2004**). "Sensitive periods in the development of the brain and behavior.", Journal of Cognitive Neuroscience **16**, 1412–1425. 9

Koehlinger, K. M., Van Horne, A. J. O., and Moeller, M. P. (**2013**). "Grammatical Outcomes of 3- and 6-Year-Old Children Who Are Hard of Hearing", Journal of Speech Language and Hearing Research **56**, 1701–1714. 26, 27, 183

Koenig, L. L., Lucero, J. C., and 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**, 3158–3170. 17, 19, 124, 135

Kondaurova, M. V. and Bergeson, T. R. (**2011**). "The Effects of Age and Infant Hearing Status on Maternal Use of Prosodic Cues for Clause Boundaries in Speech", Journal of Speech Language and Hearing Research **54**, 740–754. 41

Kondaurova, M. V., Bergeson, T. R., and Dilley, L. C. (**2012**). "Effects of deafness on acoustic characteristics of American English tense/lax vowels in maternal speech to infants", The Journal of the Acoustical Society of America **132**, 1039–1049. 40

Kosky, C. and Boothroyd, A. (**2003**). "Perception and Production of Sibilants by Children with Hearing Loss: A training study", The Volta Review **103**, 71–98. 21, 25, 66, 74, 80, 143, 237

Krause, J. C. and Braida, L. D. (**2004**). "Acoustic properties of naturally produced clear speech at normal speaking rates", The Journal of the Acoustical Society of America **115**, 362–378. 31, 37, 44, 216

Krause, J. C. and Braida, L. D. (**2009**). "Evaluating the role of spectral and envelope characteristics in the intelligibility advantage of clear speech", The Journal of the Acoustical Society of America **125**, 3346–3357. 34

Krishnan, S. (**2013**). "A developmental perspective on the integration of language production and comprehension.", Behavioral and Brain Sciences **36**, 363–364. 243

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., Sundberg, U., and Lacerda, F. (**1997**). "Cross-language analysis of phonetic units in language addressed to infants", Science **277**, 684–686. 39

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., and Nelson, T. (**2008**). "Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e)", Philosophical Transactions of the Royal Society B: Biological Sciences **363**, 979–1000. 7

Kuhl, P. K. and Meltzoff, A. N. (**1982**). "Bimodal speech perception in early infancy", The Journal of the Acoustical Society of America **71**, S77–S78. 8

Kuo, Y.-C., Rosen, S., and Faulkner, A. (**2008**). "Acoustic cues to tonal contrasts in Mandarin: Implications for cochlear implants", The Journal of the Acoustical Society of America **123**, 2815–2824. 12, 123

Kučera, M. and Francis, W. N. (**1967**). *Computational analysis of present-day American English.* (Brown University Press., Providence, RI). 184

Lachs, L., Pisoni, D. B., and Kirk, K. I. (**2001**). "Use of audiovisual information in speech perception by prelingually deaf children with cochlear implants: A first report", Ear and Hearing **22**, 236–251. 14

Lam, C. and Kitamura, C. (**2010**). "Maternal interactions with a hearing and hearing-impaired twin: similarities and differences in speech input, interaction quality and word production", Journal of Speech, Language, and Hearing Research **53**, 543–555. 40

Lam, C. and Kitamura, C. (**2012**). "Mommy, speak clearly: induced hearing loss shapes vowel hyperarticulation", Developmental Science **15**, 212–221. 40, 42

Lam, J., Tjaden, K., and Wilding, G. (**2012**). "Acoustics of Clear Speech: Effect of Instruction", Journal of Speech Language and Hearing Research **55**, 1807. 36, 42

Lane, H., Denny, M., Guenther, F. H., Hanson, H. M., Marrone, N., Matthies, M. L., Perkell, J. S., Stockmann, E., Tiede, M., and Vick, J. (**2007**). "On the structure of phoneme categories in listeners with cochlear implants", Journal of Speech Language and Hearing Research **50**, 2–14. 14

Lane, H., Wozniak, J., and Perkell, J. (**1994**). "Changes in voice-onset time in speakers with cochlear implants.", Journal of the Acoustical Society of America **96**, 56–64. 25, 80

Larsen-Freeman, D. and Long, M. H. (**1991**). *An introduction to second language acquisition research.* (Longman, London). 38

Lecanuet, J., Gautheron, B., Locatelli, A., Schaal, B., Jacquet, A., and Busnel, M. (**1998**). "What sounds reach fetuses: Biological and nonbiological modeling of the transmission of pure tones.", Developmental Psychobiology **33**, 203–219. 7

Lecanuet, J., Granier-Deferre, C., Jacquet, A.-Y., Capponi, I., and Ledru, L. (**1993**). "Prenatal discrimination of a male and female voice uttering the same sentence.", Early Development and Parenting **2**, 217–228. 7

Lederberg, A. R. and Beal-Alvarez, J. S. (**2011**). "Expressing meaning: From prelinguistic communication to building vocabulary.", in *The Oxford Handbook of Deaf Studies, Language and Education (Volume 1)*, edited by M. Marschark and P. E. Spencer, 258–275 (Oxford University Press, Oxford). 21

Lederberg, A. R. and Everhart, V. S. (**2000**). "Conversations Between Deaf Children and Their Hearing Mothers: Pragmatic and Dialogic Characteristics", Journal of Deaf Studies and Deaf Education **5**, 303–322. 40

Lederberg, A. R., Ryan, H. B., and Robbins, B. L. (**1986**). "Peer interaction in young deaf children: The effect of partner hearing status and familiarity.", Developmental Psychology **22**, 691–700. 56, 67, 242

Lee, S., Potamianos, A., and Narayanan, S. (**1999**). "Acoustics of children's speech: Developmental changes of temporal and spectral parameters", The Journal of the Acoustical Society of America **105**, 1455–1468. 16, 17, 18, 55, 123, 124, 135, 142, 177

Leinonen, E. and Letts, C. (**1997**). "Referential communication tasks: performance by normal and pragmatically impaired children", International Journal of Language & Communication Disorders **32**, 53–65. 75

Lenden, J. M. and Flipsen, Jr., P. (**2007**). "Prosody and voice characteristics of children with cochlear implants", Journal of Communication Disorders **40**, 66–81. 22, 75

Levrez, C., Bourdin, B., Le Driant, B., D'Arc, B. F., and Vandromme, L. (**2012**). "The Impact of Verbal Capacity on Theory of Mind in Deaf and Hard of Hearing Children", American Annals of the Deaf **157**, 66–77. 60, 61

Levy, Y. (**1999**). "Early metalinguistic competence: speech monitoring and repair behavior.", Developmental Psychology **35**, 822–834. 51, 52

Lewis, A. (**1993**). "Primary School Children's Understanding of Severe Learning Difficulties", Educational Psychology **13**, 133–145. 50

Liberman, A. M., Harris, K. S., Hoffman, H. S., and Griffith, B. C. (**1957**). "The discrimination of speech sounds within and across phoneme boundaries.", Journal of Experimental Psychology **54**, 358–368. 8

Liker, M., Mildner, V., and Šindija, B. (**2007**). "Acoustic analysis of the speech of children with cochlear implants: A longitudinal study", Clinical Linguistics & Phonetics **21**, 1–11. 26, 74

Lind, C., Campbell, K., Davey, A., Rodgers, L., Seipolt, C., and Akins, C. (**2010**). "Some Prosodic Characteristics of Repeated Talk following Conversation Repair Requests by Adults with Hearing Impairment", Seminars in Hearing **31**, 127–139. 36, 41, 246

Lindblom, B. (**1990**). "Explaining phonetic variation: A sketch of the H&H theory", in *Speech production and speech modelling*, edited by W. J. Hardcastle and A. Marchal, 403–439 (Kluwer Academic Publishers). 42, 43, 136, 178, 217, 233, 241, 243, 244, 246

Lisker, L. (**1978**). "In qualified defense of VOT.", Language and Speech **21**, 375–383. 18

Liu, S., Del Rio, E., Bradlow, A. R., and Zeng, F.-G. (**2004**). "Clear speech perception in acoustic and electric hearing", The Journal of the Acoustical Society of America **116**, 2374–2383. 35

Liu, S. and Zeng, F.-G. (**2006**). "Temporal properties in clear speech perception", The Journal of the Acoustical Society of America **120**, 424–432. 31

Lloyd, J. (**2003**). "Oral communication between hearing-impaired children and their normally hearing peers", Ph.D. thesis, University of Manchester. 56, 75, 76, 245

Lloyd, J., Lieven, E., and Arnold, P. (**2005**). "The oral referential communication skills of hearing-impaired children", Deafness & Education International **7**, 22–42. 64, 65, 76

Lloyd, P., Camaioni, L., and Ercolani, P. (**1995**). "Assessing referential communication skills in the primary school years: a comparative study.", British Journal of Developmental Psychology 13–29. 54

Lloyd, P., Mann, S., and Peers, I. (**1998**). "The growth of speaker and listener skills from five to eleven years", First Language **18**, 81–103. 53, 54, 56, 64, 65, 76

Löfqvist, A., Sahlén, B., and Ibertsson, T. (**2010**). "Vowel spaces in Swedish adolescents with cochlear implants", The Journal of the Acoustical Society of America **128**, 3064–3069. 14, 26, 74, 179, 241

Long, M. H. (**1983**). "Linguistic and conversational adjustments to non-native speakers", Studies in second language acquisition **5**, 177–193. 38, 194

Lovett, R. E. S., Vickers, D. A., and Summerfield, A. Q. (**2015**). "Bilateral Cochlear Implantation for Hearing-Impaired Children: Criterion of Candidacy Derived from an Observational Study", Ear and Hearing **36**, 14–23. 11

Lu, X. (**2012**). "The Relationship of Lexical Richness to the Quality of ESL Learners' Oral Narratives", The Modern Language Journal **96**, 190–208. 182

Lundy, J. E. (**2002**). "Age and language skills of deaf children in relation to theory of mind development", Journal of Deaf Studies and Deaf Education **7**, 41–56. 60, 61

MacArdle, B., Hazan, V., and Prasher, D. (**1999**). "Speech pattern audiometry in hearing impaired children.", British Journal of Audiology **33**, 383–393. 12, 13

Macaulay, C. E. and Ford, R. M. (**2006**). "Language and theory-of-mind development in prelingually deafened children with cochlear implants: a preliminary investigation", Cochlear Implants International **7**, 1–14. 60

Maniwa, K., Jongman, A., and Wade, T. (**2008**). "Perception of clear fricatives by normal-hearing and simulated hearing-impaired listeners", The Journal of the Acoustical Society of America **123**, 1114–1125. 33

Maniwa, K., Jongman, A., and Wade, T. (**2009**). "Acoustic characteristics of clearly spoken English fricatives", The Journal of the Acoustical Society of America **125**, 3962–3973. 46, 234

Markham, D. and Hazan, V. (**2004**). "The effect of talker-and listener-related factors on intelligibility for a real-word, open-set perception test", Journal of Speech Language and Hearing Research **47**, 725–737. 19

Markman, A. B. and Makin, V. S. (**1998**). "Referential communication and category acquisition.", Journal of Experimental Psychology: General **127**, 331–354. 75

Marschark, M., Green, V., Hindmarsh, G., and Walker, S. (**2000**). "Understanding theory of mind in children who are deaf", Journal of child psychology and psychiatry **41**, 1067–1073. 62

Marschark, M. and Hauser, P. C. (**2012**). *How deaf children learn: what parents and teachers need to know* (Oxford University Press, Oxford). 30, 31, 33, 48, 238, 239

Martin, A., Schatz, T., Versteegh, M., Miyazawa, K., Mazuka, R., Dupoux, E., and Cristia, A. (**2015**). "Mothers Speak Less Clearly to Infants Than to Adults: A Comprehensive Test of the Hyperarticulation Hypothesis", Psychological Science **26**, 341–347. 39

Martin, D. and Bat-Chava, Y. (**2003**). "Negotiating deaf–hearing friendships: Coping strategies of deaf boys and girls in mainstream schools", Child: care, health and development **29**, 511–521. 6

Martin, D., Bat-Chava, Y., Lalwani, A., and Waltzman, S. B. (**2010**). "Peer Relationships of Deaf Children With Cochlear Implants: Predictors of Peer Entry and Peer Interaction Success", Journal of Deaf Studies and Deaf Education **16**, 108–120. 56

Massaro, D. W., Cohen, M. M., and Gesi, A. T. (**1993**). "Long term training, transfer, and retention in learning to lipread.", Perception and Psychophysics **5J**, 549–556. 137

Masur, E. F. (**1978**). "Preschool boys' speech modifications: The effect of listeners' linguistic levels and conversational responsiveness", Child Development 924–927. 59, 217, 236, 239

Matthews, D., Lieven, E., and Tomasello, M. (**2007**). "How toddlers and preschoolers learn to uniquely identify referents for others: A training study", Child Development **78**, 1744–1759. 53

Mattys, S. and Jusczyk, P. (**2001**). "Phonotactic cues for segmentation of fluent speech by infants.", Cognition **78**, 91–121. 8

McCune, L. and Vihman, M. M. (**2001**). "Early Phonetic and Lexical Development: A Productivity Approach", Journal of Speech Language and Hearing Research **44**, 670–684. 16

McGarr, N. S. and Campbell, M. M. (**1995**). "Ephphatha: Opening inroads to understanding articulatory organisation in persons with hearing impairment (chapter 29)", in *Producing speech: Contemporary Issues*, edited by F. Bell-Berti and L. J. Raphael, 433–454 (AIP Press, New York). 21

McGarr, N. S. and Osberger, M. J. (**1978**). "Pitch deviancy and intelligibility of deaf speech", Journal of Communication Disorders 237–247. 28

McGuckian, M. and Henry, A. (**2007**). "The grammatical morpheme deficit in moderate hearing impairment.", International Journal of Language and Communication Disorders **42 (Suppl.1)**, 17–36. 27

McKee, G., Malvern, D., and Richards, B. (**2000**). "Measuring vocabulary diversity using dedicated software", Literary and linguistic computing **15**, 323–338. 185, 186

McMurray, B., Kovack-Lesh, K. A., Goodwin, D., and McEchron, W. (**2013**). "Infant directed speech and the development of speech perception: Enhancing development or an unintended consequence?", Cognition **129**, 362–378. 39

Medina, V. and Serniclaes, W. (**2009**). "Development of voicing categorization in deaf children with cochlear implant", in *Interspeech 2009*, 1–4 (Brighton). 13, 14, 66

Ménard, L., Polak, M., Denny, M., Burton, E., Lane, H., Matthies, M. L., Marrone, N., Perkell, J. S., Tiede, M., and Vick, J. (**2007**). "Interactions of speaking condition and auditory feedback on vowel production in postlingually deaf adults with cochlear implants", The Journal of the Acoustical Society of America **121**, 3790–3801. 67

Messaoud-Galusi, S., Hazan, V., and Rosen, S. (**2011**). "Investigating Speech Perception in Children With Dyslexia: Is There Evidence of a Consistent Deficit in Individuals?", Journal of Speech Language and Hearing Research **54**, 1682–1701. 98, 104

Metz, D. E., Samar, V. J., Schiavetti, N., Sitler, R. W., and Whitehead, R. L. (**1985**). "Acoustic Dimensions of Hearing-Impaired Speakers' Intelligibility", Journal of Speech Language and Hearing Research **28**, 345–355. 28, 74, 111

Mildner, V., Šindija, B., and Vrban Zrinski, K. (**2006**). "Speech perception of children with cochlear implants and children with traditional hearing aids", Clinical Linguistics & Phonetics **20**, 219–229. 13, 14, 238

Miller, S. (**2000**). "Children's Understanding of Preexisting Differences in Knowledge and Belief", Developmental Review **20**, 227–282. 48, 49

Moeller, M. P. (**2000**). "Early intervention and language development in children who are deaf and hard of hearing", PEDIATRICS **106**, 1–9. 10, 14, 15

Moeller, M. P., Hoover, B., Putman, C., Arbataitis, K., Bohnenkamp, G., Peterson, B., Lewis, D., Estee, S., Pittman, A., and Stelmachowicz, P. (**2007**a). "Vocalizations of infants with hearing loss compared with infants with normal hearing: Part II–Transition to words", Ear and Hearing **28**, 628–642. 20

Moeller, M. P., Hoover, B., Putman, C., Arbataitis, K., Bohnenkamp, G., Peterson, B., Wood, S., Lewis, D., Pittman, A., and Stelmachowicz, P. (**2007**b). "Vocalizations of infants with hearing loss compared with infants with normal hearing: Part I–Phonetic development", Ear and Hearing **28**, 605–627. 20

Moeller, M. P. and Schick, B. (**2006**). "Relations between maternal input and theory of mind understanding in deaf children", Child Development **77**, 751–766. 60, 61

Moeller, M. P., Tomblin, J. B., Yoshinaga-Itano, C., Connor, C. M., and Jerger, S. (**2007**c). "Current state of knowledge: Language and literacy of children with hearing impairment", Ear and Hearing **28**, 740–753. 9, 100

Mohammed, T. E. (**2007**). "An Investigation of Speechreading in Profoundly Congenitally Deaf British Adults", Ph.D. thesis, University College London (UCL). 137

Monsen, R. B. (**1974**). "Durational Aspects of Vowel Production in the Speech of Deaf Children", Journal of Speech Language and Hearing Research **17**, 386. 24, 74, 80, 181

Monsen, R. B. (**1976**). "Normal and reduced phonological space in the production of vowels by deaf adolescents", The Journal of the Acoustical Society of America **59**, S86. 23, 24

Monsen, R. B. (**1978**). "Toward Measuring How Well Hearing-Impaired Children Speak", Journal of Speech Language and Hearing Research **21**, 197–219. 28

Monsen, R. B. (**1979**). "Acoustic Qualities of Phonation in Young Hearing-Impaired Children", Journal of Speech Language and Hearing Research **22**, 270–288. 22, 23

Montag, J. L., AuBuchon, A. M., Pisoni, D. B., and Kronenberger, W. G. (**2014**). "Speech Intelligibility in Deaf Children After Long-Term Cochlear Implant Use", Journal of Speech Language and Hearing Research **57**, 2332–2343. 21, 28, 198, 214

Montangero, J. (**1996**). *Understanding changes in time: The development of diachronic thinking in 7 to 12-year-old children.* (Taylor and Francis, London). 49

Moore, B. C. (**2007**). *Cochlear hearing loss: physiological, psychological, and technical issues* (John Wiley and Sons, Chichester). 3, 4, 5, 11

Morrongiello, B., Robson, R., and Best, C. (**1984**). "Trading relations in the perception of speech by 5-year-old children.", Journal of Experimental Child Psychology **37**, 231–250. 8

Most, T. (**2002**). "The use of repair strategies by children with and without hearing impairment.", Language, Speech and Hearing Services in Schools **33**, 112–123. 63, 76, 196

Most, T. and Peled, M. (**2007**). "Perception of Suprasegmental Features of Speech by Children With Cochlear Implants and Children With Hearing Aids", Journal of Deaf Studies and Deaf Education **12**, 350–361. 12

Most, T., Shina-August, E., and Meilijson, S. (**2010**). "Pragmatic Abilities of Children With Hearing Loss Using Cochlear Implants or Hearing Aids Compared to Hearing Children", Journal of Deaf Studies and Deaf Education **15**, 422–437. 62, 64

Munson, B. (**2004**). "Variability in/s/ production in Children and Adults: Evidence From Dynamic Measures of Spectral Mean", Journal of Speech Language and Hearing Research **47**, 58–69. 19

Nabelek, A. and Robinson, P. (**1982**). "Monaural and binaural speech perception in reverberation for listeners of various ages.", The Journal of the Acoustical Society of America **71**, 1242–1248. 8

Nadol, J. B. and Merchant, S. N. (**2001**). “Histopathology and molecular genetics of hearing loss in the human.”, International Journal of Paediatric Otorhinolaryngology **61**, 1–15. 3

Nakamura, H., Gilbert, H., and Robb, M. (**2007**). “A comparative study of vocal fundamental frequency in the speech of Japanese and American hearing-impaired children”, Logopedics Phoniatrics Vocology **32**, 17–22. 21

Nathani, S., Oller, D. K., and Neal, A. R. (**2007**). “On the robustness of vocal development: An examination of infants with moderate to severe hearing loss and additional risk factors.”, Journal of Speech Language and Hearing Research **50**, 1425–1444. 20

NDCS (**2012**a). *Look, Smile, Chat*, London, UK, http://youngpeople.ndcsbuzz.org.uk/looksmilechat/. 1

NDCS (**2012**b). *Tips for communicating with deaf children.*, London, UK, leaflet. 30, 31, 33, 136, 239

Neumeyer, V., Harrington, J., and Draxler, C. (**2010**). “An acoustic analysis of the vowel space in young and old cochlear-implant speakers”, Clinical Linguistics & Phonetics **24**, 734–741. 26, 74, 179

Newman, R., Clouse, S., and Burnham, J. (**2001**). “The perceptual consequences of within-talker variability in fricative production”, The Journal of the Acoustical Society of America **109**, 1181–1196. 32, 74, 142, 178

Ní Almhain, M. (**2014**). “Breakdown and repair in talk between pairs of 9-year-old hearing-impaired and normally-hearing children.”, Master’s thesis, University College London (UCL). 90, 113, 196, 246

Nicholas, J. G. and Geers, A. E. (**2006**). “Effects of early auditory experience on the spoken language of deaf children at 3 years of age”, Ear and Hearing **27**, 286–298. 21, 75

Nicolaidis, K. and Sfakiannaki, A. (**2007**). “An acoustic analysis of vowels produced by Greek speakers with hearing impairment.”, in *Proceedings of ICPhS*, 1969–1972. 24

Niebuhr, O. and Michaud, A. (**2015**). "Speech data acquisition: the underestimated challenge", KALIPHO - Kieler Arbeiten zur Linguistik und Phonetik 3 **2**, 1–42. 74, 75, 76

Nilsen, E. S. and Mangal, L. (**2011**). "Which is important for preschoolers' production and repair of statements: What the listener knows or what the listener says?", Journal of Child Language **39**, 1121–1134. 52, 54

Niparko, J. K. (**2009**). *Cochlear implants: principles and practices* (Lippincott Williams and Wilkins, Philadephia, PA). 20, 21, 66, 124

Niparko, J. K., Tobey, E. A., Thal, D. J., Eisenberg, L. S., Wang, N.-Y., Quittner, A. L., Fink, N. E., and Team, C. I. (**2010**). "Spoken language development in children following cochlear implantation", Journal of the American Medical Association **303**, 1498–1506. 11, 27

Nippold, M. A. and Duthie, J. K. (**2003**). "Mental Imagery and Idiom Comprehension: A Comparison of School-Age Children and Adults", Journal of Speech Language and Hearing Research **46**, 788–799. 19

Nippold, M. A., Hesketh, L. J., Duthie, J. K., and 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**, 1048–1064. 19, 55, 183, 193

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. 16, 17, 18

Nittrouer, S. and Boothroyd, A. (**1990**). "Context effects in phoneme and word recognition by young children and older adults", The Journal of the Acoustical Society of America **87**, 2705–2715. 8, 9

Nittrouer, S., Caldwell-Tarr, A., Tarr, E., Lowenstein, J. H., Rice, C., and Moberly, A. C. (**2013**). "Improving speech-in-noise recognition for children with hearing loss: Potential effects of language abilities, binaural summation, and head shadow", International Journal of Audiology **52**, 513–525. 9

Nittrouer, S., Estee, S., Lowenstein, J. H., and 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. 16

Nittrouer, S. and Lowenstein, J. (**2009**). "Does harmonicity explain childrens cue weighting of fricative-vowel syllables?", Journal of the Acoustical Society of America **125**, 1679–1692. 9

Nittrouer, S. and Lowenstein, J. H. (**2010**). "Learning to perceptually organize speech signals in native fashion", The Journal of the Acoustical Society of America **127**, 1624–1635. 7

Nittrouer, S. and Studdert-Kennedy, M. (**1987**). "The role of coarticulatory effects on the perception of fricatives by children and adults.", Journal of Speech and Hearing Research **30**, 319–329. 9

Norbury, C. F., Bishop, D. V. M., and Briscoe, J. (**2001**). "Production of English finite verb morphology: A comparison of SLI and mild-moderate hearing impairment.", Journal of Speech, Language, and Hearing Research **44**, 165–178. 27

Ohala, J. J. (**1994**). "Acoustic study of clear speech: a test of the contrastive hypothesis", in *International Symposium on Prosody*, 75–89 (Yokohama, Japan). 46

O'Halpin, R. (**2001**). "Intonation issues in the speech of hearing impaired children: analysis, transcription and remediation", Clinical Linguistics & Phonetics **15**, 529–550. 21

Oller, D. K. (**2000**). *The emergence of the speech capacity* (Erlbaum, Mahwah, NJ). 16

Oller, K. D., Eilers, R., Bull, D., and Carney, A. (**1985**). "Prespeech vocalizations of a deaf infant: A comparison with normal metaphonological development.", Journal of Speech and Hearing Research **28**, 47–63. 20

Osberger, M. J. and McGarr, N. S. (**1982**). “Speech production characteristics of the hearing impaired”, Speech and Language: Advances in Basic Research and Practice **8**, 221–278. 22

Payton, K. L., Uchanski, R. M., and Braida, L. D. (**1994**). “Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing”, The Journal of the Acoustical Society of America **95**, 1581–1592. 35, 199

Pedersen, M. F., Møller, S., Krabbe, S., Bennett, P., and Svenstrup, B. (**1990**). “Fundamental voice frequency in female puberty measured with electroglottography during continuous speech as a secondary sex caracteristic. A comparison between voice, pubertal stages, oestrogens and androgens”, International Journal of Pediatric Otorhinolaryngology **20**, 17–24. 17

Pelegrín-García, D., Smits, B., Brunskog, J., and Jeong, C.-H. (**2011**). “Vocal effort with changing talker-to-listener distance in different acoustic environments”, The Journal of the Acoustical Society of America **129**, 1981–1990. 37

Peng, S.-C., Tomblin, J. B., Cheung, H., Lin, Y.-S., and Wang, L.-S. (**2004**). “Perception and Production of Mandarin Tones in Prelingually Deaf Children with Cochlear Implants”, Ear and Hearing **25**, 251–264. 22

Peng, S.-C., Tomblin, J. B., Spencer, L. J., and Hurtig, R. R. (**2007**). “Imitative production of rising speech intonation in pediatric cochlear implant recipients”, Journal of Speech Language and Hearing Research **50**, 1210–1227. 22

Peng, S.-C., Tomblin, J. B., and Turner, C. W. (**2008**). “Production and perception of speech intonation in pediatric cochlear implant recipients and individuals with normal hearing”, Ear and Hearing **29**, 336–351. 22

Peters, K., Remmel, E., and Richards, D. (**2009**). “Language, mental state vocabulary, and false belief understanding in children with cochlear implants”, Language, Speech, and Hearing Services in Schools **40**, 245–255. 61

Peterson, C. C. (**2004**). "Theory of mind development in oral deaf children with cochlear implants or conventional hearing aids", Journal of child psychology and psychiatry **45**, 1096–1106. 60

Peterson, C. L., Danner, F. W., and Flavell, J. H. (**1972**). "Developmental changes in children's response to three indications of communicative failure", Child Development 1463–1468. 51, 52

Pettinato, M., Tuomainen, O., Granlund, S., and Hazan, V. (**2016**). "Vowel space area in later childhood and adolescence: Effects of age, sex and ease of communication", Journal of Phonetics **54**, 1–14. 18, 29, 58, 59, 60, 77, 94, 127, 175, 177, 217

Piaget, J. (**1926**). *The language and thought of the child.* (Harcourt, New York). 47

Picheny, M. A., Durlach, N. I., and Braida, L. D. (**1985**). "Speaking Clearly for the Hard of Hearing I", Journal of Speech Language and Hearing Research **28**, 96–103. 35, 212

Picheny, M. A., Durlach, N. I., and Braida, L. D. (**1986**). "Speaking Clearly for the Hard of Hearing II", Journal of Speech Language and Hearing Research **29**, 434–446. 34, 35, 44, 45, 136, 137, 176, 178, 216, 238, 247

Pichora-Fuller, K., Goy, H., and van Lieshout, P. (**2010**). "Effect on Speech Intelligibility of Changes in Speech Production Influenced by Instructions and Communicative Environments", Seminars in Hearing **31**, 77–94. 34

Pickering, M. J. and Garrod, S. (**2004**). "Toward a mechanistic psychology of dialogue", Behavioral and Brain Sciences **27**, 169–190. 36, 243

Pickering, M. J. and Garrod, S. (**2013**). "An integrated theory of language production and comprehension", Behavioral and Brain Sciences **36**, 329–347. 36, 243

Pimperton, H., Blythe, H., Kreppner, J., Mahon, M., Peacock, J. L., Stevenson, J., Terlektsi, E., Worsfold, S., Yuen, H. M., and Kennedy, C. R. (**2014**). "The impact of universal newborn hearing screening on long-term literacy outcomes: a prospective cohort study", Archives of Disease in Childhood 1–7. 87

Pisoni, D. B. and Geers, A. E. (**2000**). "Working memory in deaf children with cochlear implants: correlations between digit span and measures of spoken language processing.", The Annals of otology, rhinology and laryngology. Supplement **185**, 92–93. 28

Pittman, A. (**2011**). "Children's Performance in Complex Listening Conditions: Effects of Hearing Loss and Digital Noise Reduction", Journal of Speech Language and Hearing Research **54**, 1224–1239. 14

Ponton, C. and Eggermont, J. (**2001**). "Of kittens and kids: altered cortical maturation following profound deafness and cochlear implant use.", Audiology an Neurootology **6**, 363–380. 9

Pratt, S. R. and Tye-Murray, N. (**2008**). "Speech impairment secondary to hearing loss (chapter 13).", in *Clinical Management of Sensorimotor Speech Disorders*, edited by M. R. McNeil, 204–234 (Thieme Medical Publishers, New York). 21, 25, 143

Pursell, S. L., Swanson, L. A., Hedrick, M. S., and Nabelek, A. K. (**2002**). "Categorical labeling of synthetic/I/and/$\varepsilon$/in adults and school-age children", Journal of Phonetics **30**, 131–137. 8

R Core Team (**2014**). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, URL `http://www.R-project.org/`. 125, 127

Rahilly, J. (**2013**). "Vowel Disorders in Hearing Impairment", in *Handbook of Vowels and Vowel Disorders*, edited by M. J. Ball and F. E. Gibbon, 364–385 (Psychology Press, New York). 20, 24

Raine, C. (**2013**). "Cochlear implants in the United Kingdom: Awareness and utilization", Cochlear Implants International **14**, S32–S37. 5

Reches, M. and Pérez Pereira, M. (**2007**). "Referential communication abilities and theory of mind development in preschool children.", Journal of Child Language **34**, 21–52. 54

Redford, M. A. (**2014**). "The perceived clarity of children's speech varies as a function of their default articulation rate", The Journal of the Acoustical Society of America **135**, 2952–2963. 18, 19, 28, 199, 214

Redford, M. A. and Gildersleeve-Neumann, C. E. (**2009**). "The development of distinct speaking styles in preschool children", Journal of Speech Language and Hearing Research **52**, 1434–1448. 59, 175

Remmel, E. and Peters, K. (**2009**). "Theory of Mind and Language in Children With Cochlear Implants", Journal of Deaf Studies and Deaf Education **14**, 218–236. 61

Richards, B. (**1987**). "Type/Token Ratios: what do they really tell us?", Journal of Child Language **14**, 201–209. 183

Rinker, T. W. (**2013**). *qdap: Quantitative Discourse Analysis Package*, University at Buffalo/SUNY, Buffalo, New York, URL `http://github.com/trinker/qdap`, version 2.2.0. 125

Roberts, R. J. and Patterson, C. (**1983**). "Perspective taking and referential communication: the question of correspondence reconsidered.", Child Development **54**, 1005–1014. 54

Rodríguez, M. S. and Lana, E. T. (**1996**). "Dyadic Interactions Between Deaf Children and Their Communication Partners", American Annals of the Deaf **141**, 245–251. 67

Romeo, R., Hazan, V., and Pettinato, M. (**2013**). "Developmental and gender-related trends of intra-talker variability in consonant production", The Journal of the Acoustical Society of America **134**, 3781–3792. 18, 30, 32, 55, 96, 111, 143, 146, 147, 148, 173, 175, 177

Rotteveel, L., Snik, A., Vermeulen, A. M., Cremers, C., and Mylanus, E. (**2008**). "Speech perception in congenitally, prelingually and postlingually deaf children expressed in an equivalent hearing loss value", Clinical Otolaryngology **33**, 560–569. 11

Roy, B. C., Frank, M. C., and Roy, D. (**2009**). "Exploring word learning in a high-density longitudinal corpus", in *31st Annual Meeting of the Cognitive Science Society* (Cognitive Science Society, Inc.). 39, 217

Ruffin, C. V., Kronenberger, W. G., Colson, B. G., Henning, S. C., and 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**, 289–296. 11, 15

Russell, P. A., Hosie, J. A., Gray, C. D., Scott, C., Hunter, N., Banks, J. S., and Macaulay, M. C. (**1998**). "The development of theory of mind in deaf children", Journal of child psychology and psychiatry **39**, 903–910. 48, 60

Ryalls, J. and Larouche, A. (**1992**). "Acoustic integrity of speech production in children with moderate and severe hearing impairment", Journal of Speech Language and Hearing Research **35**, 88–95. 21

Sachs, J. and Devin, J. (**1976**). "Young children's use of age-appropriate speech styles in social interaction and role-playing", Journal of Child Language **3**, 81–98. 57, 193, 194

Sacks, H., Schegloff, E., and Jefferson, G. (**1974**). "A simplest systematics for the organization of turn-taking for conversation.", Language **50**, 696–735. 41, 90

Saffran, J., Aslin, R., and Newport, E. (**1996**). "Statistical learning by 8-month-old infants.", Science **274**, 1926–1928. 8

Sandgren, O., Ibertsson, T., Andersson, R., Hansson, K., and Sahlén, B. (**2010**). "'You sometimes get more than you ask for': responses in referential communication between children and adolescents with cochlear implant and hearing peers", International Journal of Language & Communication Disorders **46**, 375–385. 64, 76

Sankowska, J., García Lecumberri, M. L., and Cooke, M. (**2011**). "Vowel shortening in altered speech modes elicited via an interactive task", in *International Conference of the Phonetic Sciences (ICPhS) XVII*, 1–4 (Hong Kong). 38, 45, 77, 111

Sarant, J. Z., Blamey, P. J., Dowell, R. C., Clark, G. M., and Gibson, W. (**2001**). "Variation in speech perception scores among children with cochlear implants", Ear and Hearing **22**, 18–28. 11

Scarborough, R., Brenier, J., Zhao, Y., Hall-Lew, L., and 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. 37, 38

Schertz, J. (**2013**). "Exaggeration of featural contrasts in clarifications of misheard speech in English", Journal of Phonetics **41**, 249–263. 45, 178, 234

Schick, B., De Villiers, P., De Villiers, J., and Hoffmeister, R. (**2007**). "Language and theory of mind: A study of deaf children", Child Development **78**, 376–396. 60, 61

Schmidt, C. and Paris, S. (**1984**). "Verbal communicative skills in children.", Advances in Child Development and Behavior **18**, 2–47. 47, 52, 241

Schober, M. (**1993**). "Spatial perspective taking in conversation.", Cognition **47**, 1–24. 43, 217, 233, 241

Schum, D. J. (**1996**). "Intelligibility of Clear and Conversational Speech ofYoung and Elderly Talkers", Journal of the American Academy of Audiology **7**, 212–218. 35

Seewald, R. C. and Brackett, D. (**1984**). "Spoken language modifications as a function of the age and hearing ability of the listener", The Volta Review **86**, 20–35. 56, 57

Serry, T. A. and Blamey, P. J. (**1999**). "A 4-year investigation into phonetic inventory development in young cochlear implant users", Journal of Speech Language and Hearing Research **42**, 141–154. 25, 75

Shannon, R. (**2002**). "The relative importance of amplitude, temporal and spectral cues for cochlear implant processor design.", American Journal of Audiology **11**, 124–127. 5

Shatz, M. and 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. 57, 59, 184, 193, 194

Silbert, N. and de Jong, K. (**2008**). "Focus, prosodic context, and phonological feature specification: Patterns of variation in fricative production", The Journal of the Acoustical Society of America **123**, 2769–2779. 177

Silvestre, N., Ramspott, A., and Pareto, I. D. (**2007**). "Conversational skills in a semistructured interview and self-concept in deaf students", Journal of Deaf Studies and Deaf Education **12**, 38–54. 61

Simon, C. and Fourcin, A. J. (**1978**). "Crosslanguage study of speechpattern learning", The Journal of the Acoustical Society of America **63**, 925–935. 8

Simpson, A., Hersbach, A. A., and McDermott, H. J. (**2005**). "Improvements in speech perception with an experimental non-linear frequency compression hearing device.", International Journal of Audiology **44**, 281–292. 12

Simpson, A., Hersbach, A. A., and McDermott, H. J. (**2006**). "Frequency-compression outcomes in listeners with steeply sloping audiograms.", International Journal of Audiology **45**, 619–629. 12

Sims, V. K. and Chin, M. G. (**2002**). "Responsiveness and perceived intelligence as predictors of speech addressed to cats", Anthrozoos: A Multidisciplinary Journal of The Interactions of People & Animals **15**, 166–177. 40, 42, 241

Sininger, Y. S., Grimes, A., and Christensen, E. (**2010**). “Auditory development in early amplified children: Factors influencing auditory-based communication outcomes in children with hearing loss”, Ear and Hearing **31**, 166–185. 10

Skelt, L. (**2010**). ““Are You Looking at Me?” The Influence of Gaze on Frequent Conversation Partners’ Management of Interaction with Adults with Acquired Hearing Impairment”, Seminars in Hearing **31**, 116–126. 41

Skelt, L. (**2011**). “Dealing with misunderstandings: The sensitivity of repair in hearing-impaired conversation (chapter 7).”, in *Hearing Aids Communication: Integrating Social Interaction, Audiology and User Centered Design to Improve Communication with Hearing Loss and Hearing Technologies.*, edited by M. Egbert and A. Deppermann, 56–66 (Verlag für Gesprächsforschung). 245

Skelt, L. M. (**2013**). “Recipient gaze and the resolution of overlapping talk in hearing impaired interaction”, Journal of Interactional Research in Communication Disorders **4**, 71–94. 41

Skoruppa, K. and Rosen, S. (**2014**). “Processing of Phonological Variation in Children With Hearing Loss: Compensation for English Place Assimilation in Connected Speech”, Journal of Speech Language and Hearing Research **57**, 1127–1134. 12

Smiljanić, R. and Bradlow, A. R. (**2005**). “Production and perception of clear speech in Croatian and English”, The Journal of the Acoustical Society of America **118**, 1677–1688. 34, 44, 136, 178, 237, 238

Smiljanić, R. and Bradlow, A. R. (**2008**a). “Stability of temporal contrasts across speaking styles in English and Croatian”, Journal of Phonetics **36**, 91–113. 44, 45, 238

Smiljanić, R. and Bradlow, A. R. (**2008**b). “Temporal organization of English clear and conversational speech”, The Journal of the Acoustical Society of America **124**, 3171–3182. 34, 35

Smiljanić, R. and Bradlow, A. R. (**2009**). "Speaking and Hearing Clearly: Talker and Listener Factors in Speaking Style Changes", Language and Linguistics Compass **3**, 236–264. 32, 34, 136, 198

Smiljanić, R. and Bradlow, A. R. (**2011**). "Bidirectional clear speech perception benefit for native and high-proficiency non-native talkers and listeners: Intelligibility and accentedness", The Journal of the Acoustical Society of America **130**, 4020–4031. 140

Smiljanić, R. and Sladen, D. (**2013**). "Acoustic and Semantic Enhancements for Children With Cochlear Implants", Journal of Speech Language and Hearing Research **56**, 1085–1096. 15, 35

Smith, A. and Goffman, L. (**1998**). "Stability and patterning of speech movement sequences in children and adults.", Journal of Speech Language and Hearing Research **41**, 18–30. 16, 17, 55

Smith, L. A. and Williams, J. M. (**2001**). "Children's understanding of the physical, cognitive and social consequences of impairments", Child: care, health and development **27**, 603–617. 50

Smith, L. A. and Williams, J. M. (**2005**). "Developmental differences in understanding the causes, controllability and chronicity of disabilities", Child: care, health and development **31**, 479–488. 50

Smith, N. A. and Trainor, L. J. (**2008**). "Infant-directed speech is modulated by infant feedback", Infancy **13**, 410–420. 41, 42, 241

Snow, C. E. (**1972**). "Mothers speech to children learning language", Child Development **43**, 549–565. 39

Snow, D. and Ertmer, D. (**2009**). "The development of intonation in young children with cochlear implants: a preliminary study of the influence of age at implantation and length of implant experience", Clinical Linguistics & Phonetics **23**, 665–679. 22, 75

Soares, A. P., Machado, J., Costa, A., Iriarte, Á., Simões, A., de Almeida, J. J., Comesaña, M., and Perea, M. (**2014**). "On the advantages of word frequency and contextual diversity measures extracted from subtitles: The case of Portuguese", The Quarterly Journal of Experimental Psychology **68**, 680–696. 184

Sonnenschein, S. (**1988**). "The development of referential communication: Speaking to different listeners", Child Development 694–702. 59

Sparrow, K. and Hird, K. (**2010**). "The Effectiveness of Communication Strategy Training with Adult Cochlear Implantees", Seminars in Hearing **31**, 165–176. 42

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. 15, 183

Spencer, P., Koester, L. S., and 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**, 512–518. 56, 57, 67

Spencer, P. E. (**2003**). "Parent-child interaction", in *The Young Deaf or Hard of Hearing Child: A Family-Centred Approach to Early Education.*, edited by B. Bodner-Johnson and M. Sass-Lehrer (Brookes Publishing). 40

Spilton, D. and Lee, L. C. (**1977**). "Some determinants of effective communication in four-year-olds", Child Development 968–977. 52

Stacey, P. C., Fortnum, H. M., Barton, G. R., and Summerfield, A. Q. (**2006**). "Hearing-impaired children in the United Kingdom, I: Auditory performance, communication skills, educational achievements, quality of life, and cochlear implantation", Ear and Hearing **27**, 161–186. 5

Stathopoulos, E. T., Huber, J. E., and Sussman, J. E. (**2011**). "Changes in Acoustic Characteristics of the Voice Across the Life Span: Measures From Individuals 4–93 Years of Age", Journal of Speech Language and Hearing Research **54**, 1011–1021. 17

Stathopoulos, E. T. and Sapienza, C. M. (**1997**). "Developmental changes in laryngeal and respiratory function with variations in sound pressure level", Journal of Speech Language and Hearing Research **40**, 595–614. 17

Stelmachowicz, P. G., Pittman, A. L., Hoover, B. M., and Lewis, D. E. (**2002**). "Aided perception of/s/and/z/by hearing-impaired children", Ear and Hearing **23**, 316–324. 12, 80

Stelmachowicz, P. G., Pittman, A. L., Hoover, B. M., Lewis, D. E., and 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**, 556–562. 5, 10, 12, 26, 80

Stoel-Gammon, C. (**2011**). "Relationships between lexical and phonological development in young children", Journal of Child Language **38**, 1–34. 16

Streiner, D. L. and Norman, G. R. (**2002**). *Health Measurement Scales: A Practical Guide to their Development and Use, 3rd Edition.* (Oxford Medical Publications., Oxford). 202

Sundqvist, A., Lyxell, B., Jönsson, R., and Heimann, M. (**2014**). "International Journal of Pediatric Otorhinolaryngology", International Journal of Pediatric Otorhinolaryngology **78**, 538–544. 61

Svirsky, M. A., Teoh, S.-W., and Neuburger, H. (**2004**). "Development of Language and Speech Perception in Congenitally, Profoundly Deaf Children as a Function of Age at Cochlear Implantation", Audiology and Neuro-Otology **9**, 224–233. 11

Syrett, K. and Kawahara, S. (**2014**). "Production and perception of listener-oriented clear speech in child language", Journal of Child Language **41**, 1373–1389. 58, 59, 60, 140, 175, 213, 214, 236, 242

Szagun, G. (**2004**). "Learning by ear: on the acquisition of case and gender marking by German-speaking children with normal hearing and with cochlear implants.", Journal of Child Language **31**, 1–30. 27

Szagun, G. and Stumper, B. (**2013**). "Age or Experience? The Influence of Age at Implantation and Social and Linguistic Environment on Language Development in Children With Cochlear Implants.", Journal of Speech, Language and Hearing Research **55**, 1640–1654. 11

Tamayo, J. M. (**1987**). "Frequency of use as a measure of word difficulty in bilingual vocabulary test construction and translation", Educational and Psychological Measurement **47**, 893–902. 182

Taylor, R. and Paisley, S. (**2000**). "The clinical and cost effectiveness of advances in hearing aid technology report to the national institute for clinical excellence", URL `http://www.gserve.nice.org.uk/nicemedia/pdf/hearing_hta_report.pdf`. 5

The GIMP Team (**2012**). *GIMP: GNU Image Manipulation Program*, URL `http://www.gimp.org/`. 81

Titze, I. R. (**1989**). "On the relation between subglottal pressure and fundamental frequency in phonation", The Journal of the Acoustical Society of America **85**, 901–906. 31, 123, 205

Tjaden, K., Lam, J., and Wilding, G. (**2013**). "Vowel Acoustics in Parkinson's Disease and Multiple Sclerosis: Comparison of Clear, Loud, and Slow Speaking Conditions", Journal of Speech Language and Hearing Research **56**, 1485–1502. 68

Tjaden, K., Sussman, J. E., and Wilding, G. E. (**2014**). "Impact of Clear, Loud, and Slow Speech on Scaled Intelligibility and Speech Severity in Parkinson's Disease and Multiple Sclerosis", Journal of Speech Language and Hearing Research **57**, 779–792. 68, 235

Tjaden, K. and Wilding, G. E. (**2004**). "Rate and Loudness Manipulations in DysarthriaAcoustic and Perceptual Findings", Journal of Speech Language and Hearing Research **47**, 766–783. 68, 140

Tobey, E. A., Geers, A. E., Brenner, C., Altuna, D., and Gabbert, G. (**2003**). "Factors associated with development of speech production skills in children implanted by age five.", Ear and Hearing **24**, 36S–45S. 25

Tobey, E. A., Geers, A. E., Douek, B. M., Perrin, J., Skellet, R., Brenner, C., and Toretta, G. (**2000**). "Factors associated with speech intelligibility in children with cochlear implants", The annals of otology, rhinology and laryngology. Supplement **185**, 28–30. 21

Tobey, E. A., Geers, A. E., Sundarrajan, M., and Lane, J. (**2011**). "Factors influencing speech production in elementary and high-school aged cochlear implant users", Ear and Hearing **32**, 27S–38S. 21

Tomblin, J. B., Peng, S.-C., Spencer, L. J., and Lu, N. (**2008**). "Long-term trajectories of the development of speech sound production in pediatric cochlear implant recipients", Journal of Speech Language and Hearing Research **51**, 1353–1368. 25

Tsui, I. Y. H. and 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**, 507–525. 12, 13, 23, 179

Tye-Murray, N. (**2003**). "Conversational Fluency of Children Who Use Cochlear Implants", Ear and Hearing **24**, 82S–89S. 62, 242

Tye-Murray, N., Spencer, L., and Woodworth, G. G. (**1995**). "Acquisition of speech by children who have prolonged cochlear implant experience", Journal of Speech Language and Hearing Research **38**, 327–337. 14, 19, 21

Tye-Murray, N. and Witt, S. A. (**1996**). "Conversational moves and conversational styles of adult cochlear-implant users", Journal of the Academy of Rehabilitative Audiology **29**, 11–25. 196

Uchanski, R. M. (**1988**). "Spectral and temporal contributions to speech clarity for hearing impaired listeners.", Ph.D. thesis, Massachusetts Institute of Technology (MIT). 177

Uchanski, R. M., Choi, S. S., Braida, L. D., Reed, C. M., and Durlach, N. I. (**1996**). "Speaking Clearly for the Hard of Hearing IVFurther Studies of the Role of Speaking Rate", Journal of Speech Language and Hearing Research **39**, 494–509. 45

Uchanski, R. M. and Geers, A. E. (**2003**). "Acoustic Characteristics of the Speech of Young Cochlear Implant Users: A Comparison with Normal-Hearing Age-Mates", Ear and Hearing **24**, 90S–105S. 21, 23, 25, 26, 66, 74, 80, 111, 139, 143, 179, 180

Uther, M., Knoll, M. A., and Burnham, D. (**2007**). "Do you speak E-NG-L-I-SH? A comparison of foreigner- and infant-directed speech", Speech Communication **49**, 2–7. 38, 39

Uziel, A. S., Sillon, M., Vieu, A., Artieres, F., Piron, J.-P., Daures, J.-P., and Mondain, M. (**2007**). "Ten-year follow-up of a consecutive series of children with multichannel cochlear implants", Otology & neurotology **28**, 615–628. 11, 15, 183, 195

van Buuren, R. A., Festen, J. M., and Plomp, R. (**1995**). "Evaluation of a wide range of amplitude-frequency responses for the hearing impaired", Journal of Speech Language and Hearing Research **38**, 211–221. 199

Van Engen, K. J., Baese-Berk, M., Baker, R. E., Choi, A., Kim, M., and Bradlow, A. R. (**2010**). "The Wildcat Corpus of Native-and Foreign-accented English: Communicative Efficiency across Conversational Dyads with Varying Language Alignment Profiles", Language and Speech **53**, 510–540. 77, 107, 293

Van Engen, K. J., Chandrasekaran, B., and Smiljanić, R. (**2012**). "Effects of Speech Clarity on Recognition Memory for Spoken Sentences", PLoS ONE **7**, 1–8. 34, 35, 136

van Heuven, W. J. B., Mandera, P., Keuleers, E., and Brysbaert, M. (**2014**). "SUBTLEX-UK: A new and improved word frequency database for British English.", Quarterly Journal of Experimental Psychology **67**, 1176–1190. 184

van Rooij, J. C. and Plomp, R. (**1991**). "The effect of linguistic entropy on speech perception in noise in young and elderly listeners", The Journal of the Acoustical Society of America **90**, 2985–2991. 33

Vandam, M., Ide-Helvie, D., and Moeller, M. P. (**2011**). "Point vowel duration in children with hearing aids and cochlear implants at 4 and 5 years of age", Clinical Linguistics & Phonetics **25**, 689–704. 18, 24, 26, 74, 180

Vandell, D. L., Anderson, L. D., Ehrhardt, G., and Wilson, K. S. (**1982**). "Integrating hearing and deaf preschoolers: An attempt to enhance hearing children's interactions with deaf peers", Child Development 1354–1363. 56, 57

Vandell, D. L. and George, L. B. (**1981**). "Social interaction in hearing and deaf preschoolers: Successes and failures in initiations", Child Development 627–635. 56, 57

Vihman, M. M. and de Boysson-Bardies, B. (**1994**). "The nature and origins of ambient language influence on infant vocal production and early words", Phonetica **51**, 159–169. 16

Vohr, B., Jodoin-Krauzyk, J., Tucker, R., Johnson, M. J., Topol, D., and Ahlgren, M. (**2008**). "Early Language Outcomes of Early-Identified Infants With Permanent Hearing Loss at 12 to 16 Months of Age", PEDIATRICS **122**, 535–544. 10

Volden, J., Magill-Evans, J., Goulden, K., and Clarke, M. (**2006**). "Varying Language Register According to Listener Needs in Speakers with Autism Spectrum Disorder", Journal of Autism and Developmental Disorders **37**, 1139–1154. 36, 68

von Hapsburg, D. and Davis, B. L. (**2006**). "Auditory sensitivity and the prelinguistic vocalizations of early-amplified infants", Journal of Speech Language and Hearing Research **49**, 809–822. 20

Vongpaisal, T., Trehub, S. E., Schellenberg, E. G., van Lieshout, P., and Papsin, B. (**2010**). "Children with cochlear implants recognize their mother s voice.", Ear & Hearing **31**, 555–566. 12

Vorperian, H. K. and Kent, R. D. (**2007**). "Vowel acoustic space development in children: A synthesis of acoustic and anatomic data", Journal of Speech Language and Hearing Research **50**, 1510–1545. 18

Vorperian, H. K., Kent, R. D., Lindstrom, M. J., Kalina, C. M., Gentry, L. R., and Yandell, B. S. (**2005**). "Development of vocal tract length during early childhood: A magnetic resonance imaging study", The Journal of the Acoustical Society of America **117**, 338–350. 16

Vosoughi, S. and Roy, D. (**2012**). "A longitudinal study of prosodic exaggeration in child-directed speech", in *International Conference on Speech Prosody*, 1–4 (Shanghai, China). 39

Vouloumanos, A. and Werker, J. (**2004**). "Tuned to the signal: The privileged status of speech for young infants.", Developmental Science **7**, 270–276. 7

Wagner, P., Trouvain, J., and Zimmerer, F. (**2015**). "In defence of stylistic diversity in speech research.", Journal of Phonetics **48**, 1–12. 35

Wake, M., Hughes, E. K., Poulakis, Z., Collins, C., and Rickards, F. W. (**2004**). "Outcomes of Children with Mild-Profound Congenital Hearing Loss at 7 to 8 Years: A Population Study", Ear and Hearing **25**, 1–8. 21, 26, 66, 183

Wake, M., Poulakis, Z., Hughes, E. K., Carey-Sargeant, C., and Rickards, F. W. (**2005**). "Hearing impairment: a population study of age at diagnosis, severity, and language outcomes at 7-8 years", Archives of Disease in Childhood **90**, 238–244. 26, 183

Walsh, B. and Smith, A. (**2002**). "Articulatory movements in adolescents: evidence for pro- tracted development of speech motor control processes.", Journal of Speech Language and Hearing Research **45**, 1119–1133. 16, 18, 55

Warren-Leubecker, A. and Bohannon, J. N. (**1982**). "The effects of expectation and feedback on speech to foreigners", Journal of Psycholinguistic Research **11**, 207–215. 39, 42, 217, 241

Warren-Leubecker, A. and Bohannon, J. N. (**1983**). "The effects of verbal feedback and listener type on the speech of preschool children", Journal of Experimental Child Psychology **35**, 540–548. 57, 59, 217, 236

Wassink, A. B., Wright, R. A., and Franklin, A. D. (**2007**). "Intraspeaker variability in vowel production: An investigation of motherese, hyperspeech, and Lombard speech in Jamaican speakers", Journal of Phonetics **35**, 363–379. 44

Wellman, H. M., Cross, D., and Watson, J. (**2001**). "Metaanalysis of theory-ofmind development: the truth about false belief", Child Development **72**, 655–684. 48

Wellman, H. M. and Lempers, J. D. (**1977**). "The naturalistic communicative abilities of two-year-olds", Child Development 1052–1057. 51

Weppelman, T. L., Bostow, A., Schiffer, R., Elbert-Perez, E., and Newman, R. S. (**2003**). "Children's use of the prosodic characteristics of infant-directed speech", Language & Communication **23**, 63–80. 57, 59, 140

Werker, J. and Tees, R. (**1984**). "Cross-language speech perception: Evidence for perceptual reorganization during the first year of life.", Infant Behavior and Development **7**, 49–63. 8

Wester, M., García Lecumberri, M. L., and Cooke, M. (**2014**). "DIAPIX-FL: A symmetric corpus of problem-solving dialogues in first and second languages", in *Interspeech 2014*, 1–5. 77

Whitehurst, G. J. (**1976**). "The development of communication: Changes with age and modeling", Child Development 473–482. 53

Whiteside, S. P., Dobbin, R., and Henry, L. (**2003**). "Patterns of variability in voice onset time: a developmental study of motor speech skills in humans", Neuroscience Letters **347**, 29–32. 19

Whitfield, J. A. and Goberman, A. M. (**2014**). "Articulatory-acoustic vowel space: Application to clear speech in individuals with Parkinson's disease", Journal of Communication Disorders **51**, 19–28. 68

Wiggin, M., Sedey, A. L., Awad, R., Bogle, J. M., and Yoshinaga-Itano, C. (**2013**). "Emergence of Consonants in Young Children with Hearing Loss", The Volta Review **113**, 127–148. 24, 25, 75

Wimmer, H. and Perner, J. (**1983**). "Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception.", Cognition **13**, 103–128. 48

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`. 185

Xie, Y.-H., Potmesil, M., and Peters, B. (**2014**). "Children Who Are Deaf or Hard of Hearing in Inclusive Educational Settings: A Literature Review on Interactions With Peers", Journal of Deaf Studies and Deaf Education **19**, 423–437. 6

Xu, L., Chen, X., Lu, H., Zhou, N., Wang, S., Liu, Q., Li, Y., Zhao, X., and Han, D. (**2011**). "Tone perception and production in pediatric cochlear implants users", Acta Oto-laryngologica **131**, 395–398. 22

Xu, N., Burnham, D., Kitamura, C., and Vollmer-Conna, U. (**2013**). "Vowel Hyperarticulation in Parrot-, Dog- and Infant-Directed Speech", Anthrozoos: A Multidisciplinary Journal of The Interactions of People & Animals **26**, 373–380. 40

Yoshinaga-Itano, C., Baca, R. L., and Sedey, A. L. (**2010**). "Describing the trajectory of language development in the presence of severe to profound hearing loss: A closer look at children with cochlear implants versus hearing aids", Otology & neurotology **31**, 1268–1274. 14, 183

Yoshinaga-Itano, C. and Sedey, A. (**2000**). "Speech development of deaf and hardofhearing children in early childhood: Interrelationships with language and hearing", The Volta Review **100**, 181–212. 20, 21

Yoshinaga-Itano, C., Sedey, A., Coulter, D., and Mehl, A. (**1998**). "The language of early- and later-identified children with hearing loss.", Pediatrics **102**, 1161–1171. 10, 21

Young, A. and Tattersall, H. (**2007**). "Universal Newborn Hearing Screening and Early Identification of Deafness: Parents' Responses to Knowing Early and Their Expectations of Child Communication Development", Journal of Deaf Studies and Deaf Education **12**, 209–220. 4

Yule, G. (**1997**). *Referential Communication Tasks* (Lawrence Erlbaum Associates, Mahwah, NJ). 75

Zebrowitz, L. A., Brownlow, S., and Olson, K. (**1992**). "Baby talk to the baby-faced", Journal of Nonverbal Behavior **16**, 143–158. 39, 42

Zeng, F.-G. (**2004**). "Trends in cochlear implants", Trends in amplification **8**, 1–34. 5

Zeng, F.-G. and Djalilian, H. (**2010**). "Hearing impairment (chapter 14).", in *Oxford Handbook of Auditory Science: Hearing.*, edited by D. Moore, P. Fuchs, C. Plack, A. Rees, and A. Palmer (Oxford University Press, Oxford). 3, 4, 5

Zeng, F.-G. and Liu, S. (**2006**). "Speech Perception in Individuals With Auditory Neuropathy", Journal of Speech Language and Hearing Research **49**, 367–380. 35

Ziv, M., Most, T., and Cohen, S. (**2013**). "Understanding of Emotions and False Beliefs Among Hearing Children versus Deaf Children", Journal of Deaf Studies and Deaf Education **18**, 161–174. 61, 62