Speech production and speech production accuracy in young children: hearing and hearing impaired children with a cochlear implant

Spraakproductie en accuraatheid van spraakproductie bij jonge kinderen: normaalhorende kinderen en kinderen met een cochleair implantaat

Proefschrift voorgelegd tot het behalen van de graad van doctor in de Taalkunde aan de Universiteit Antwerpen te verdedigen door

Jolien Faes

Promotor: Prof. dr. Steven Gillis

Antwerpen, 2017
# Table of contents

Acknowledgements .......................................................................................................................... 9  

1. Chapter 1: General introduction ................................................................................................. 11  
   1.1. Phonological development: general introduction ......................................................... 14  
   1.2. Phonological development in children with NH and children with CI ................................ 16  
      1.2.1. From babbling to first word productions ................................................................. 16  
      1.2.2. Comparison of children with NH and children with CI’s first word productions .......... 18  
      1.2.3. Factors affecting children’s first word productions ................................................. 20  
      1.2.4. CI-specific factors affecting language development .............................................. 21  
   1.3. Aims and outline of the dissertation .................................................................................... 23  

2. Chapter 2: Method ....................................................................................................................... 27  
   2.1. Subjects – corpora ............................................................................................................... 29  
      2.1.1. CI group .................................................................................................................. 29  
      2.1.2. NH group ............................................................................................................... 32  
         2.1.2.1. NH longitudinal corpus .................................................................................... 32  
         2.1.2.2. NH cross-sectional corpus .............................................................................. 32  
   2.2. Data collection and transcription ....................................................................................... 36  
   2.3. Statistical analyses ............................................................................................................ 41  

Appendix 1. Overview of the number of words per corpus .......................................................... 42  

3. Chapter 3: Phonemic accuracy development in children with cochlear implants up to five years of age by using Levenshtein distance .............................................. 47  
   3.1. Introduction ....................................................................................................................... 50  
      3.1.1. Measures of phonemic accuracy ............................................................................. 51  
      3.1.2. Phonemic accuracy in children with NH and children with CI ................................ 55  
   3.2. Method ............................................................................................................................... 60  
      3.2.1. Participants .............................................................................................................. 60
4.3.2. Cross-sectional comparisons between children with NH and children with CI .......................................................... 104
   4.3.2.1. PWV ..................................................................... 104
   4.3.2.2. The Variability score ............................................. 106
4.4. Discussion ........................................................................... 111
   4.4.1 Intraword variability in children with CI ..................... 112
   4.4.2 Cross-sectional comparisons of children with CI and age-
        matched children with NH ........................................... 114
   4.4.3 Explanation of the results, clinical implications and future
        research ....................................................................... 118

5. Chapter 5: The effect of word frequency on the phonemic accuracy in
   children with cochlear implants and peers with normal hearing ....... 121
   5.1. Introduction ..................................................................... 123
      5.1.1. Frequency effects in children with NH ..................... 125
      5.1.2. Frequency effects in children with CI ..................... 127
   5.2. Methods .......................................................................... 129
      5.2.1 Participants ................................................................. 129
      5.2.2. Procedures ............................................................... 129
      5.2.3. Data analyses ........................................................... 129
      5.2.4. Statistical analyses ................................................... 131
   5.3. Results .............................................................................. 133
      5.3.1 Longitudinal analysis of children with CI .................. 133
      5.3.2. Cross-sectional comparisons between children with CI and
              children with NH ...................................................... 136
      5.3.3. Relationship between frequency and the type of words .... 138
   5.4. Discussion ......................................................................... 139
      5.4.1. Word frequency and word accuracy in children with NH .. 139
      5.4.2. Word frequency and word accuracy in children with CI .... 140
      5.4.3. Differences between children with CI and children with
              NH ........................................................................ 143
      5.4.4. Conclusion ................................................................. 145
6. Chapter 6: Consonant cluster production in Dutch-speaking children with cochlear implants: a comparison with peers with normal hearing

6.1. Introduction ........................................................................................................................................ 149
   6.1.1. Consonant cluster production in children with NH ................................................................. 150
   6.1.2. Consonant cluster production in children with CI ................................................................. 152
   6.1.3. The current study ....................................................................................................................... 156

6.2. Method .............................................................................................................................................. 157
   6.2.1. Participants ................................................................................................................................. 157
   6.2.2. Procedure .................................................................................................................................... 157
   6.2.3. Amount of data .......................................................................................................................... 157
   6.2.3. Data analyses ............................................................................................................................. 158
   6.2.4. Statistical analyses ...................................................................................................................... 159

6.3. Results ............................................................................................................................................... 161
   6.3.1. Longitudinal analyses of children with CI ............................................................................... 161
      6.3.1.1. The likelihood of CC clusters and the likelihood of accurate CC clusters ...................... 161
      6.3.1.2. Types of inaccuracies: the likelihood of complete deletion, reduction and simplification .... 163
      6.3.1.3. Reduction processes ........................................................................................................... 166
   6.3.2. Comparisons between children with CI and children with NH .................................................. 169
      6.3.2.1. The likelihood of CC clusters and the likelihood of accurate CC clusters ...................... 170
      6.3.2.2. Types of inaccuracies: the likelihood of complete deletion, reduction and simplification .... 172
      6.3.2.3. Reduction processes ........................................................................................................... 174

6.4. Discussion ......................................................................................................................................... 177
   6.4.1. Longitudinal analyses of children with CI .............................................................................. 177
      6.4.1.1. Development with age ........................................................................................................ 178
6.4.1.2. Age at implantation .................................................. 179
6.4.2. Cross-sectional comparisons between children with CI and
children with NH .......................................................... 180
6.4.3. Conclusions .................................................................. 182

7. Chapter 7: Word initial fricative production in children with cochlear
implants and their peers with normal hearing matched on lexicon size
............................................................................................................ 183

7.1. Introduction ............................................................................ 185
7.1.1. Measures of comparison .................................................. 186
7.1.2. Fricative production of children with CI: comparisons with
age-mates with NH .................................................................... 190
7.2. Method .................................................................................. 192
7.2.1. Participants ........................................................................ 192
7.2.2. Data collection and transcription ...................................... 193
7.2.3. Dutch consonant inventory .............................................. 194
7.2.4. Data analyses ...................................................................... 194
7.2.5. Statistical analyses ............................................................ 195
7.3. Results .................................................................................. 196
7.3.1. Descriptive statistics on the substitution patterns .......... 196
7.3.2. Comparisons on chronological age ................................. 197
7.3.3. Comparisons on lexicon size ........................................... 202
7.4. Discussion .............................................................................. 204
7.4.1. Comparisons on chronological age ................................. 204
7.4.2. Comparisons on lexicon size ........................................... 208
7.4.4. Comparison of comparisons ........................................... 209

8. Chapter 8: Syntagmatic and paradigmatic development of cochlear
implanted children in comparison with peers with normal hearing up to
age seven ............................................................................................ 213

8.1. Introduction ............................................................................ 216
8.1.1. Mean length of utterance (MLU) ...................................... 217
8.1.2. Mean size of paradigm (MSP) ........................................... 220
8.2. Method ......................................................................................................................... 222
  8.2.1. Participants ........................................................................................................ 222
  8.2.2. Data collection and transcription ...................................................................... 222
  8.2.3. Language measures and statistical analyses .................................................. 223
8.3. Results ......................................................................................................................... 225
  8.3.1. Mean length of utterance (MLU) ................................................................. 225
  8.3.2. Mean size of paradigm (MSP) ......................................................................... 226
8.4. Discussion .................................................................................................................. 227
  8.4.1. Syntagmatic development ............................................................................. 228
  8.4.2. Paradigmatic development ........................................................................... 230
  8.4.3. Relationship between syntagmatic and paradigmatic development in children with CI ................................................................. 232
8.5. Conclusion .................................................................................................................. 233

9. Chapter 9: Conclusions .............................................................................................. 235
  9.1. The outcomes of the present dissertation ............................................................ 237
    9.1.1. Speech production at the word level ....................................................... 237
    9.1.2. Speech production at the segmental level ............................................... 240
  9.2. Relevance of the dissertation and further research ........................................... 242

10. Summary – Nederlandse samenvatting ..................................................................... 247
  10.1. Summary .............................................................................................................. 249
  10.2. Nederlandse samenvatting ................................................................................. 253

11. References ................................................................................................................... 259
DANKWOORD - ACKNOWLEDGMENTS

Dit doctoraat zou er niet geweest zijn zonder de hulp en ondersteuning van talloze helden. Zij verdienen meer dan wat ook een plaats in het dankwoord van mijn doctoraatsthesis.

Eerst en vooral wil ik mijn promotor Prof. Dr. Steven Gillis bedanken voor zijn toewijding. Hij heeft me in de voorbije jaren telkens bijgestaan wanneer ik het moeilijk had, geluisterd naar mijn vragen en ze op gepaste manier beantwoord, geduldig geweest met mijn ongeduld en mij telkens weer op het juiste spoor gezet wanneer ik het zelf niet meer zag. Zijn feedback, ondersteuning, kennis en hulp zijn onontbeerlijk geweest. Zonder hem zou er dan ook geen sprake geweest zijn van deze doctoraatsthesis. Bedankt voor alles. Verder heeft hij mij de kans gegeven om me ook naast het onderzoekswerk te ontplooien. Ik vond het zeer uitdagend, aangenaam, verhelderend en vooral leuk om ook een deel van de lessen te verzorgen, bachelor- en masterstudenten te begeleiden, etc. Dit zorgde voor de nodige afwisseling.

Verder wil ik Prof. Dr. Sven de Maeyer bedanken om mij wegwijs te maken in de wondere wereld van statistiek. Hij heeft zeer veel betekend voor me de afgelopen vier jaren: hij was er altijd als ik vragen had, zelfs de voor hem meest evidentie. Ook wil ik Joris Gillis van harte bedanken. Hij is de motor achter enkele pythonscripts die ik onmogelijk zelf had kunnen schrijven. Hij heeft dan ook een groot aandeel in de voltooiing van deze dissertatie. Zelfs wanneer hij het zelf erg druk had, vond hij de tijd om bepaalde scripts te debuggen en om mij de nodige informatie te geven. Ontzettend bedankt voor al jullie tijd en moeite, Sven en Joris. Ik apprecieer het enorm.

I would also like to thank the members of my commission and my judge for their time, work and efforts. Thank you for reading my dissertation and attending my PhD defense.

Daarnaast zijn er natuurlijk de CLiPS-collega’s. Er is enorm veel gelachen, maar er was ook plaats voor een serieuze noot als dat nodig was. Ook de discussies, snelle vragen en brainstormmomenten over ons onderzoek op de derde verdieping zelf zullen
me blijven. Tot slot zal ik met weemoed, maar een goed gevoel terugkijken op onze badmintonmiddagen.

Het leven is meer dan werken alleen. Daarom mag een welgemeende dankjewel voor mijn familie, vrienden en kennissen zeker niet ontbreken. Elk op hun eigen manier zorgden zij ervoor dat ik me wist te ontspannen. Alle avondjes uit, alle etentjes, alle treinritten, alle ijsjes, maar ook alle uren samen sporten, dansen, lachen, shoppen, gezelschapspelletjes spelen en gewoon gezellig samenzijn hebben me enorm veel deugd gedaan. Ik koester die momenten en draag ze voor altijd mee in mijn hart. Vier mensen wil ik speciaal in de bloemetjes zetten en dat zijn mijn ouders, mijn broer en Stijn. Zonder hen had ik dit niet tot een goed einde kunnen brengen. Mijn ouders; mama en papa, bedankt om mij alle te kansen te geven, om mij bij te staan in hoogtes en in laagtes, om te zijn wie jullie zijn, om mee kwaad, droevig of teleurgesteld te zijn wanneer ik dat was en om mee blij en euforisch te zijn zelfs wanneer jullie geen idee hadden waar ik het precies over had. Ik kon en kan altijd bij jullie terecht en koester dat. Ik zie jullie graag. Daarnaast verdient mijn broer meer dan een pluim. Bram, bedankt om mee te denken over dingen waar je zelf niet in thuis was, om dingen na te lezen, om talloze discussies te voeren – al dan niet over ons onderzoek-, om de gezellige avonden bij jullie of ons thuis en op café, om tips te geven vanuit je eigen ervaringen en vooral ook om me meer dan eens op weg te helpen bij het programmeren. Het is van ontelbare waarde geweest en zal het altijd zijn. Tot slot, Stijn, liefde van mijn leven: bedankt om er altijd te zijn, om alles vanop de eerste rij mee te maken en om de rust te zijn die ik nodig had. Bedankt om me te motiveren wanneer dat nodig was, om te relativeren wanneer ik dat zelf niet kon, om blij te zijn voor mij, om de sussende woorden, en om mij graag te zien zoals ik ben. Ik hou van je.
Chapter 1
1 General introduction

This dissertation aims to study the phonological development of Dutch-speaking congenitally deaf children with cochlear implants. More specifically, the phonemic accuracy of their spontaneous is compared to that of normally hearing peers as well as their phonemic accuracy. In this sense, this dissertation compares the phonological development of two groups of children that differ in their degree of hearing.

About two out of 1,000 children are born with a hearing impairment of more than 40 dB (De Raeve, 2016). There are different degrees of hearing loss: slight (up to 25 dB), mild (26 – 40 dB), moderate and moderate-severe (41 – 70 dB), severe (71 – 90 dB) and profound hearing loss (more than 90 dB) (Clark, 1981). Considering children with a bilateral hearing loss, about 45% of the children has a severe-to-profound hearing loss (20-25% severe, 25-30% profound) (De Raeve, 2016). In Belgium, neonatal hearing screening has increased the early identification of children with a congenital hearing loss, and, as a result, also the early cochlear implantation of children with a severe-to-profound hearing loss.

Children with a severe-to-profound hearing loss are potential candidates for cochlear implantation, if their hearing loss is due to a deficit in the cochlea, i.e. sensorineural hearing loss, and if they do not benefit from acoustic hearing aids. A cochlear implant consists of an external part and an internal part. In the external part a microphone captures environmental sound, a sound processor transforms the sound into a digital code, which is transmitted to the inner part of the implant. The internal part consists of an electrode array inserted in the cochlea, which converts the digital code into electrical signals. These electrodes then directly stimulate the auditory nerve, which sends the signal to the brain. In other words, a cochlear implant is an electronic device, which takes over the role of the impaired cochlea by directly stimulating the auditory nerve.

After a period of auditory deprivation, cochlear implants provide severe-to-profound hearing impaired children access to sound in general, and hence also to speech sounds. But, the speech signal is degraded when compared to the acoustic speech signal available in normal hearing (Drennan & Rubinstein, 2008; Wilson, 2006). Thus, a cochlear implant only partly restores the hearing of children with a
severe-to-profound hearing loss. Therefore, children with cochlear implants may be considered as children with a moderate hearing loss (Svirsky, Robbins, Kirk, Pisoni & Miyamoto, 2000).

1.1 Phonological development: general introduction

Congenitally deaf children with cochlear implants (henceforth: children with CI) have a delay in auditory experience as compared to normally hearing children (henceforth: children with NH). Their auditory experience starts later due to their severe-to-profound hearing loss and, after implantation, they experience a degraded signal when compared to children with NH (Drennan & Rubinstein, 2008; Wilson, 2006). Nevertheless, the restored access to auditory information has a positive effect on their spoken language development. Their speech perception improves (Tyler, Fryauf-Bertschy, Kelsay, Gantz, Woodworth & Parkinson, 1997) as well as their speech production (e.g. Grogan, Barker, Dettman & Blamey, 1995; Serry & Blamey, 1999). In the present dissertation, the phonological development of children with CI is investigated in comparison to that of peers with NH. Speech production at the word level, i.e. phonemic accuracy, is considered as well as speech production at the segmental level, i.e. phonemic accuracy and the different types of inaccuracies of some specific speech sounds. Section 1.3 presents a detailed overview of the precise aims.

Even though cochlear implantation enhances speech perception and production, studies suggest that reduced language experience and reduced auditory input has an impact on phonological skills and speech production in comparison to children with NH. For children with CI, delays are for instance found in word learning (Houston, Stewart, Moberly, Hollich & Miyamoto, 2012), speech processing skills and efficiency in encoding and maintaining phonological representations (AuBuchon, Pisoni & Kronenberger, 2015; Spencer & Oleson, 2008). For speech production, these skills are crucial. For instance, a less precise phonological representation of a particular word may very well impact the phonemic accuracy of that word in production. If these phonological skills are less developed in children with CI, their speech production may well be affected.
Children’s phonological development, e.g. speech production skills such as phonemic accuracy, has a far-reaching impact on other aspects of children’s development. For instance, good early speech production skills are found to be related to later language development in general, but also to later neurocognitive development such as concentration, the ability to manipulate information in the working memory, etc. (Castellanos, Pisoni, Kronenberger & Beer, 2016). Furthermore, well-developed early speech production skills impact children’s later reading achievements in both children with NH and children with CI (Catts, 1997; Desjardin, Ambrose & Eisenberg, 2009; Spencer, Barker & Tomblin, 2003; Spencer & Oleson, 2008). In other words, when children learn to read, they benefit from well-developed phonological skills.

In addition, phonological development contributes to children’s social and emotional development (Preisler, Tvingstedt & Ahlström, 2002; Wong, Ching, Leigh, Cupples, Button, Marnane, Whitfield, Gunnourie & Martin, 2016). Phonological development includes i.a. the phonemically accurate production of words (henceforth: speech production accuracy). It has been shown that speech production accuracy is closely related to speech intelligibility (Ingram, 2002). Phonemically more accurately produced words are more intelligible to listeners, although intelligibility also depends on fluency, prosody, etc. Speech production accuracy plays an essential role in children’s intelligibility, which is crucial in interaction. Social interactions require good mutual intelligibility in order to avoid misunderstanding and communication breakdown. In parent-child interaction, parents know how their children speak and what they mean. Outsiders, however, have a more difficult job understanding children they are not familiar with. In order to avoid such a communication breakdown, interaction partners often use repair strategies such as repeating words/sentences (initiated by the speaker or by the listener) (Tye-Murray, Witt, Schum & Slobaski, 1994). However, when one remains unintelligible, the conversation will inevitably end. Also in child-peer interaction, a good mutual understanding is essential to avoid communication breakdown. Repeated communication breakdowns in peer interactions might exclude children from peer interactions (Preisler et al., 2002), both on their own initiative (they may engage less in a conversation), and on that of peers (they may unconsciously avoid conversations with these children). Children with CI, who
supposedly have less developed phonological skills, may well be especially vulnerable in their interactions with peers with NH. However, extensive social (peer) interactions are vital to children’s social well-being (Ginsburg, 2007). In other words, individuals well-being depends i.a. on their social interactions, to which speech intelligibility and thus phonemic accuracy are essential.

Early access to auditory input and more extensive language experience is beneficial for developing phonological processing skills and phonological representations of words and sounds (Spencer & Oleson, 2008; Weisleder & Fernald, 2013). This is beneficial for later language development, neurocognitive development, reading development and social well-being. Therefore, it is important to investigate whether children with CI eventually reach similar levels of phonological development as compared to children with NH. In what follows, our aims are expressed in more detail relative to the background of the existing literature.

1.2 Phonological development in children with NH and children with CI

1.2.1 From babbling to first word productions
Speech perception is indispensible for spoken language acquisition. In other words, the linguistic environment has an influence on spoken language acquisition. As far as babbling is concerned, it has been shown that the frequent syllable types and sound patterns of the ambient language are represented in children’s prelexical utterances (Vihman, Ferguson & Elbert, 1986). The same is found in children’s first word productions, which reflect the statistically frequent sound patterns of the ambient language (Velleman & Vihman, 2002). These observations highlight the importance of early auditory input.

The babble of children with NH with a Germanic and/or Romance background typically consist of CV-sequences, with a high proportion of stops, nasals and glides and a frequent use of labial and coronal sounds (e.g. Stoel-Gammon & Cooper, 1984). The same applies to their first word productions. Children with NH’s first word productions typically contain CV-syllables with a high frequency of stops, nasals and glides, labials and coronals (Beers, 1995; Ferguson & Farwell, 1975; Fikkert &
Altvater-Mackensen, 2013; Leonard, Newhoff & Mesalam, 1980; StoelGammon, 1985; Van Severen, 2012). This suggests a continuity between children’s babbles and first words. This continuity is also present at an individual level: a child’s sound pattern preferences in babbling are also reflected in the child’s first words (McCune & Vihman, 2001; Vihman, 1992). In children with CI, babble and first words also contain stops, nasals and glides at a labial and coronal place of articulation (Schramm, Bohnert & Keilmann, 2009). In other words, there is also continuity between babbles and first words in children with CI (Moreno-Torres, 2014).

Radical template phonology (Vihman & Croft, 2007) proposes that children initially have some word templates based on their own production (babbling, also called vocal motor schemes) and the input they receive (input experience). At the beginning of word production and phonological development, children have a limited number of word shapes. They adapt other words in such a way that they fit these templates. For instance, a Dutch child with a word template <pV> adapts his/her word productions to this template: /pu/ for /blum/ flower, pe for /spel/ play and also for /bet/ teddy bear and pa for /part/ horse. Word templates can be seen as a match of the phonetic structure of their own productions and the model in the input (Vihman & Croft, 2007), also for onomatopoeia (Laing, 2014). Children appear to have a preference for particular word templates and adapt their own productions to those templates. Gradually, children develop an expanded range of word templates. According to the A-map model (McAllister Buyn, Inkelas & Rose, 2016), children’s phonological development is guided by two pressures: accuracy and precision. The first one indicates a tendency to match adult production even if this causes performance errors. The second one reflects the tendency to have a stable production, even if this production does not match adult input. Both pressures have a weight in children’s phonology and this weight changes with maturation. In the earliest stages, children’s productions are highly variable and thus accuracy pressure plays the main role. However, after several erroneous productions, the weight of the precision pressure increases. This causes more stable, but erroneous productions, as motor control is not yet optimal. This can also be seen as word templates, based on children’s own production (precision) and the input they receive (accuracy). As
children gain more motor control, the effect of precision diminishes and hence word templates expand and productions become more accurate. Studies suggest that this expansion of word templates first concerns the suprasegmental structure, i.e. the prosodic structure of templates, and later the segmental structure of templates (Ibertsson, Willstedt-Svensson, Radeborg & Sahlén, 2008; Sundström, Samuelsson & Lyxell, 2014). This present dissertation analyses i.a. phonemic accuracy at the word level and phonemic accuracy of word-initial consonant clusters (see section 1.3). Therefore, both suprasegmental and segmental accuracy are taken into account. Suprasegmental performance is related to the prosodic structure of the target words/sounds (e.g. the number of syllables, the number of consonants in the onset), whereas segmental performance is associated with the accurate production of phonemes (words and word-initial clusters).

1.2.2 Comparison of children with NH and children with CI’s first word productions

The developmental trajectory of children with CI’s productions from babble to first words is similar to that of children with NH (Moreno-Torres, 2014; Schramm et al., 2009). However, when compared to peers with NH, it is often reported that children with CI’s phonological development lags behind that of peers with NH. For instance with respect to babbling, Molemans (2011) showed that Dutch-speaking children with CI start to babble later, babble less frequently and their babbles are less complex. Nevertheless, those children with CI had a good opportunity to catch up, as they are early implanted, i.e. before 20 months of age. With respect to word productions, for instance Schauwers, Taelman, Gillis and Govaerts (2008) showed that Dutch-speaking children with CI produce less complex words up to age two-and-a-half: children with CI produce more monosyllables as compared to their peers with NH. In addition, up to that same age, their word productions are less accurate than those of children with NH (Schauwers et al., 2008; Van den Berg, 2012). Similarly, the word productions of English-speaking children with CI are reported to be less complex, less accurate, and more variable when compared to peers with NH up to five years of age (Eriks-Brophy, Gibson & Tucker, 2013; Ertmer & Goffman, 2011; Ertmer, Kloiber, Jung, Kirleis &
Bradford, 2012; Warner-Czyz & Davis, 2008). So far, there is no literature on the variability in speech production of children with CI acquiring Dutch. It may be expected that Dutch-speaking children with CI’s word productions are also more variable as compared to peers with NH similar to the findings for children with CI acquiring English.

Thus, the literature suggests that children with CI’s first word productions are less accurate, more variable and less complex. Children with CI are delayed regarding the phonological aspects of first word productions. A delay of children with CI is also found in other domains such as the development of morphology and syntax (Szagun, 2001). However, for vocabulary, morphology and syntax, Nicholas and Geers (2007) found a catch-up for early implanted English-speaking children with CI at the age of four and a half. Similarly, Boons, De Raeve, Langereis, Peeraer, Wouters and Van Wieringen (2013) found a catch up for morphological and syntactic development in early implanted Dutch-speaking children with CI, and Wie (2010) for receptive and expressive syntax in early implanted Norwegian-speaking children with CI. However, little is known about the long-term phonological development of children with CI thus far. The few long-term studies which have been carried out have focussed on standardised tests instead of spontaneous speech, and children acquiring other languages than Dutch (e.g. for English: Eriks-Brophy et al., 2013). It remains to be seen whether children with CI catch up on their peers with NH for complexity, accuracy, and consistency of spontaneous speech production beyond the age of two.

The present dissertation examines the long-term phonological development of early implanted Dutch-speaking children with CI up to the age of seven. Do those children reach similar levels of phonological development in their spontaneous speech production as compared to their peers with NH? An in-depth discussion of the literature regarding speech production accuracy at the word level and at the segmental level as well as speech production variability is provided, when relevant, in the introduction of each chapter.
1.2.3 Factors affecting children’s first word productions

Children’s first word productions differ in complexity, accuracy and consistency (Ingram, 2002). However, there are several factors, which affect for instance accuracy and consistency of first word productions and the order of phonological and lexical acquisition. In what follows, only those factors are discussed which are relevant for the present dissertation, i.e. phonemic complexity and word frequency. Nevertheless, apart from complexity and frequency of the target words, numerous other factors affect the production of words and sounds such as neighbourhood density, functional load, age of acquisition, etc. (i.a. Newman and German (2002), Sosa and Stoel-Gammon (2012), Storkel (2004) and Stokes and Surendran (2005)).

What is the effect of complexity of target words and sounds on the phonological development? Complexity can be defined in different ways. For instance in terms of the number of phonemes, Storkel (2004) showed that early acquired words are short words. When complexity is defined in terms of the number of syllables in the target word, similar observations are made in children with NH: there is a benefit for less complex words. For instance, children with NH acquire monosyllabic words before multisyllabic ones (Maekawa & Storkel, 2006), target words with fewer syllables are more accurate (Gathercole, Willis, Emslie & Baddeley, 1991; Willis & Gathercole, 2001) and less variable (Sosa, 2015). Moreover, simple sounds are acquired before complex ones, e.g. singleton consonants before consonant clusters (Saaristo-Helin, Kunnari & Savinainen-Makkonen, 2011) or simple vowels before diphthongs (Hua & Dodd, 2000). In children with CI, the complexity of target words is also found to affect their speech production. Children with CI acquire simple vowels and singleton consonants before diphthongs and consonant clusters (Blamey, Barry, Bow, Sarant, Paatsch & Wales, 2001). In addition, target words with more syllables are less accurately produced (Carter, Dillon & Pisoni, 2002; Willstedt-Svensson, Löfqvist, Almqvist & Sahlén, 2004) and target words with increased phonological complexity (i.e. words with fricatives, affricates and liquids) are more variable and less accurate in children with CI (Macrae, 2013). Thus, in both groups of children, the complexity of target words affects production.
A second factor, which is considered in the present dissertation, is the effect of frequency. The frequency of sounds and words in the input as well as the frequency of sounds and words in children’s own productions are found to affect speech production. For instance at the segmental level, the frequency of phoneme sequences in child-directed speech, i.e. phonotactic probability, affects the appearance of those sounds in children’s speech production. Phonemes and phoneme sequences with a higher phonotactic probability in the target language are produced more accurately in children with NH (Edwards, Beckman & Munson, 2004; Munson, 2001). In children with CI, an effect of phonotactic probability on production accuracy is only found in children with bilateral implants (Guo, McGregor & Spencer, 2015; Todd, Edwards & Litovsky, 2011).

In addition to frequency effects at the segmental level, the literature also reports effects of word frequency in both child-directed speech and the child’s own productions. These frequency effects are related to the order of lexical acquisition. In general, it is found that higher word frequency in both child-directed speech and the children’s own productions has a positive effect on lexical acquisition: highly frequent words are acquired earlier (children with NH: Goodman, Dale & Li, 2008; children with CI: Han, Storkel, Lee & Yoshinaga-Itano, 2015). Next to lexical development, it may be expected that word frequency affects phonological development as well. In the present dissertation, we investigate if word frequency also affects phonological development in both children with NH and children with CI. More information on word frequency effects is provided in chapter 5.

1.2.4 CI-specific factors affecting language development
The previous paragraphs broadly discussed the relevant aspects of phonological development of children with NH and children with CI and reviewed some factors affecting language development in both groups of children. However, there are numerous other important factors that affect the language development of children with CI (Gillis, 2017; Szagun & Schramm, 2016), such as communication modality (only oral communication or oral and sign communication), preoperative hearing levels, audiological rehabilitation, age at identification of the hearing impairment, the time at cochlear implantation, length of implant use, etc. In what follows, three factors
General introduction

are discussed: length of implant use, preoperative hearing levels and the time at cochlear implantation.

First of all, studies have clearly demonstrated that children with CI’s language skills improve with length of implant use. Improvements with longer device use have been found for morphology (Blamey et al., 2001; Flipsen & Kangas, 2014; Nicholas & Geers, 2007; Schauweers, 2006), vocabulary (Nicholas & Geers, 2007; Schramm, Bohnert & Keilmann, 2010), sentence complexity (Nicholas & Geers, 2007), phoneme accuracy (Blamey et al., 2001; Eriks-Brophy et al., 2013; Gantz, Tyler, Woodworth, Tye-Murray & Fryauf-Bertschy, 1994; Tye-Murray, Spencer & Woodworth, 1995), speech perception (Tyler et al., 1997) and speech intelligibility (Chin, Tsai & Gao, 2003). Thus, the time of implant use is important when studying children with CI’s language development.

Secondly, it has been shown that the amount of residual hearing before implantation affects the language outcomes after cochlear implantation. Children with more residual hearing before implantation have better comprehension and expression scores after implantation (Niparko, Tobey, Thal, Eisenberg, Wang, Quittner & Fink, 2010). For instance Houston et al. (2012) showed that children with CI with more residual hearing before implantation perform better on novel word learning tasks than children with CI with less or no residual hearing before implantation. Moreover, Szagun (2001) has shown that vocabulary and MLU outcomes after implantation are better if children had more residual hearing before implantation. The residual hearing in the children with CI of the present dissertation was well below the speech level. More information is presented in the method section (chapter 2).

Finally, early implantation is beneficial for the speech and language outcomes of children with CI (Connor, Hieber, Arts & Zwolan, 2000). Nevertheless, there is no consensus on the exact age. For instance Dettman, Dowell, Choo, Arnott, Abrahams, Davis, Dronan, Leigh, Constantinescu, Cowan and Briggs (2016) and Levine, Stother-Garcia, Golinkhoff and Hirsh-Pasek (2016) support cochlear implantation before children’s first birthday, whereas for instance Schorr, Fox, van Wassenhove and Knudsen (2005) found better outcomes for children implanted before the age of two and a half in comparison with children implanted later. Nevertheless, it is assumed
that cochlear implantation approximately before the age of four results in better language outcomes. Early implantation (before the age of four), has been shown to be beneficial for grammatical development (Boons et al., 2013; Nikolopoulos, Dyar, Archbold & O'Donoghue, 2004), speech production (Leigh, Detmman, Dowell & Briggs, 2013), speech production accuracy (Connor, Craig, Raudenbush, Heavner & Zwolan, 2006), speech perception (Nikolopoulos, O'Donoghue & Archbold, 1999; Svirsky, Teoh & Neuburger, 2003), receptive vocabulary (Schorr, Roth & Fox, 2008), novel word learning (Houston et al., 2012), speech intelligibility (Nikolopoulos et al., 1999), phonological memory (Schorr et al., 2008) and even for the auditory-visual fusion on speech comprehension (Schorr et al., 2005). The children with CI in the present dissertation were all implanted before the age of 20 months. More information is provided in the method section.

1.3 Aims and outline of the dissertation

This chapter has provided a general introduction of the topic in this dissertation, i.e. speech production and production accuracy in children with cochlear implants. In each individual chapter, a more in-depth discussion of the relevant literature will be provided. This dissertation is a compilation of published papers. As a consequence, there will be some overlap between the different chapters. In order to avoid some of the reduplication, chapter 2 gives a general overview of the methodology, concerning the participants, the data collection and the data transcription. Nevertheless, each chapter has some chapter-specific methods as well, which are discussed when appropriate.

Which specific research questions are addressed in this study? The chapters can be seen in two clusters. The first cluster of chapters aims to examine the effect of some word-related factors influencing speech production and speech production accuracy (chapters 3 – 5). More specifically, whole-word accuracy and variability are examined and the possible influences of particular target word characteristics are investigated. The second cluster of chapters investigates the development of more specific speech sounds (chapters 6 – 8). Each chapter assesses whether there are differences between
General introduction

children with NH and children with CI and examines whether children with CI are able to catch up with their age-matched peers with NH (long-term approach).

First, it is examined whether certain word-related characteristics have an influence on their production. Chapter 3 investigates whether syllable length and complexity affect phonemic accuracy in children with CI. Thus, are longer words and more complex words less accurate than shorter and less complex ones? A relatively new methodology is used to measure whole-word phonemic accuracy, namely Levenshtein distance. The Levenshtein distance between each adult target and child replica is calculated. Hence, Levenshtein distance gives an indication of the phonemic accuracy of children’s word productions. A large distance indicates a less accurate production and a smaller distance a more accurate production. Children with CI are compared to peers with NH from word onset up to five years of age. Moreover, the effects of age at implant activation and length of device use in children with CI are considered.

Chapter 4 addresses the question whether variability in word productions is similar in children with CI and peers with NH up to the age of five. Therefore, the different child productions of each target word are listed per age and per child. The proportion of variability and different variability profiles (consistent correct, consistent incorrect, variable with hits, variable with no hits) are calculated. In addition, it is investigate whether longer words are more variable than shorter ones. Thus, the effect of word length on the variability in production is examined.

Chapter 5 explores the effect of input frequency on phonemic accuracy is explored. Are more frequent words in child-directed speech produced more accurately by children themselves or not? And, is this effect similar in both groups of children? As in chapter 3, phonemic accuracy is measured using Levenshtein distance, which considers whole-word accuracy. Comparisons between children with NH and children with CI range between word onset and seven years of age.

Secondly, the production and accuracy development of two specific groups of speech sounds is examined. While chapters 3 and 4 show that complexity has an effect on production (accuracy and variability), chapters 6 and 7 examine the development of two specific groups of speech sounds often reported as complex sounds, i.e. consonant clusters and fricatives. In chapter 6, the development of word-initial consonant
clusters in children with CI is studied and compared to that of children with NH. The aim is threefold: study (1) the frequency of consonant clusters in both groups of children, (2) the accuracy of consonant clusters in both groups of children and (3) the development of inaccurate productions. Moreover, the types of inaccuracies and their development over time will be scrutinized. Analyses are performed from word onset up to seven years of age.

Chapter 7 studies the development of fricatives in both groups of children. The aim is to study (1) their frequency, (2) their accuracy and (3) the development of inaccurate productions. Moreover, chapter 7 also evaluates the possibility of comparing children with CI and children with NH on lexical age rather than on chronological age. As this can only be tested by means of longitudinal data, only the production of fricatives between word onset and two and a half years of age is considered. Moreover, the outcomes in this chapter indicate that fricatives hardly ever occur in children with CI at the age of 24 months. In the literature, it is hypothesized that the presence of a single fricative at the age of 18 months is related to better expressive language development in vocabulary, sentence length and grammar. Vice versa, this would suggest that sentence length and grammatical development are problematic in children with CI as fricatives are missing in their speech, even at the age of 24 months. In chapter 8, this hypothesis is considered: the paradigmatic (inflection) and syntagmatic (sentence length) development of children with CI is compared to that of children with NH between ages two and seven.

Finally, chapter 9 presents a general conclusion and some directions for further research, followed by a Dutch summary in chapter 10.
Chapter 2

Chapter 2
2 Method

2.1 Subjects – corpora

This present dissertation compares two groups of children: a group of congenitally deaf children with a cochlear implant (CI) and a group of age-matched children with normal hearing (NH). All children lived in middle-to-high class families in Flanders, i.e. the northern part of Belgium, and are monolingual Dutch. In the next paragraphs, details for each group of children are elaborated.

2.1.1 CI group

The CI corpus used in this thesis is collected and transcribed by Schauwers (2006). Ten children with CI (3 boys and 7 girls) were followed longitudinally from the moment their device was activated up to age seven: monthly from implant activation up to 30 months after implantation and yearly between three and seven years of age. One child (S10) dropped out of the study at the age of 25 months and was therefore only included when comparisons up to that age are made, resulting in a corpus of 9 children with CI when comparing with NH children between three and seven. Children with CI were raised orally and used only a limited amount of lexical signs. Table 1 gives an overview of the CI participants. All children had a congenitally profound hearing loss with a mean Pure Tone Average (PTA) before implantation of 113.00 (SD = 8.72) decibels Hearing Level (dB HL) and 112.56 (SD = 9.13) dB HL when S10 is not included. Thresholds are given in the better ear and were tested at 500, 1000 and 2000 Hertz. The causes of deafness are a mutation in the connexin-26 gene (S1, S2, S4, S5 and S6), a mutation in the connexin-31 gene (S9), a cytomegalovirus infection (S3), genetic (S7) and unknown (S8 and S10). No other health or developmental problems are reported during data collection. Hearing impairment was detected in a neonatal hearing-screening program (Philips, Corthals, De Raeve, D’haenens, Maes, Bockstael, Keppler, Swinnen, De Vel, Vinck & Dhooge, 2009). As part of the follow-up of the initial diagnosis, all children were fitted with bilateral hearing aids within 1 to 5 months after the detection of the hearing loss and they wore those devices for several months. The mean age at fitting of the hearing aids was 4.28 months (SD = 3.12) and 4.40 months (SD = 3.28) excluding S10. With
Method

hearing aids, the mean PTA improved to 96.70 dB HL (SD = 28.70) for all children and to 94.67 (SD = 29.67) dB HL for all children except S10. These thresholds remained all well below speech level. Consequently all children were implanted with a Nucleus-24 implant before 20 months of age. The mean age at implantation was 12.05 months (SD = 4.96) for all children with CI and 11.92 months (SD = 5.25) when S10 is not included. None of the children had contralateral hearing with a hearing aid. About a month after implantation the cochlear implant was fitted. The mean age at implant activation was 13.21 months (SD = 5.10) and 13.12 months (SD = 5.40) when S10 is not included. After implantation, the PTA improved in all children. At age two the mean PTA of all children improved to 40.10 dB HL (SD = 8.24) and 39.78 dB HL (SD = 8.67) without S10. At age five the mean PTA of all children – except S10 of which no data are available – even improved to 32.33 dB HL (SD = 7.11). The mean age at which the first words appeared was 18.50 months (SD = 2.92) including S10 and 18.11 months (SD = 2.80) excluding S10. Words were identified in accordance to Vihman and McCune (1994). The median age at word onset was 19 months (range 15 – 23 months) including S10 and 18 months (range 15 – 23 months) excluding S10. Six out of ten children received a second CI during data collection. Implant types of the second CI were Nucleus-24 (S7 and S9), Nucleus Freedom (S3 and S5), Digisonic SP20 (S2) and unknown (S1). Mean age at second implantation is 52.50 months (SD = 27.03). Table 1 gives an overview of all the characteristics of each. All children received speech and language therapy twice a week. All information presented in this paragraph and more details can be found in Schauwers (2006).
<table>
<thead>
<tr>
<th>ID</th>
<th>Gender</th>
<th>PTA&lt;sup&gt;a&lt;/sup&gt; unaided</th>
<th>PTA&lt;sup&gt;a&lt;/sup&gt; HA&lt;sup&gt;b&lt;/sup&gt;</th>
<th>PTA&lt;sup&gt;a&lt;/sup&gt; CI&lt;sup&gt;c&lt;/sup&gt; (Age&lt;sup&gt;c&lt;/sup&gt; 24)</th>
<th>PTA&lt;sup&gt;a&lt;/sup&gt; CI&lt;sup&gt;c&lt;/sup&gt; (Age&lt;sup&gt;c&lt;/sup&gt; 60)</th>
<th>Age&lt;sup&gt;e&lt;/sup&gt; HA&lt;sup&gt;b&lt;/sup&gt; fitting</th>
<th>Age&lt;sup&gt;e&lt;/sup&gt; 1&lt;sup&gt;st&lt;/sup&gt; CI</th>
<th>Age activation&lt;sup&gt;e&lt;/sup&gt; 1&lt;sup&gt;st&lt;/sup&gt; CI</th>
<th>Age&lt;sup&gt;e&lt;/sup&gt; 2&lt;sup&gt;nd&lt;/sup&gt; CI</th>
<th>Age&lt;sup&gt;e&lt;/sup&gt; first word</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>F</td>
<td>120</td>
<td>120</td>
<td>48</td>
<td>35</td>
<td>9.10</td>
<td>13.49</td>
<td>14.89</td>
<td>75.00</td>
<td>20.00</td>
</tr>
<tr>
<td>S2</td>
<td>F</td>
<td>120</td>
<td>120</td>
<td>30</td>
<td>27</td>
<td>1.13</td>
<td>6.69</td>
<td>7.66</td>
<td>56.00</td>
<td>16.00</td>
</tr>
<tr>
<td>S3</td>
<td>F</td>
<td>115</td>
<td>113</td>
<td>33</td>
<td>25</td>
<td>1.59</td>
<td>10.00</td>
<td>11.66</td>
<td>70.00</td>
<td>20.00</td>
</tr>
<tr>
<td>S4</td>
<td>M</td>
<td>113</td>
<td>117</td>
<td>48</td>
<td>42</td>
<td>10.00</td>
<td>18.16</td>
<td>19.34</td>
<td>-</td>
<td>20.00</td>
</tr>
<tr>
<td>S5</td>
<td>M</td>
<td>93</td>
<td>47</td>
<td>38</td>
<td>32</td>
<td>4.79</td>
<td>16.89</td>
<td>17.89</td>
<td>76.00</td>
<td>18.00</td>
</tr>
<tr>
<td>S6</td>
<td>M</td>
<td>120</td>
<td>107</td>
<td>53</td>
<td>37</td>
<td>1.69</td>
<td>8.76</td>
<td>9.66</td>
<td>-</td>
<td>16.00</td>
</tr>
<tr>
<td>S7</td>
<td>F</td>
<td>117</td>
<td>107</td>
<td>42</td>
<td>23</td>
<td>4.00</td>
<td>5.16</td>
<td>6.13</td>
<td>15.00</td>
<td>15.00</td>
</tr>
<tr>
<td>S8</td>
<td>F</td>
<td>112</td>
<td>58</td>
<td>38</td>
<td>42</td>
<td>2.00</td>
<td>19.46</td>
<td>21.13</td>
<td>-</td>
<td>23.00</td>
</tr>
<tr>
<td>S9</td>
<td>F</td>
<td>103</td>
<td>63</td>
<td>28</td>
<td>28</td>
<td>5.26</td>
<td>8.69</td>
<td>9.69</td>
<td>23.00</td>
<td>15.00</td>
</tr>
<tr>
<td>S10</td>
<td>F</td>
<td>91 ↓ 117</td>
<td>45 ↓ 115</td>
<td>43</td>
<td>N.A.</td>
<td>3.20</td>
<td>13.23</td>
<td>14.13</td>
<td>-</td>
<td>22.00</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>113.00</td>
<td>96.70</td>
<td>40.10</td>
<td>32.33</td>
<td>4.28</td>
<td>12.05</td>
<td>13.21</td>
<td>52.50</td>
<td>18.50</td>
</tr>
<tr>
<td>SD&lt;sup&gt;d&lt;/sup&gt;</td>
<td></td>
<td>8.72</td>
<td>28.70</td>
<td>8.24</td>
<td>7.11</td>
<td>3.12</td>
<td>4.96</td>
<td>5.10</td>
<td>27.03</td>
<td>2.92</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>112.56</td>
<td>94.67</td>
<td>39.78</td>
<td>32.33</td>
<td>4.40</td>
<td>11.92</td>
<td>13.12</td>
<td>52.50</td>
<td>18.11</td>
</tr>
<tr>
<td>SD&lt;sup&gt;d&lt;/sup&gt;</td>
<td></td>
<td>9.13</td>
<td>29.67</td>
<td>8.67</td>
<td>7.11</td>
<td>3.28</td>
<td>5.25</td>
<td>5.40</td>
<td>27.03</td>
<td>2.80</td>
</tr>
</tbody>
</table>

<sup>a</sup> PTA = Pure Tone Average (in dB HL = decibels Hearing Level)
<sup>b</sup> hearing aids
<sup>c</sup> ↓ = progressive hearing loss
<sup>d</sup> Ages are presented in months
<sup>e</sup> including subject S10, <sup>f</sup> excluding subject S10
N.A. = not available
- = no second CI
Method

2.1.2 NH group
The children of the CI corpus are compared to a group of children with normal hearing (NH). *Kind & Gezin*, i.e. the Flemish infant welfare centre, checked those children’s hearing approximately three weeks after birth. The hearing screening is nation wide and consists of an otoacoustic emission test. The corpus of children with NH is subdivided into two different corpora: (1) a longitudinal corpus of children with NH (henceforth, NHlongitudinal) and (2) a cross-sectional corpus of children with NH (henceforth, NHcross-sectional).

1.1.1.1 NHlongitudinal corpus
The NHlongitudinal corpus was collected and transcribed by Molemans (2011), Van den Berg (2012) and Van Severen (2012) and includes 30 NH children (S11 – S40). The children in this corpus (16 boys and 14 girls) were followed longitudinally and monthly between six months and two years. All children had normal hearing and no health or developmental problems were reported during data collection. The mean age at which the first words appeared was 14.47 months (SD = 0.86) and the median age at first words was 14 months (range 14 - 19 months). None of the children scored less than percentile 1 on the N-CDI test, i.e. the Dutch version of the CDI test (Zink & Lejaegere, 2002). More detailed information of the participants and about the data collection can be found in Molemans (2011), Van den Berg (2012) and Van Severen (2012).

1.1.1.2 NHcross-sectional corpus
A cross-sectional design was set-up for the children with NH between ages two and seven. The NHcross-sectional corpus was also collected by Molemans (2011), Van den Berg (2012) and Van Severen (2012). However, the transcription of this corpus was not yet available and was carried out by the author of this dissertation, as discussed later in the method section (section 2.2). A total of 63 NH children (S41 – S103) participated: 10 two-year-olds (mean = 24.62 months, SD = 1.06 months), 9
three-year-olds (mean = 35.88 months, SD = 1.09 months), 12 four-year-olds (mean = 48.35 months, SD = 1.39 months), 11 five-year-olds (mean = 60.41 months, SD = 1.32), 11 six-year-olds (mean = 72.43 months, SD = 2.80) and 10 seven-years-olds (mean = 83.32 months, SD = 1.74). These children were all recorded only once. All children in this corpus had normal hearing and no health or developmental problems were reported during data collection. Table 2, presents the exact ages and gender of each participant, and Table 3 specifies the exact time of the transcribed video recordings.
Table 2. Age and gender characteristics of the subjects in the NHcross-sectional corpus

<table>
<thead>
<tr>
<th>Two-year-olds</th>
<th>Three-year-olds</th>
<th>Four-year-olds</th>
<th>Five-year-olds</th>
<th>Six-year-olds</th>
<th>Seven-year-olds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Age&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Age&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Age&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Age&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Age&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>S41 F&lt;sup&gt;b&lt;/sup&gt;</td>
<td>24</td>
<td>S51 M</td>
<td>35</td>
<td>S60 F</td>
<td>48</td>
</tr>
<tr>
<td>S42 M&lt;sup&gt;c&lt;/sup&gt;</td>
<td>25</td>
<td>S52 F</td>
<td>35</td>
<td>S61 F</td>
<td>48</td>
</tr>
<tr>
<td>S43 M</td>
<td>25</td>
<td>S53 F</td>
<td>35</td>
<td>S62 M</td>
<td>46</td>
</tr>
<tr>
<td>S44 F</td>
<td>23</td>
<td>S54 M</td>
<td>38</td>
<td>S63 F</td>
<td>50</td>
</tr>
<tr>
<td>S45 M</td>
<td>26</td>
<td>S55 F</td>
<td>37</td>
<td>S64 F</td>
<td>47</td>
</tr>
<tr>
<td>S46 F</td>
<td>24</td>
<td>S56 F</td>
<td>36</td>
<td>S65 M</td>
<td>48</td>
</tr>
<tr>
<td>S47 M</td>
<td>23</td>
<td>S57 F</td>
<td>36</td>
<td>S66 M</td>
<td>47</td>
</tr>
<tr>
<td>S48 F</td>
<td>27</td>
<td>S58 F</td>
<td>35</td>
<td>S67 F</td>
<td>49</td>
</tr>
<tr>
<td>S49 F</td>
<td>25</td>
<td>S59 M</td>
<td>36</td>
<td>S68 F</td>
<td>49</td>
</tr>
<tr>
<td>S50 M</td>
<td>25</td>
<td>S69 M</td>
<td>49</td>
<td>S81 M</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S70 F</td>
<td>49</td>
</tr>
<tr>
<td>Mean</td>
<td>24.62</td>
<td>35.88</td>
<td>48.35</td>
<td>60.41</td>
<td>72.43</td>
</tr>
<tr>
<td>SD</td>
<td>1.06</td>
<td>1.09</td>
<td>1.39</td>
<td>1.32</td>
<td>2.80</td>
</tr>
</tbody>
</table>

<sup>a</sup>Ages are presented in months  
<sup>b</sup>Female,  
<sup>c</sup>Male
Table 3. Exact time in minutes and seconds of transcriptions per subject in the NHcross-sectional corpus

<table>
<thead>
<tr>
<th>Two-year-olds</th>
<th>Three-year-olds</th>
<th>Four-year-olds</th>
<th>Five-year-olds</th>
<th>Six-year-olds</th>
<th>Seven-year-olds</th>
</tr>
</thead>
<tbody>
<tr>
<td>S41</td>
<td>15:39</td>
<td>S51</td>
<td>10:35</td>
<td>S60</td>
<td>10:41</td>
</tr>
<tr>
<td>S42</td>
<td>35:00</td>
<td>S52</td>
<td>10:19</td>
<td>S61</td>
<td>10:34</td>
</tr>
<tr>
<td>S43</td>
<td>Unknown</td>
<td>S53</td>
<td>13:12</td>
<td>S62</td>
<td>11:42</td>
</tr>
<tr>
<td>S44</td>
<td>15:29</td>
<td>S54</td>
<td>07:12</td>
<td>S63</td>
<td>14:36</td>
</tr>
<tr>
<td>S45</td>
<td>Unknown</td>
<td>S55</td>
<td>08:12</td>
<td>S64</td>
<td>10:13</td>
</tr>
<tr>
<td>S46</td>
<td>26:18</td>
<td>S56</td>
<td>09:08</td>
<td>S65</td>
<td>17:40</td>
</tr>
<tr>
<td>S47</td>
<td>13:00</td>
<td>S57</td>
<td>05:09</td>
<td>S66</td>
<td>10:00</td>
</tr>
<tr>
<td>S49</td>
<td>17:03</td>
<td>S59</td>
<td>14:42</td>
<td>S68</td>
<td>18:07</td>
</tr>
<tr>
<td>S50</td>
<td>14:11</td>
<td></td>
<td></td>
<td>S69</td>
<td>12:47</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S70</td>
<td>22:50</td>
<td>S82</td>
<td>13:13</td>
</tr>
<tr>
<td>Mean&lt;sup&gt;a&lt;/sup&gt;</td>
<td>19.54</td>
<td>10.23</td>
<td>15.40</td>
<td>14.12</td>
<td>22:57</td>
</tr>
<tr>
<td>SD&lt;sup&gt;a&lt;/sup&gt;</td>
<td>7.50</td>
<td>3.21</td>
<td>6.06</td>
<td>3.83</td>
<td>8.58</td>
</tr>
</tbody>
</table>

<sup>a</sup> expressed in minutes

Please note that sometimes, two children (for instance brothers and/or sisters) are recorded simultaneously while playing and that therefore video length that is used for transcription can be longer (e.g. S71 and S101).
2.2 Data collection and transcription

The corpora used in the present dissertation are existing, already transcribed corpora. In what follows, we explain the exact procedure of the transcription of the CI corpus and the NHlongitudinal corpus, applied by Schauwers (2006), Molemans (2011), Van den Berg (2012) and Van Severen (2012). However, for the NHcross-sectional corpus only the video recordings and an orthographic transcription of the selected speech samples are already available. In order to complete the transcriptions of this corpus, we followed the exact same procedures as described in the next paragraphs for the CI corpus and the NHlongitudinal corpus.

Approximately one-hour video recordings (60- to 90-minutes) are made at the children’s home with a JVC digital camera (type GZ-MG77E) with built-in microphone. Spontaneous speech samples of the interactions between the child and its (primary) caregiver(s) are gathered up to age five (e.g. reading a picture book, having a meal, playing, etc.). At ages six and seven, large busy pictures and the picture book “Frog, where are you?” (Mayer, 1969) are presented to the children, which formed a topic to initiate spontaneous interactions. After each recording, approximately 20 minutes in which the child was vocally the most active are selected in order to keep the transcription time within reasonable limits. On average, the transcription process lasted 14 hours, from the video recording at the child’s home up to the complete transcription of each 20 minute selection (Molemans, 2011; Molemans, Van den Berg, Van Severen & Gillis, 2012; Schauwers, 2006; Van den Berg, 2012; Van Severen, 2012). Only complete, finished interactions are selected and long pauses and noisy passages are excluded.

Speech samples in which no lexical but only prelexical child utterances appeared are excluded, as only lexical utterances were of interest in the present thesis. In other words, only data from word onset onwards are used in the present thesis and only lexical items are analysed. Next, transcriptions of the 20-minute selections are made in the program CLAN and according to the CHAT conventions (MacWhinney, 2000). For the identification of words, the procedure articulated by Vihman and McCune (1994) is used. More specifically, in order to be counted as a word, a child’s production has to
meet a number of criteria relative to its shape, its context of use, and its relation to other vocalisations. The criteria based on vocalisation shape comprise for instance an exact match of the vocalisation with the target production. Next, the criteria based on context involve for instance maternal identification, i.e. the identification of the vocalisation as a word by the mother. Finally, the criteria based on the relation to other vocalisations cover for instance the absence of inappropriate use, i.e. the vocalisation is only used in plausible contexts.

For the transcription of the children’s productions, both an orthographic and a phonemic transcription in DISC symbols and with stress marking are made. These transcriptions are based on the video recordings. A phonemic transcription with stress marking of the target word is retrieved from the lexical database Fonilex, which is ‘a pronunciation database containing the phonetic transcription of the most frequent word forms of Dutch as spoken in Flanders’ (Mertens, 2001). The phonetic transcription of normal pronunciation is selected and automatically added to the children’s productions. After target words are added, both the child’s production and the target word are syllabified with stress marking. Finally, the child’s utterances are aligned with the target forms at the phoneme level: the phonetic characters of the transcription pairs are automatically aligned with a computer program implementing a dynamic alignment algorithm (based on ADAPT (Elffers, Van Bael & Strik, 2005)). The alignments are verified manually and corrected if needed. For the transcription of the adult’s productions, an orthographic transcription is made based on the video recordings. The phonetic transcription in DISC symbols and with stress marking is retrieved from the Fonilex database (Mertens, 2001). The phonetic form of normal pronunciation is added automatically. This phonemic transcription is verified manually and specific attention is given to phenomena such as the deletion of final /n/ in Dutch. Target words are added and both adult’s productions and target words are syllabified with stress marking. Finally, an alignment between the target and the adult’s production is made. Exact same procedures as for the transcription of children’s productions are used.

Overall, 154,367 spoken lexical items, distributed in 80,191 utterances (prelexical utterances are excluded), are available, of which 71,280 words in 35,549 utterances for the CI corpus, 43,765 words in 35,504 utterances for the NHlongitudinal corpus and
Method

39322,651 words in 9,138 utterances for the NHcross-sectional corpus. With respect to child directed speech, a total of 686,296 spoken lexical items in 194,719 utterances appeared, of which 258,186 words in 72,689 utterances for the CI corpus, 370,026 words in 109,395 utterances for the NHlongitudinal corpus and 71,259 words in 15,316 utterances for the NHcross-sectional corpus. In Table 4, an overview of various metrics is shown for the speech samples in each corpus. Both the number of words and the number of utterances in children’s own speech and in child directed speech is given. In Appendix 1, the exact number of words per child per age is given for all corpora.

Table 4. Word and lexical utterance counts of the speech samples per corpus

<table>
<thead>
<tr>
<th></th>
<th>Children's utterances</th>
<th>Child directed speech</th>
<th>CI corpus</th>
<th>NH longitudinal corpus</th>
<th>NH cross-sectional corpus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># words</td>
<td># utterances</td>
<td># words</td>
<td># utterances</td>
<td># words</td>
</tr>
<tr>
<td>All speech samples</td>
<td>71 280</td>
<td>35 549</td>
<td>258 186</td>
<td>72 689</td>
<td>370 026</td>
</tr>
<tr>
<td>Per speech sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>320.50</td>
<td>190.5</td>
<td>1 196</td>
<td>349.5</td>
<td>1 217.86</td>
</tr>
<tr>
<td>Mean</td>
<td>336.22</td>
<td>167.68</td>
<td>396.80</td>
<td>84.26</td>
<td>308.16</td>
</tr>
<tr>
<td>SD</td>
<td>272.33</td>
<td>93.09</td>
<td>329.48</td>
<td>76.61</td>
<td>243.11</td>
</tr>
<tr>
<td>Min</td>
<td>1</td>
<td>1</td>
<td>1012</td>
<td>301</td>
<td>290</td>
</tr>
<tr>
<td>Max</td>
<td>1412</td>
<td>357</td>
<td>3 110</td>
<td>575</td>
<td>3 130</td>
</tr>
</tbody>
</table>

Reliability of phonemic transcriptions is computed on 10% of the CI corpus and the NHlongitudinal corpus. Two measures are used: percentages of agreement and kappa scores. In line with Cucchiarini (1996), Kappa scores are calculated in order to consider...
the possible influence of chance. Kappa scores are interpreted as follows: 0 – 0.20 *slight agreement*, .021 – 0.40 *fair agreement*, 0.41 – 0.60 *moderate agreement*, 0.61 – 0.80 *substantial agreement* and 0.81 – 0.99 *almost perfect agreement* (Landis & Koch, 1977). With respect to the CI corpus, only intrarater reliability is checked. The general percentage of agreement is 81.63% and the general Kappa score 0.87, which is interpreted as “almost perfect” (Landis & Koch, 1977). These numbers are the average reliability outcomes on the re-annotation of vowel place, vowel height, consonant place and consonant manner. Exact numbers equal 73.50% (Kappa 0.60) for vowel place, 84.40% (Kappa 0.69) for vowel height, 82.90% (Kappa 0.76) for consonant place and 85.7% (Kappa 0.79) for consonant manner. More detailed information about data collection and transcription of the CI corpus can be found in Schauwers (2006).

With respect to the NHlongitudinal speech samples, both intrarater reliability and interrater reliability are checked. Re-annotation of the full code (both consonants and vowels) resulted in a percentage of agreement of 63.69% for interrater reliability and 81.51% for intrarater reliability. Kappa scores are 0.60 for interrater reliability and 0.80 for intrarater reliability in the NHlongitudinal speech samples. These Kappa scores are on the edge of “moderate” to “substantial” and on the edge of “substantial” to “almost perfect” respectively (Landis & Koch, 1977). In Table 5, the percentages of agreement and Kappa scores of re-annotation for articulatory features can be found. This table is based on the results in Van den Berg (2012) and Van Severen (2012). More detailed information on data collection and transcription of the NHlongitudinal corpus can also be found in these theses. If reliability in the NHlongitudinal speech samples is determined on only the same articulatory features as in the CI speech samples (vowel place, vowel height, consonant place and consonant manner), intrarater reliability equals 90.25% (Kappa 0.85) and interrater reliability equals 79.02% (Kappa 0.68). The latter score is highly similar to that of the CI corpus.
Method

Table 5. Reliability outcomes for the NHlongitudinal corpus

<table>
<thead>
<tr>
<th>Consonants</th>
<th>Intrarater reliability</th>
<th>Interrater reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full code</td>
<td>% of agreement</td>
<td></td>
</tr>
<tr>
<td></td>
<td>84.17</td>
<td>70.43</td>
</tr>
<tr>
<td></td>
<td>Kappa score</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.68</td>
</tr>
<tr>
<td>Place of articulation</td>
<td>% of agreement</td>
<td></td>
</tr>
<tr>
<td></td>
<td>92.08</td>
<td>81.14</td>
</tr>
<tr>
<td></td>
<td>Kappa score</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.71</td>
</tr>
<tr>
<td>Voice</td>
<td>% of agreement</td>
<td></td>
</tr>
<tr>
<td></td>
<td>87.50</td>
<td>78.70</td>
</tr>
<tr>
<td></td>
<td>Kappa score</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.62</td>
</tr>
<tr>
<td>Manner of articulation</td>
<td>% of agreement</td>
<td></td>
</tr>
<tr>
<td></td>
<td>91.72</td>
<td>81.03</td>
</tr>
<tr>
<td></td>
<td>Kappa score</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.71</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Vowels</th>
<th>% of agreement</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Full code</td>
<td>78.84</td>
<td>56.95</td>
</tr>
<tr>
<td></td>
<td>0.76</td>
<td>0.52</td>
</tr>
<tr>
<td>Tense/lax</td>
<td>83.48</td>
<td>66.68</td>
</tr>
<tr>
<td></td>
<td>0.81</td>
<td>0.60</td>
</tr>
<tr>
<td>Place of articulation</td>
<td>88.25</td>
<td>75.69</td>
</tr>
<tr>
<td></td>
<td>0.82</td>
<td>0.61</td>
</tr>
<tr>
<td>Tongue height</td>
<td>88.93</td>
<td>77.23</td>
</tr>
<tr>
<td></td>
<td>0.82</td>
<td>0.66</td>
</tr>
<tr>
<td>Roundness</td>
<td>90.85</td>
<td>81.94</td>
</tr>
<tr>
<td></td>
<td>0.78</td>
<td>0.59</td>
</tr>
</tbody>
</table>

With respect to the NHcross-sectional corpus, intrarater reliability of orthographic transcriptions is computed on 65% of the data. A second investigator retranscribed the complete 20-minute selections orthographically. Thereof, two different percentages of agreement are calculated: firstly, the agreement of the number of words per utterance and secondly, the agreement on the words themselves. The median agreement of number of words per utterance equals 90.88% (range: 81.50% - 97.25%). The median agreement on the identified words themselves equals 81.38% (range: 69.50% - 92.63%). For this corpus, video recordings as well as the video selections and orthographic transcriptions are available. For the present thesis, phonemic transcriptions are added for both the children’s productions as the adult productions. Therefore as well as for the syllabification and alignment, the exact same procedure as used to transcribe the CI corpus and the NHlongitudinal corpus is followed.
2.3 Statistical analyses

In this dissertation, the data are analysed by means of multilevel modelling (MLM). Multilevel models, also called hierarchical linear models, are used for the longitudinal analyses and fixed occasion multilevel models for the cross-sectional comparisons (more information is provided in each chapter). Multilevel models have two parts: a fixed part and a random part. The fixed part comprises the independent variables, whereas the random part considers the variation in the dataset. The data of the present dissertation are structured hierarchically into three levels: individual words, various observations at consecutive ages and different children. These levels are nested. Words are nested within the different consecutive observations, which are nested within individual children. In other words, there is some variation in the present dataset resulting from the nesting of variables at different levels: At various ages, different words may be used and those may differ between children. In contrast to for instance ANOVA’s, MLM takes this variation into account (Baayen, 2008; Woltman, Feldstain, MacKay & Rocchi, 2012). Thus, MLM captures for instance the variation in the amount of word tokens and the variation in the sample sizes (i.e. different number of children in each group). In each chapter, the fixed and the random effects is specified for each analysis.
Method

Appendix 1. Overview of the number of words per corpus – ages are represented in months

<table>
<thead>
<tr>
<th>CI corpus</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
<th>26</th>
<th>27</th>
<th>28</th>
<th>29</th>
<th>30</th>
<th>31</th>
<th>32</th>
<th>33</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>62</td>
<td>83</td>
<td>30</td>
<td>139</td>
<td>114</td>
<td>123</td>
<td>141</td>
<td>239</td>
<td>185</td>
<td>205</td>
<td>276</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S2</td>
<td>5</td>
<td>3</td>
<td>41</td>
<td>26</td>
<td>48</td>
<td>21</td>
<td>46</td>
<td>98</td>
<td>172</td>
<td>204</td>
<td>176</td>
<td>238</td>
<td>212</td>
<td>240</td>
<td>257</td>
<td>159</td>
<td>240</td>
<td>323</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3</td>
<td>8</td>
<td>13</td>
<td>36</td>
<td>35</td>
<td>46</td>
<td>58</td>
<td>257</td>
<td>278</td>
<td>227</td>
<td>358</td>
<td>241</td>
<td>324</td>
<td>342</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S4</td>
<td>11</td>
<td>2</td>
<td>6</td>
<td>11</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>53</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S5</td>
<td>3</td>
<td>47</td>
<td>25</td>
<td>106</td>
<td>211</td>
<td>377</td>
<td>215</td>
<td>289</td>
<td>337</td>
<td>456</td>
<td>319</td>
<td>446</td>
<td>414</td>
<td>569</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S6</td>
<td>9</td>
<td>5</td>
<td>77</td>
<td>43</td>
<td>53</td>
<td>140</td>
<td>82</td>
<td>304</td>
<td>122</td>
<td>298</td>
<td>269</td>
<td>265</td>
<td>347</td>
<td>299</td>
<td>352</td>
<td>441</td>
<td>263</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S7</td>
<td>3</td>
<td>19</td>
<td>20</td>
<td>55</td>
<td>108</td>
<td>118</td>
<td>170</td>
<td>171</td>
<td>102</td>
<td>252</td>
<td>232</td>
<td>703</td>
<td>429</td>
<td>721</td>
<td>512</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S8</td>
<td>6</td>
<td>57</td>
<td>106</td>
<td>38</td>
<td>80</td>
<td>138</td>
<td>302</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S9</td>
<td>28</td>
<td>25</td>
<td>53</td>
<td>117</td>
<td>161</td>
<td>158</td>
<td>171</td>
<td>260</td>
<td>320</td>
<td>355</td>
<td>398</td>
<td>587</td>
<td>603</td>
<td>665</td>
<td>602</td>
<td>543</td>
<td>776</td>
<td>630</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S10</td>
<td>34</td>
<td>99</td>
<td>55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<p>| S1        | 278| 250| 353| 351| 368| 429| 479| 168| 427| 797|
| S2        | 329|
| S3        | 974| 516| 531| 541|
| S4        | 151| 321| 302| 234| 262| 175| 108| 279| 771|
| S5        | 469| 436| 873| 593| 764| 827| 844| 532|
| S6        | 264| 315|
| S7        | 437| 1177|
| S8        | 478| 394| 694| 310| 356| 493| 440| 774| 447| 599| 1099|
| S9        | 435| 884| 831| 1168|
| S10       | 42 |</p>
<table>
<thead>
<tr>
<th></th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>S11</td>
<td>43</td>
<td>70</td>
<td>32</td>
<td>111</td>
<td>90</td>
<td>128</td>
<td>190</td>
<td>239</td>
<td>270</td>
<td>293</td>
<td>385</td>
</tr>
<tr>
<td>S12</td>
<td>45</td>
<td>70</td>
<td>105</td>
<td>71</td>
<td>78</td>
<td>125</td>
<td>227</td>
<td>325</td>
<td>249</td>
<td>272</td>
<td>330</td>
</tr>
<tr>
<td>S13</td>
<td>5</td>
<td>102</td>
<td>110</td>
<td>177</td>
<td>108</td>
<td>111</td>
<td>213</td>
<td>229</td>
<td>257</td>
<td>322</td>
<td>452</td>
</tr>
<tr>
<td>S14</td>
<td>41</td>
<td>95</td>
<td>82</td>
<td>98</td>
<td>182</td>
<td>205</td>
<td>355</td>
<td>587</td>
<td>527</td>
<td>602</td>
<td>602</td>
</tr>
<tr>
<td>S15</td>
<td>8</td>
<td>10</td>
<td>12</td>
<td>11</td>
<td>98</td>
<td>135</td>
<td>155</td>
<td>192</td>
<td>287</td>
<td>328</td>
<td>328</td>
</tr>
<tr>
<td>S16</td>
<td>15</td>
<td>102</td>
<td>110</td>
<td>177</td>
<td>108</td>
<td>111</td>
<td>213</td>
<td>229</td>
<td>257</td>
<td>322</td>
<td>452</td>
</tr>
<tr>
<td>S17</td>
<td>1</td>
<td>4</td>
<td>14</td>
<td>14</td>
<td>59</td>
<td>40</td>
<td>40</td>
<td>72</td>
<td>72</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>S18</td>
<td>15</td>
<td>47</td>
<td>56</td>
<td>83</td>
<td>204</td>
<td>174</td>
<td>258</td>
<td>288</td>
<td>245</td>
<td>245</td>
<td>245</td>
</tr>
<tr>
<td>S19</td>
<td>11</td>
<td>33</td>
<td>45</td>
<td>50</td>
<td>116</td>
<td>204</td>
<td>204</td>
<td>204</td>
<td>204</td>
<td>204</td>
<td>204</td>
</tr>
<tr>
<td>S20</td>
<td>1</td>
<td>4</td>
<td>11</td>
<td>53</td>
<td>71</td>
<td>97</td>
<td>87</td>
<td>181</td>
<td>223</td>
<td>217</td>
<td>229</td>
</tr>
<tr>
<td>S21</td>
<td>1</td>
<td>8</td>
<td>43</td>
<td>85</td>
<td>153</td>
<td>182</td>
<td>250</td>
<td>250</td>
<td>250</td>
<td>250</td>
<td>250</td>
</tr>
<tr>
<td>S22</td>
<td>8</td>
<td>30</td>
<td>14</td>
<td>60</td>
<td>162</td>
<td>163</td>
<td>151</td>
<td>303</td>
<td>252</td>
<td>343</td>
<td>252</td>
</tr>
<tr>
<td>S23</td>
<td>1</td>
<td>14</td>
<td>31</td>
<td>101</td>
<td>83</td>
<td>101</td>
<td>107</td>
<td>204</td>
<td>128</td>
<td>243</td>
<td>321</td>
</tr>
<tr>
<td>S24</td>
<td>62</td>
<td>62</td>
<td>25</td>
<td>68</td>
<td>42</td>
<td>45</td>
<td>43</td>
<td>97</td>
<td>161</td>
<td>162</td>
<td>199</td>
</tr>
<tr>
<td>S25</td>
<td>69</td>
<td>81</td>
<td>107</td>
<td>50</td>
<td>131</td>
<td>78</td>
<td>150</td>
<td>182</td>
<td>226</td>
<td>205</td>
<td>323</td>
</tr>
<tr>
<td>S26</td>
<td>17</td>
<td>23</td>
<td>32</td>
<td>15</td>
<td>81</td>
<td>96</td>
<td>131</td>
<td>130</td>
<td>236</td>
<td>293</td>
<td>260</td>
</tr>
<tr>
<td>S27</td>
<td>50</td>
<td>92</td>
<td>115</td>
<td>146</td>
<td>116</td>
<td>135</td>
<td>198</td>
<td>307</td>
<td>302</td>
<td>441</td>
<td>476</td>
</tr>
<tr>
<td>S28</td>
<td>21</td>
<td>34</td>
<td>60</td>
<td>16</td>
<td>63</td>
<td>43</td>
<td>172</td>
<td>160</td>
<td>216</td>
<td>155</td>
<td>219</td>
</tr>
<tr>
<td>S29</td>
<td>45</td>
<td>33</td>
<td>44</td>
<td>64</td>
<td>138</td>
<td>153</td>
<td>118</td>
<td>179</td>
<td>244</td>
<td>321</td>
<td>385</td>
</tr>
<tr>
<td>S30</td>
<td>2</td>
<td>4</td>
<td>21</td>
<td>91</td>
<td>110</td>
<td>137</td>
<td>133</td>
<td>219</td>
<td>352</td>
<td>273</td>
<td>288</td>
</tr>
<tr>
<td>S31</td>
<td>54</td>
<td>56</td>
<td>133</td>
<td>115</td>
<td>208</td>
<td>83</td>
<td>213</td>
<td>162</td>
<td>156</td>
<td>284</td>
<td>369</td>
</tr>
<tr>
<td>S32</td>
<td>26</td>
<td>11</td>
<td>40</td>
<td>87</td>
<td>156</td>
<td>178</td>
<td>216</td>
<td>269</td>
<td>219</td>
<td>308</td>
<td>303</td>
</tr>
<tr>
<td>S33</td>
<td>80</td>
<td>77</td>
<td>180</td>
<td>113</td>
<td>203</td>
<td>150</td>
<td>283</td>
<td>276</td>
<td>323</td>
<td>417</td>
<td>295</td>
</tr>
<tr>
<td>S34</td>
<td>13</td>
<td>23</td>
<td>17</td>
<td>15</td>
<td>19</td>
<td>47</td>
<td>68</td>
<td>66</td>
<td>122</td>
<td>96</td>
<td>68</td>
</tr>
<tr>
<td>S35</td>
<td>2</td>
<td>47</td>
<td>22</td>
<td>34</td>
<td>42</td>
<td>68</td>
<td>145</td>
<td>159</td>
<td>159</td>
<td>159</td>
<td>159</td>
</tr>
<tr>
<td>S36</td>
<td>4</td>
<td>31</td>
<td>30</td>
<td>68</td>
<td>53</td>
<td>136</td>
<td>147</td>
<td>268</td>
<td>308</td>
<td>230</td>
<td>330</td>
</tr>
<tr>
<td>S37</td>
<td>23</td>
<td>46</td>
<td>142</td>
<td>144</td>
<td>143</td>
<td>162</td>
<td>109</td>
<td>122</td>
<td>195</td>
<td>277</td>
<td>329</td>
</tr>
<tr>
<td>S38</td>
<td>20</td>
<td>35</td>
<td>3</td>
<td>6</td>
<td>16</td>
<td>33</td>
<td>151</td>
<td>116</td>
<td>143</td>
<td>248</td>
<td>231</td>
</tr>
<tr>
<td>S39</td>
<td>19</td>
<td>26</td>
<td>26</td>
<td>50</td>
<td>100</td>
<td>138</td>
<td>94</td>
<td>143</td>
<td>266</td>
<td>155</td>
<td>156</td>
</tr>
</tbody>
</table>

NH longitudinal corpus
### Method

NHcross-sectional corpus

<table>
<thead>
<tr>
<th></th>
<th>24</th>
<th>36</th>
<th>48</th>
<th>60</th>
<th>72</th>
<th>84</th>
</tr>
</thead>
<tbody>
<tr>
<td>S41</td>
<td>676</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S42</td>
<td></td>
<td>877</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S43</td>
<td></td>
<td></td>
<td>611</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S44</td>
<td></td>
<td></td>
<td>798</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S45</td>
<td></td>
<td></td>
<td>691</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S46</td>
<td></td>
<td></td>
<td>819</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S47</td>
<td></td>
<td></td>
<td>642</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S48</td>
<td></td>
<td></td>
<td>782</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S49</td>
<td></td>
<td></td>
<td>788</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S50</td>
<td></td>
<td></td>
<td>725</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S51</td>
<td></td>
<td></td>
<td></td>
<td>310</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S52</td>
<td></td>
<td></td>
<td></td>
<td>179</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S53</td>
<td></td>
<td></td>
<td></td>
<td>274</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S54</td>
<td></td>
<td></td>
<td></td>
<td>380</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S55</td>
<td></td>
<td></td>
<td></td>
<td>338</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S56</td>
<td></td>
<td></td>
<td></td>
<td>229</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S57</td>
<td></td>
<td></td>
<td></td>
<td>201</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S58</td>
<td></td>
<td></td>
<td></td>
<td>501</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S59</td>
<td></td>
<td></td>
<td></td>
<td>356</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>714</td>
<td></td>
</tr>
<tr>
<td>S61</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>759</td>
<td></td>
</tr>
<tr>
<td>S62</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>615</td>
<td></td>
</tr>
<tr>
<td>S63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>761</td>
</tr>
<tr>
<td>Method</td>
<td>( S89 )</td>
<td>( S90 )</td>
<td>( S91 )</td>
<td>( S92 )</td>
<td>( S93 )</td>
<td>( S94 )</td>
</tr>
<tr>
<td>--------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>844</td>
<td>615</td>
<td>500</td>
<td>625</td>
<td>660</td>
<td>1130</td>
</tr>
</tbody>
</table>
Chapter 3
3 Phonemic accuracy development in children with cochlear implants up to five years of age by using Levenshtein Distance

A version of this chapter was published as a paper. Reference: Faes, J., Gillis, J., & Gillis, S. (2016). Phonemic accuracy development in children with cochlear implants up to five years of age by using Levenshtein Distance. Journal of Communication Disorders, 59, 40 – 58, doi: 10.1016/j.jcomdis.2015.09.004

ABSTRACT
Phonemic accuracy of children with cochlear implants (CI) is often reported to be lower in comparison age-matched children with normal hearing (NH). This study compares phonemic accuracy development in the spontaneous speech of Dutch-speaking children with CI and age-matched peers with NH. A dynamic cost model of Levenshtein distance is used to compute the accuracy of each word token. We set up a longitudinal design with monthly data for comparisons up to age two and a cross-sectional design with yearly data between three and five years of age. The main finding is that phonemic accuracy steadily increases throughout the period studied. Children with CI’s accuracy is lower than that of their age mates with NH, but this difference is not statistically significant in the earliest stages of lexical development. But accuracy of children with CI initially improves significantly less steeply than that of peers with NH. Furthermore, the number of syllables in the target word and target word’s complexity influence children’s accuracy, as longer and more complex target words are less accurately produced. Up to age four, children with CI are significantly less accurate than children with NH with increasing word length and word complexity. This difference has disappeared at age five. Finally, hearing age is shown to influence accuracy development of children with CI, while age of implant activation is not.

Keywords: Phonemic accuracy; children with CI; Levenshtein distance; target word complexity; target word syllable length
Levenshtein distance

3.1 Introduction
This study describes the development of phonemic accuracy of children with a cochlear implant (henceforth, CI) acquiring Dutch as their native language. As speech production depends on speech perception (Altvater-Mackensen & Fikkert, 2010; Jusczyk, 1992; Stoel-Gammon, 2011; Stoel-Gammon & Sosa, 2007), spoken language development in children with CI is a topic, which has received considerable attention in the literature. The ultimate question is whether the spoken first language proficiency of children with CI eventually reaches a level comparable to that of age mates with normal hearing (henceforth, NH). In this respect Nicholas and Geers (2007) published an aptly titled paper: “Will they catch up?” in which they investigated whether children with a CI, after an initial delay in comparison to children with NH, eventually close the gap and achieve age-appropriate spoken language levels in terms of vocabulary, sentence complexity and morphology. This study elaborates on this question by investigating speech production of children with CI and more specifically phonemic accuracy. Phonemic accuracy is based on phonemic transcriptions instead of narrow phonetic transcriptions. In other words, we focus on broad production accuracy of phonemes, regardless of phonetic variance. Speech production accuracy is important since it affects children’s speech intelligibility (Ingram, 2002) and, consequently, their communicative effectiveness. The development of phonemic accuracy is traced in a group of children with CI and a control group of children with NH, using Levenshtein distance, a measure that is relatively new in language acquisition research. We study three research questions: (1) How does phonemic accuracy develop immediately after cochlear implantation up to age 2;0? (2) How does it develop with prolonged linguistic experience, i.e., up to five years of age? And (3) Are there effects of length of implant use and age at implantation in accuracy development of children with CI. In what follows, we first discuss various measures of phonemic accuracy used in the literature and then, we elaborate on the main findings of phonemic accuracy development in the literature.
3.1.1 Measures of phonemic accuracy

In the literature two methods are frequently used to express phonemic accuracy: \textit{Percentage of Consonants Correct} (PCC, Shriberg, Austin, Lewis, McSweeny & Wilson, 1997; Shriberg & Kwiatkowski, 1982) and \textit{phonological mean length of utterance} (pMLU, Ingram, 2002). Both measures increase when production accuracy improves (Tyler & Lewis, 2005).

Shriberg and Kwiatkowski (1982) developed the first phonological yardstick, PCC. This metric indicates the percentage of correctly produced consonants. The procedure is fairly simple: A child’s rendition of a particular adult target word is lined up with a transcription of the actual target, and the number of overlapping consonants is counted. For instance, if a child says /bɔk/ for the adult target Dutch word /bɔks/ ‘fight’, two out of three consonants are correctly rendered. This procedure is repeated for all the words in a speech sample and the end result is the percentage of correctly produced consonants.

Although PCC is frequently used in child language studies, this metric faces three problems. First, only consonants are taken into consideration, while vowels are not considered, which means that the measure only partly reflects the child’s accuracy. This problem can easily be dealt with by the \textit{Percentage of Phonemes Correct} (PPC, Shriberg et al., 1997), which considers consonants as well as vowels. Secondly, in both PCC and PPC, a rendition of a particular target word may differ from the intended target in various ways. It may involve deletions, insertions and/or substitutions. In Table 1, the target word /bɔks/ ‘fight’ is compared with (example) renditions involving a deletion, a substitution and two examples of insertions. Deleting one consonant (as in /bɔks/) or substituting one (as in /pɔks/) leads to a PCC accuracy score of 66%. This result is expected in the sense that both renditions do not equal the adult target, and hence an accuracy score of 100% should not be expected. Inserting a consonant or a vowel (as in /bɒks/ and /bɾɔks/) leads to a higher PCC accuracy score. In fact, only in the case of insertions, the PCC is 100%, which implies that the child’s rendition equals the adult target, but it does not. So the second problem with PCC is that substitutions and deletions are “penalized” as errors or deviations from the adult target, while insertions are not. In fact, this means that there is an implicit weighting...
of the errors or deviations. Some errors have a more heavy weight than others. In the case of PCC insertions do not influence the value of PCC at all, while substitutions and deletions have an equal influence on the value of PCC. Nevertheless, such errors have varying effects on speech accuracy and speech intelligibility. Thirdly, the phonemic accuracy of less complex words, such as monosyllabic words or words with singleton consonants, is expected to be higher than the accuracy of more complex words, such as multisyllabic words or words with consonant clusters. For instance, Gathercole, Willis, Emslie and Baddeley (1991) pointed out that accuracy decreases with increasing target syllable length. Including word length may have an added value, but in computing PPC word length is not taken into account.

Table 1. Examples of PCC calculations

<table>
<thead>
<tr>
<th>Target</th>
<th>PCC child</th>
<th>PCC</th>
<th>Accuracy rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replica</td>
<td>/bɔks/</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>1) Deletion</td>
<td>/bɔk/</td>
<td>2</td>
<td>66%</td>
</tr>
<tr>
<td>2) Substitution</td>
<td>/pɔks/</td>
<td>2</td>
<td>66%</td>
</tr>
<tr>
<td>3) Insertion</td>
<td>/bɔkəs/</td>
<td>3</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>/bɔks/</td>
<td>3</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 2. Examples of pMLU calculations

<table>
<thead>
<tr>
<th>Target</th>
<th>pMLU child</th>
<th>pMLU</th>
<th>PWP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replica</td>
<td>/bɔks/</td>
<td>5</td>
<td>0.71 (71%)</td>
</tr>
<tr>
<td>1) Deletion</td>
<td>/bɔk/</td>
<td>6</td>
<td>0.86 (86%)</td>
</tr>
<tr>
<td>2) Substitution</td>
<td>/pɔks/</td>
<td>7</td>
<td>1.00 (100%)</td>
</tr>
</tbody>
</table>

In 2002, Ingram proposed phonological mean length of utterance (pMLU) as an alternative phonological measure. pMLU combines word length with PCC. pMLU is calculated by counting the number of phonemes of each word production, regardless of their accuracy. This means for instance that if the adult target contains three segments (/kæt/, ‘cat’) and the child produces three segments then the child’s rendition is credited with three pMLU points. An additional point is added to the pMLU score for each correctly produced consonant (Ingram, 2002). To measure the degree of
accuracy, Ingram (2002) developed the proportion of whole-word proximity (PWP). PWP is calculated by dividing the child’s pMLU by the target language’s pMLU. If the child’s rendition equals the adult target, PWP equals 1: What is targeted and what is actually produced are identical. Therefore, PWP is “an indirect measure of the child’s intelligibility” (Ingram, 2002). In Table 2, the pMLU and PWP scores of three child renditions of the target word /bɔks/ are given as an example. Both deleting and substituting a consonant lead to an imperfect PWP score (lower than 1.00). However, the insertion of /ə/ in /bɔkəs/, as shown in example (3), is not penalized, as the child’s pMLU score equals the target pMLU. Consequently, a maximum PWP score of 1.00 is reached.

pMLU faces three problems. First, pMLU reflects not only phonological but also morphological development (Taelman, Durieux & Gillis, 2005). At least for Dutch, pMLU is higher when inflected words, like plural nouns, finite verb forms etc., are included in the corpus (Taelman et al., 2005). Secondly, since pMLU is dependent on word length, languages with inherently longer words will have higher pMLU rates. This was illustrated by Saaristo-Helin, Savinainen-Makkonen and Kunnari (2006), who report higher pMLU rates for Finnish than English, with the former having longer words than the latter. This does not only make cross-linguistic comparison of children’s pMLU highly problematic, it also has the pernicious effect of the same error being weighted differently according to language. If a child deletes a phoneme in a language with longer words, the pMLU and associated PWP will be higher than if a deletion had occurred in a language with shorter words. Therefore, the same error incurs a higher penalty in the second language. Thirdly, as for PCC, an implicit weighting of speech errors is present, as Table 2 indicates. While deletions and substitutions are penalized as errors, insertions are not. The insertion of /ə/ in /bɔkəs/ is not penalized and has no influence on the accuracy rate (100%). Moreover, the pMLU metric penalizes deletions more heavily than substitutions. Nevertheless, a more explicit and desirable weighting in terms of frequency and complexity of errors is absent. The same is true for PCC/PPC. Speech errors differ from one another both in frequency and severity. Firstly, some errors are more common in children than other ones. Secondly, some speech errors are more severe as they influence
Levenshtein distance

intelligibility more than other speech errors. Therefore, it might be advisable to penalize frequent and less severe errors less than infrequent and severe errors. These are integrated into the Levenshtein distance measure.

In this study, Levenshtein distance (LD) is used as an alternative to track production accuracy and its development. LD is a commonly used technique to measure the difference between character strings in various scientific fields such as computer science, bio-informatics, and dialectology. LD between two character strings expresses their difference. A small distance between two strings indicates that they are very similar, and a large distance means that they are dissimilar. In the present context LD is used to compute the distance between the phonemic transcription of an adult target word and the phonemic transcription of a child’s rendition of that target word. It can be expected that the distance between the adult target and the child’s rendition is relatively large when a child starts using conventional words. And over time, the distance is expected to diminish, meaning that the child’s productions become more and more adult-like. Without going into the technical details of LD (see Method section), LD is determined by counting the number of edit operations needed to transform one string of characters into another one (Heeringa, 2004; Nerbonne & Heeringa, 2010; Wieling, Margaretha & Nerbonne, 2011). Three edit operations are assumed: Substitutions, deletions and insertions. The more edit operations required to transform one string (i.e. an adult target word) into another (i.e. the child’s rendition of that word), the more distant these two strings are. Moreover each edit operation has a cost associated with it and the number of edit operations needed determines the LD between the two character strings. For example, suppose the cost of each edit operation is 1, consider the difference between /bet/ and /be/ . The difference between these two character strings is the deletion of /t/, so LD equals 1, because the other characters are identical.

LD can not only be used to measure the distance between two character strings, DNA sequences or dialectal variants, but also to measure the distance between a word spoken by a child and the adult pronunciation of that word. Only a few studies have used LD to measure phonemic accuracy. Riches, Loucas, Baird, Charman and Simonoff (2011) have shown that LD error rates of adolescents with Specific
Language Impairment are higher than adolescent with Autism plus Language Impairments. Sanders and Chin (2009) have shown that the distance between word productions of children with CI and the adult targets measured by LD correlates with intelligibility judgments of naïve listeners. The larger the distance the less intelligible children’s speech was judged. In this study phonemic accuracy, as measured by LD, is studied in a group of children with CI and a group of children with NH. The use of this measure is relatively new in studies on language acquisition and has some important advantages. First, insertions are taken into account in the LD measure, in contrast to PCC and pMLU. In other words, all edit operations are considered in LD, while this is not the case in PCC and pMLU. Moreover, in the current study, a dynamic algorithm of LD is used. This means that the LD computations are based on a model of adult spoken language and not on a priori defined weights. In adult language, some phonemes typically appear more frequently than others, some phonemic variations are frequent in spontaneous speech (e.g., the deletion of /n/ after schwa at the end of words, as in /eta/ instead of /etan/ ‘eat’). These frequency differences are taken into account when computing the distance between the adult model and the child’s rendition. Infrequent speech errors as well as severe speech errors are penalized more heavily than frequent and less severe ones, due to the algorithm’s way of working (this is fully explained in the method section (2.3.).

3.1.2 Phonemic accuracy in children with NH and children with CI

This study investigates the phonemic accuracy of children with CI immediately after implantation up to the age of 2;0 (years;months) and with prolonged linguistic experience (i.e. up to age 5;0) and compare these to peers with NH. In children with NH, phonemic accuracy reaches ceiling percentages around the age of 3;0 (English: Dodd, Holm, Hua & Crosbie, 2003; Irwin & Wong, 1983; Finnish: Saaristo-Helin, 2009). For English, Warner-Czyz (2005) has shown that production accuracy in spontaneous speech is higher in children with NH as compared to children with CI up to six months after meaningful word onset. In spontaneous speech, the accuracy of Dutch speaking children with NH is found to be higher in comparison to children with CI. Between ages 2;1 and 2;6 pMLU and PWP reach median scores that are significantly higher in NH children (pMLU: Ca. 6.3, PWP: Ca. 0.8) than in children
Levenshtein distance

with CI (pMLU: Ca. 4.5, PWP: Ca. 0.7) (Schauwers, Taelman, Gillis & Govaerts, 2008b). At age 2;0, median proportions of phonemic accuracy at the word level are significantly higher for children with NH (0.66) than for age-matched children with CI (0.58) (van den Berg, 2012). Age at implantation of the participants with CI reported in Schauwers et al. (2008b) and in van den Berg (2012) was 1;0 (SD = 0;5). In the present article, the same children with CI as in Schauwers et al. (2008b) and van den Berg (2012) were studied. Their research is expanded in two ways. Firstly, the children with CI are studied up age 5;0, and secondly, a more fine-grained accuracy measure, viz. Levenshtein distance, is used. Therefore we expect the children with CI to be less accurate than peers with NH at age 2;0.

Nicholas and Geers (2007) report that English speaking children with CI implanted at around age 2;6 are more likely to catch up with age-matched peers with NH at age 4;6 than later implanted children with CI. Their conclusion holds for several linguistic domains, namely morphology, syntax and vocabulary. The present article investigates whether Nicholas & Geers’ finding can be replicated in another linguistic domain, viz. the phonemic accuracy development of early implanted children with CI. As children grow older their renditions of words can be expected to approximate the adult target forms. We investigate whether in the age span studied, children with CI reach a comparable level of accuracy as children with NH.

Thus, the speech production skills are delayed in children with CI at the early ages, but for instance Leigh, Detmman, Dowell and Briggs (2013) wonder about the long-term development of phonological skills of those children. This study will answer this question for Dutch-speaking children with CI. Studies of phonemic accuracy of children with CI acquiring Dutch as compared to their peers with NH are restricted to very young children (under age 2;0 in Schauwers et al. (2008b), van den Berg (2012)). In children acquiring English, production accuracy at older ages is shown to be significantly higher in children with NH than in children with CI. However, these results should be considered with care when generalizing them to Dutch and more long-term follow-ups are desirable (Leigh et al., 2013). Approximately at age 3;6, PPC is 83% for children with NH and 53% for children with CI (Ertmer, Kloiber, Jung, Kirleis & Bradford, 2012). At age 4;0, the same trend emerges in the PCC of
word initial consonants in a short sentence repetition task. For children with NH, all initial consonants except fricatives and affricates (86% accuracy) reach ceiling accuracy, while for children with CI initial consonant accuracy is only 62% (Ertmer & Goffman, 2011). At ages 3;0, 4;0 and 5;0, a Goldman-Fristoe Words and Sounds Test of Articulation (GFTA-2 Words and Sounds) score of 108.05 in children with NH and only 90 in hearing impaired children is reported by Eriks-Brophy, Gibson and Tucker (2013). This difference is statistically significant. However, 72% of the hearing impaired children reached average scores at age 5;0 (Eriks-Brophy et al., 2013) and thus caught up with their peers with NH. Note, however, that in Eriks-Brophy et al. (2013) and in Ertmer et al. (2012) mean age at implantation was 2;0 (SD = 1;1) and 1;6 (SD = 0;6) respectively, whereas in our study, children with CI were implanted much earlier, resulting in a mean age at implantation of 1;0 (SD = 0;5). Moreover, the hearing impaired participants reported on in Eriks-Brophy et al. (2013) were children with cochlear implants as well as children with hearing aids. All hearing impaired children were compared as a group to children with NH. In contrast, we compare children with NH with children with CI only. Moreover, we follow Dutch-speaking children with CI and children with NH up to the age of five. Unlike Ertmer and Goffman (2011), Ertmer et al. (2012) and Eriks-Brophy et al. (2013), who use standardized tests, we study spontaneous speech of children with CI and children with NH to compare accuracy development.

Numerous factors affect phonemic accuracy, including target word length and target word complexity. The number of syllables in the target word as well as the complexity of the target, determined by i.a. the presence of consonant clusters, influence children’s accuracy. This influence is attested in various studies reported in literature using nonword repetition tasks. For instance Gathercole et al. (1991) reported decreasing phonemic accuracy in typically developing children with increasing target nonword syllable length. In a repetition task, they were less accurate when target nonword syllable length increased from two to four syllables. Burkholder-Juhasz, Levi, Dillon and Pisoni (2007) and Von Mentzer, Lyxell, Sahlén, Dahlström, Lindgren, Ors, Kallioinen, Engström and Uhlén (2015) found similar results for children with CI and Nittouer, Caldwell-Tarr, Sansom, Twersky and Lowensthein
Levenshtein distance

(2014) obtained similar results for both children with NH and children with CI. Carter, Dillon and Pisoni (2002) indicated that syllable length in the target nonword influences accurate repetition of suprasegmental features in a nonword repetition task in children with CI. The reproduction of the correct number of syllables and the reproduction of the correct primary stress decreases when the target nonword has more syllables (Carter et al., 2002). Next, while Carter et al. (2002) did not find an effect of nonword complexity on suprasegmental accuracy in children with CI, nonword complexity influences phonemic, and thus segmental, accuracy in typically developing children (Macrae, 2013) and in children with CI (Von Mentzer et al., 2015). Macrae (2013) found lower consonant accuracy scores in nonwords with late developing sounds like fricatives, liquids and affricates and nonwords with consonant clusters; Von Mentzer et al. (2015) found more consonant omissions and substitutions when consonant clusters were present in the target nonword. Thus, in children with NH and children with CI, consonant accuracy decreases with increasing phonological complexity. Furthermore, frequent phonological processes in child language emphasize the influence of target word length and target word complexity on accuracy. For instance, reduplication, weak syllable deletion and cluster reduction occur frequently in child language (Johnson & Reimers, 2010): Monosyllabic, but mainly multisyllabic words are simplified by reduplicating the first syllable or by deleting weak unstressed syllables. In words with consonant clusters, clusters are simplified and only one consonant usually remains (Johnson & Reimers, 2010; Jongstra, 2003). Due to these phonological processes, children’s accuracy decreases considerably. The effects of target word complexity and target word syllable length were shown in nonword repetitions tasks for both children with NH and children with CI. In the present study, however, the influence of both factors on the phonemic accuracy of children with NH and children with CI’s spontaneous speech is examined.

For children with CI, two factors have been shown to affect language development, namely age of implantation and length of implant use, i.e. ‘hearing experience’. Numerous studies have pointed out the benefit of early implantation on language development (e.g. Nicholas & Geers, 2007; Schauwers, Gillis & Govaerts, 2008a), but only a few studies examined its effect on phonemic accuracy (Connor, Craig,
For Dutch, van den Berg (2012) found lower accuracy scores in later implanted children (range 5 – 20 months) and Schauwers et al. (2008b) found lower pMLU scores in children implanted in the second year of life as compared to children implanted in the first year of life. For English, Connor et al. (2006) concluded that children implanted before age 2;6 have better consonant production accuracy than later implanted age-mates. Though age of implantation is a major factor, length of implant use has also been indicated to influence language development (e.g. Szagun & Stumper, 2012). Szagun and Stumper (2012) even report that age of implantation does not affect language development of children with CI implanted within the sensitive period, i.e. before age 4;0, in contrast to length of implant use. For English-speaking children, accuracy is reported to increase with longer implant use (Blamey, Barry, Bow, Sarant, Paatsch & Wales, 2001; Eriks-Brophy et al., 2013; Gantz, Tyler, Woodworth, Tye-Murray & Fryauf-Bertschy, 1994; Tobey, Geers, Brenner, Altuna & Gabbert, 2003; Tomblin, Spencer & Lu, 2008; Tye-Murray, Spencer & Woodworth, 1995). After four years of implant use, overall phonemic accuracy is 62.9% (Tomblin et al., 2008). Accuracy increases after six years of implant use to 76.28% according to Tomblin et al. (2008) and approximately 86% according to Blamey et al. (2001). In Blamey et al. (2001), mean age at implantation was 3;9 (SD = 0;1) and in Tomblin et al. (2008), mean age at implantation was 4;6 (SD = 2;1). In the present study, we compare long-term accuracy development of children with NH and children with CI up to age 5;0, and at that age the children with CI have up to 4;6 years of device use. To the best of our knowledge, long-term accuracy development of Dutch children with CI has not been studied yet.

The present paper studies the impact of implantation in the first two years of life after early detection. As mean age at implantation in Blamey et al. (2001) and Tomblin et al. (2008) were above three years of age, it was impossible to investigate the influence of early implantation. In our study, phonemic accuracy of children with CI implanted at a much younger age (before age 1;8) is investigated, allowing us to study the impact of early implantation on accuracy. We expect the children with CI to benefit from their early implantation, eventually resulting in children with CI catching
Levenshtein distance

up with their peers with NH by age 5;0. Additionally, the impact of length of implant use on accuracy is examined.

3.2 Method

3.2.1 Participants
A longitudinal design was set up to assess language development of children with CI (S1 – S9): 9 children with CI were followed from word onset up to five years of age (Schauwers, 2006). The children were followed monthly from the moment of implant activation up to 2;6 years after activation. Additionally, yearly data were collected between ages 3;0 and 5;0. Here, the data will be analyzed from the appearance of first spoken words (median = 1;6, range 1;3 – 1;11). In chapter 2, more detailed information on the children with CI can be found.

The control group consisted of the complete NHlongitudinal corpus and the NHcross-sectional corpus between ages 3;0 and 5;0. With respect to the NHlongitudinal corpus, the data will be examined from the appearance of the first spoken words (median = 1;2, range 1;2 – 1;5). With respect to the NHcross-sectional corpus, a total of 32 children with NH participated: 9 three year olds (mean = 3;0, SD = 0;1), 12 four year olds (mean = 4;0 months, SD = 0;1), 11 five year olds (mean = 5;0, SD = 0;1). More information can be found in chapter 2.

3.2.2 Procedure
The data used in the current study are gathered as a part of a larger corpus, i.e. the CLiPS Child Language Corpus (CCLC). Only samples with lexical items were analyzed. Detailed descriptions of the data annotation procedures and the reliability of the transcriptions are reported in the method section (chapter 2).

For the children with CI, a total of 58,686 word tokens, with a median of 5,606 per child (range: 4,079 – 10,520) were available. Up to age 2;0, 3,406 word tokens were available, with a median of 320 word tokens per child (range 16 – 1140). Between ages 3;0 and 5;0, a total of 29,326 word tokens were available, with a median of 2,944 per child (range 1,927 – 5,002). For the longitudinal analysis of only children with CI, all 58,686 word tokens were used. For the children with NH, a total of 59,019 word
tokens were available. For the children with NH up to age 2;0, data consisted of a total of 42,535 word tokens, with a median of 1438.5 word tokens per child (range 455 – 2,889). For the NH children between ages 3;0 and 5;0 a total of 16,484 word tokens were available, with a median of 525.5 word tokens per child (range 178 – 965).

3.2.3 Levenshtein distance (LD)

Conceptually, computing Levenshtein distance amounts to comparing word after word a phonemic transcription of a child’s own word productions with the adult equivalent of those words. If both transcriptions are identical, then their distance is in principle zero and the child’s production can be said to be adult-like. If that is not the case, LD measures the distance between the adult target and the child’s rendition of that target. LD is a way to measure that distance by computing the minimal edit distance between the two transcriptions: How can one transcription be transformed into the other by a (minimal) set of edit operations (deletions, insertions, substitutions)? Given that set of operations, each operation is given a weight or “cost”. Wieling, Prokic and Nerbonne (2009) proposed a procedure for inducing those weights automatically from a corpus of transcribed speech, instead of (arbitrarily) assigning a weight to each edit operation a priori. Their proposal is essentially the following: Suppose a target word contains /e/ and that segment is rendered as /ə/, then that rendition is closer to the target than when /e/ would have been substituted by /u/. For instance, in Dutch vowel reductions – such as rendering /e/ as /ə/ – are frequent in spontaneous adult speech (Swerts, Kloots, Gillis & De Schutter, 2003). Hence such substitutions are expected to occur frequently when comparing a standard transcription deriving from a phonemic lexicon with a (broad) transcription of spontaneous speech, while the substitution of /e/ by /u/ is very infrequent. Hence the distance between /e/ and /ə/ is smaller than that between /e/ and /u/. In other words, the “cost” for transforming /e/ into /ə/ should be smaller than the “cost” for transforming /e/ into /u/. This cost model is dynamically derived from the corpus, and frequency information is crucial in computing the model: The “cost” for transforming a segment into another one is smaller if this pairing occurs relatively frequently. Note in bypassing that Wieling, Margaretha and Nerbonne (2012) found strong significant correlations between the induced distances (or costs) and the acoustic distances they measured.
Levenshtein distance

The basic algorithm for computing LD in this study was taken from Wieling et al. (2012); Wieling et al. (2009), who computed LD using a dynamic cost model. In what follows the algorithm will be described and exemplified first. After the description of the basic algorithm, the adjustments to fit the purposes of the current study will be discussed. Computing the distance between an adult target and a child’s rendition of that target, is not exactly the same as computing the distance between two adult dialects in Wieling et al. (2009). Moreover, the corpus analyzed in the present study was a longitudinal one, which requires some additional adjustments to be made.

The first phase in computing LD consists of aligning two strings of segments, such as an adult target word and a child’s rendition of that target. For instance, the pair of transcriptions in (1) represent the target word /spelə/ (the Dutch word for to play) and a child’s rendition of that word as /pe/:

(1) Adult target  s p e l ə
      Child’s rendition  p e

In principle there are many different possible alignments, but the algorithm incorporates a binary same-different strategy, trying to line up matches between segments and avoiding mismatches, and trying to maximize the number of matches. Moreover, the algorithm uses a VC-sensitive strategy allowing only vowels to be lined up with vowels and consonants with consonants (Wieling et al., 2009). Practically speaking the algorithm is implemented using a dynamic programming algorithm (Wagner & Fisher, 1974) that seeks the minimal edit distance. Applying these principles to example (1), the net result will be (2).

(2) Adult target  s p e l ə
      Child’s rendition  . p e .

Note that both the target and the rendition in (2) are of equal length since “filler characters” (represented by a dot) have been inserted in the shorter character string, which stand for an empty position or “zero segment”. The /p/ and /e/ are aligned with
the matching segments in the target and the rendition, /s/ in the target has been deleted in the child’s rendition (introducing a dot), /l/ and /a/ are deleted as well as.

The second step consists of determining the cost of an insertion, substitution or deletion. This is accomplished by implementing a dynamic cost model (Wieling et al., 2012; Wieling et al., 2009). For this purpose the program runs through a transcription and considers each target – rendition pair and computes the Pointwise Mutual Information (PMI) of each segment. For instance, given the initial alignment represented in (2), the probability of the pairings (/s/ - /./), (/p/ - /p/), (/e/ - /e/), etc. are computed. The PMI of all these pairs is calculated using equation (3):

\[
PMI(x, y) = \log_2 \frac{p(x,y)}{p(x)p(y)}
\]

where \(p(x,y)\) denotes the probability of encountering the pair \((x,y)\) in the alignment, \(p(x)\) the probability of encountering segment \(x/\), and \(p(y)\) the probability of segment \(y/\). Take the pair \((/s/ - /./)\) in example (2), the PMI of that pair is the probability of \(s/\) in the target being lined up with the empty character \(./\) in the child’s rendition, divided by the product of the probability of \(s/\) in the adult corpus and the probability of \(./\) in the child corpus. Mutatis mutandis, the PMI of the pair \((/p/ - /p/)\) is the probability of \(p/\) in the target paired with \(p/\) in the rendition, divided by the probability of \(p/\) in the adult corpus and its probability in the child corpus. The division is actually meant to normalize the probability of the pairing \(p(x,y)\) in (3), or the statistical dependence of segments \(x/\) and \(y/\), with respect to the probability of \(x/\) and \(y/\) being statistically independent (Wieling et al., 2012).

This basic procedure is repeated for the consecutive child utterances and their targets in a transcription. Evidently, the PMI value for each pair of segments will change as more words are processed. For instance, in addition to the pairing \((/p/ - /p/)\) in (2), alternative pairings may be encountered such as \((/p/ - /b/)\) or \((/p/ - /./)\), i.e., we may encounter examples of \(p/\) being substituted by \(b/\), or \(p/\) being deleted. After processing an entire transcription, the net result is a first alignment of both transcriptions, and a list of segment pairs, each with a PMI value. The PMI value of a pair is converted into a cost (or a distance, for that matter) by subtracting it from the
Levenshtein distance

maximum PMI value. The rationale is that if the PMI value is very high, meaning that a particular pairing of segments occurs very frequently in the data, the resulting cost (distance) should be small. In this way, frequently co-occurring segments will have a much smaller cost (distance) than segments that do not co-occur very often.

As an example of how LD is computed, consider the two child renditions of the adult word /spelə/ (Eng.: to play) in (4):

(4) Adult target s p e lə
   Child’s rendition 1 . p e .
   Child’s rendition 2 . p e lə

In order to calculate the LD of the 2 child renditions (/pe/ and /pelə/), the proportions of each segment and each pair of segments in the alignment have to be computed first. Take the pair /s/-./: suppose the proportion of /s/ in the transcription of the adult targets equals 0.075 and the proportion of ./ in the child’s renditions equals 0.05, and in the aligned transcriptions the proportion of the pair /s/-./ equals 0.004, meaning that in 0.4% of the cases the child deletes /s/. Inserting these figures in equation (3) results in the PMI: \( \log_2 (0.004 / 0.075 * 0.05) = 0.09 \). By way of example, proportions are given for the relevant segments and pairs in (4), and the PMI of each pair is calculated using the formula in (3). The result is shown in (5):

<table>
<thead>
<tr>
<th>(5)</th>
<th>Pair</th>
<th>P(x,y)</th>
<th>P(x)</th>
<th>P(y)</th>
<th>PMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>/s/-./</td>
<td>0.0040</td>
<td>0.0750</td>
<td>0.0500</td>
<td>0.09</td>
</tr>
<tr>
<td>b.</td>
<td>/p/-/p/</td>
<td>0.9900</td>
<td>0.0005</td>
<td>0.0005</td>
<td>21.92</td>
</tr>
<tr>
<td>c.</td>
<td>/e/-/e/</td>
<td>0.3500</td>
<td>0.2500</td>
<td>0.2500</td>
<td>2.49</td>
</tr>
<tr>
<td>d.</td>
<td>/l/-/l/</td>
<td>0.3000</td>
<td>0.0900</td>
<td>0.0900</td>
<td>5.21</td>
</tr>
<tr>
<td>e.</td>
<td>/l/-./</td>
<td>0.0100</td>
<td>0.0900</td>
<td>0.0500</td>
<td>1.15</td>
</tr>
<tr>
<td>f.</td>
<td>/ə/-/ə/</td>
<td>0.6000</td>
<td>0.1500</td>
<td>0.1500</td>
<td>4.74</td>
</tr>
<tr>
<td>g.</td>
<td>/ə/-./</td>
<td>0.0200</td>
<td>0.1500</td>
<td>0.0500</td>
<td>1.42</td>
</tr>
</tbody>
</table>
In order to calculate the LD for /pe/ and /pelə/, the maximum PMI in the alignments is determined. In example (5) the maximum PMI equals 21.92. Subsequently the relevant pairings are subtracted from the maximal PMI and the resulting values are summed. For /pe/, this results in subtracting the values in rows a, b, c, e and g in example (5) from the maximum PMI and adding up the resulting values. For /pelə/, this results in adding lines a, b, c, d and f in example (5) after subtracting them from the maximal PMI. This results in a LD of 82.53 for /pe/ and 75.15 for /pelə/. This example shows that a word that is has a more identical alignment (/pelə/) has a smaller LD. Moreover, intuitively, /pelə/ is closer to the target than /pe/.

After the first pass through the transcript, the procedure is repeated. The Levenshtein algorithm is used to generate a new alignment, but this time with an alternative weighting scheme, viz. the dynamically computed cost model. In the first pass, matching segments received a “cost” of 0 and non-matching segments (insertions, substitutions or deletions) a “cost” of 1. From the second pass onwards, the segment distances computed during the previous pass are used in constructing a new alignment. This iterative process of aligning targets and renditions and computing segment distances, is stopped once two consecutive alignments are identical, and, hence, convergence is reached.

When the final cost model is computed, the LD between an adult target and a child’s rendition is the sum of distances between the individual segments. But since LD is relative to a word’s length, the LD score was normalized for word length by computing the average cost score per word (Heeringa, 2004). Applying this to the examples in (4) and (5), this means dividing the resulting LD of both child renditions by 5, i.e. the word length in phonemes. This eventually results in 16.51 for /pe/ and 15.03 for /pelə/. The rationale for this normalization is a follows: The score of a perfect match is most likely a non-zero distance. Consequently a short incorrectly produced word possibly receives a smaller LD value than a long perfectly produced word solely because the latter is longer. This undesirable effect of word length is prevented as far as possible by calculating the average LD.

Applying LD to transcriptions of children’s speech requires two more adaptations of the procedure for computing LD with a dynamic cost model. The first adaptation
Levenshtein distance concerns the very first step in constructing the cost model dynamically. In constructing the cost model in the first pass, the cost model takes 0 for a segmental match and 1 for a mismatch, and subsequently PMI is computed. But suppose a child systematically deletes a particular segment in the renditions of adult targets. In that case the pairing of that segment with an empty segment will be highly frequent, and hence will be preferred to even the correct pairing. In order to circumvent this undesirable situation, the procedure was extended to a two-phase procedure. In the initial phase, a cost model of the adult language was constructed. For this purpose the transcription of the actual adult speech was aligned with a target transcription derived from the Fonilex database. The resulting cost model was computed for the entire corpus of adult speech and was considered to be a good estimate of adult speech patterns. Subsequently, in the second phase, children’s speech was aligned with the adult targets. In the first pass the cost model derived from the adult transcriptions was used, so that situations such as a deletion being less costly than the adult model were avoided. But an implication of this restructuring of the procedure was that particular pairs of segments encountered in lining up child and adult transcriptions were not represented in the initial (adult) cost model. For instance, substitution of a closed front vowel by an open back vowel is hardly expected in adult language. Hence the pairing of /i/ with /a/ is not expected in the adult cost model, so its probability has to be estimated in some way. For this purpose the second adaptation of the procedure consisted of introducing Katz Smoothing (Chen & Goodman, 1998; Katz, 1987) before calculating the PMI. Katz Smoothing assigns unobserved pairs of segments an estimated probability. They receive a small probability and thus a large cost (or distance) in the dynamic cost model.

Summing up, given the corpus of dyadic interactions, the speech of adults in the entire corpus was first used in order to construct a cost model, which reflects the characteristics of adult spoken Dutch. Next, the Levenshtein distance was computed for each transcription of each individual child at the various ages under consideration.
3.2.4 Data analyses

Dynamic Levenshtein distance (LD) of each phonemic transcription of a word token was computed automatically after excluding substandard words for which there was no standard pronunciation in the Fonilex database. Phonemic accuracy development in terms of LD was investigated longitudinally between ages 1;2 and 2;0 for the children with NH and between word onset after implantation up to age 2;0 for the children with CI. Phonemic accuracy was investigated cross-sectionally at ages 3;0, 4;0 and 5;0. Age ranges were 2;10 – 3;4 for analyses at age 3;0, 3;9 – 4;3 for analyses at age 4;0 and 4;11 – 5;3 for analyses at age 5;0. Outliers were determined by the interquartile rule and omitted from further analyses. All statistical analyses were carried out in JMP® Pro 11 by means of multilevel modeling (MLM). Multilevel models were used for the longitudinal analyses and fixed occasion multilevel models for the cross-sectional comparisons.

For the first research goal the phonemic accuracy of children with NH and children with CI was examined longitudinally up age 2;0. The fixed effects were the children’s hearing status (henceforth, HearingStatus), their ages in months (henceforth, Age), the length of the adult target words in syllables (henceforth, UtteranceSyllableLength), and the ratio of utterance length in phonemes of the target words over the utterance length in syllables of the target words (henceforth, WordComplexity). This last variable captures the complexity of the target words: The ratio is higher in more complex words. Random intercepts and slopes were introduced to model the variation between children.

For the second research goal, cross-sectional comparisons between children with NH and children with CI at ages 3;0, 4;0 and 5;0 were made. Fixed effects were HearingStatus, UtteranceSyllableLength and WordComplexity. At each age a random effect of child was included.

For the third research goal, the impact of age of implant activation and length of implant use was examined. For this purpose, all available data of the children with CI were used, i.e. also the monthly speech samples up to 2;6 years after implantation. Fixed effects were UtteranceSyllableLength, WordComplexity, length of implant use in months (henceforth, HearingAge) and age of CI activation in months (henceforth,
Levenshtein distance (Clactivation) as well as their interaction. Random intercepts and slopes were introduced to model inter-subject variation. In all analyses a significance level of p<0.05 was set.

3.3 Results

3.3.1 Phonemic accuracy development up to age 2;0

Analyses in this section discuss the longitudinal development of LD from word onset up to age 2;0 for children with NH and children with CI. Table 3 displays the results of fitting the mixed effect model. The predicted values of the model are plotted in Figure 1. As Figure 1 demonstrates, LD of children with CI is higher than LD of children with NH. This means that the accuracy of children with CI is lower than the accuracy of age-mates with NH. However, no significant main effect of HearingStatus is found, meaning that although children with CI’s accuracy is lower than that of children with NH, and hence their LD is higher, this difference does not reach statistical significance. In addition, Figure 1 shows a linear decrease of LD with age, and as can be inferred from Table 3 the effect of Age is significant (p = 0.0007). No interaction effect between HearingStatus and Age is found, indicating that the development depicted in Figure 1 is highly similar for both groups of children.
Table 3 shows main effect results and the parameter estimates of UtteranceSyllableLength and WordComplexity. First of all, UtteranceSyllableLength is found to influence LD significantly. LD is higher with increasing UtteranceSyllableLength, thus accuracy is lower in longer words. Secondly, WordComplexity is found to influence LD as well: LD increases with increasing WordComplexity, which means that accuracy is lower in more complex words.

Importantly, some interaction effects of UtteranceSyllableLength and WordComplexity with Age and HearingStatus are found. First of all, Table 3 shows an interaction between both Age and UtteranceSyllableLength and Age and WordComplexity. This means that the influence of UtteranceSyllableLength and WordComplexity on LD decreases with time, or, in other words, that the increase of LD with increasing syllable length and word complexity becomes less steep as children grow older. Secondly, Table 3 shows that whereas there is no interaction between UtteranceSyllableLength and HearingStatus, there is an interaction between HearingStatus and WordComplexity. The predicted values of the model are plotted in
Levenshtein distance

Figure 2. It appears that the influence of UtteranceSyllableLength is the same for both groups of children, but WordComplexity influences LD differently in both groups of children. WordComplexity is found to influence LD more severely in children with CI than in children with NH: The increase of LD with increasing WordComplexity is steeper in children with CI than in children with NH.

Table 3. Fixed effect results and parameter estimates of the fixed effects up to age 2;0

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>Estimate</th>
<th>SE</th>
<th>F ratio</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1</td>
<td>18.0992</td>
<td>0.6072</td>
<td>/</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Age</td>
<td>1</td>
<td>-0.0992</td>
<td>0.0283</td>
<td>12.27</td>
<td>0.0007</td>
</tr>
<tr>
<td>HearingStatus</td>
<td>1</td>
<td>0.9297</td>
<td>0.4860</td>
<td>3.66</td>
<td>0.0636</td>
</tr>
<tr>
<td>HearingStatus x Age</td>
<td>1</td>
<td>-0.0429</td>
<td>0.0226</td>
<td>3.59</td>
<td>0.0653</td>
</tr>
<tr>
<td>UtteranceSyllableLength</td>
<td>1</td>
<td>2.1255</td>
<td>0.1764</td>
<td>145.22</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>UtteranceSyllableLength x HearingStatus</td>
<td>1</td>
<td>-0.0462</td>
<td>0.0345</td>
<td>1.79</td>
<td>0.1806</td>
</tr>
<tr>
<td>UtteranceSyllableLength x Age</td>
<td>1</td>
<td>-0.0524</td>
<td>0.0080</td>
<td>42.94</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>WordComplexity</td>
<td>1</td>
<td>1.2746</td>
<td>0.1447</td>
<td>77.57</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>WordComplexity x HearingStatus</td>
<td>1</td>
<td>0.0912</td>
<td>0.0269</td>
<td>11.52</td>
<td>0.0007</td>
</tr>
<tr>
<td>WordComplexity x Age</td>
<td>1</td>
<td>-0.0232</td>
<td>0.0066</td>
<td>12.24</td>
<td>0.0005</td>
</tr>
</tbody>
</table>

df = degrees of freedom
SE = Standard error
NH is the reference category

Figure 2. Interaction effect of HearingStatus with UtteranceSyllableLength (left pane) and WordComplexity (right pane) between ages 1;2 and 2;0 (Predicted values)
3.3.2 Phonemic accuracy development from ages 3;0 to 5;0

The analyses in this section pertain to the longitudinal samples of the children with CI and to the cross-sectional samples of the children with NH. Therefore, the results will be discussed for each age separately.

At age 3;0, all main effects except HearingStatus were significant, as indicated in Table 4. UtteranceSyllableLength as well as WordComplexity influence LD significantly: LD increases with increasing UtteranceSyllableLength and with increasing WordComplexity. Furthermore, the estimated LD is higher in children with CI as compared to children with NH, but this effect is not statistically significant. Note however that significant interactions of HearingStatus with UtteranceSyllableLength and WordComplexity are found. Consequently, the increase of LD with increasing UtteranceSyllableLength and increasing WordComplexity is steeper in children with CI as compared to children with NH. These effects are plotted in Figure 3.

Table 4. Fixed effect results and parameter estimates of the fixed effects at age 3;0

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>Estimate</th>
<th>SE</th>
<th>F ratio</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1</td>
<td>15.7047</td>
<td>0.1535</td>
<td>/</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>HearingStatus</td>
<td>1</td>
<td>0.0898</td>
<td>0.1535</td>
<td>0.34</td>
<td>0.5620</td>
</tr>
<tr>
<td>UtteranceSyllableLength</td>
<td>1</td>
<td>0.4069</td>
<td>0.0358</td>
<td>129.27</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>UtteranceSyllableLength x HearingStatus</td>
<td>1</td>
<td>0.2058</td>
<td>0.0358</td>
<td>33.06</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>WordComplexity</td>
<td>1</td>
<td>0.3157</td>
<td>0.0330</td>
<td>91.76</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>WordComplexity x HearingStatus</td>
<td>1</td>
<td>0.1043</td>
<td>0.0330</td>
<td>10.01</td>
<td>0.0016</td>
</tr>
</tbody>
</table>

df = degrees of freedom  
SE = Standard error  
NH is the reference category
Figure 3. Interaction effects of HearingStatus with UtteranceSyllableLength and WordComplexity on LD (age 3;0, predicted values)

At age 4;0, the same trends emerge as at age 3;0. Table 5 presents the fixed effect results and the estimates of the fixed effect results. All effects are similar. No significant effect of HearingStatus is found. Furthermore, LD increases with increasing UtteranceSyllableLength and increasing WordComplexity. as was observed at age 3;0, the increase of LD with increasing UtteranceSyllableLength and increasing WordComplexity is higher in children with CI than in children with NH. In Figure 4 the predicted LD values of the model are plotted.
Table 5. Fixed effect results and parameter estimates of the fixed effects at age 4;0

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>Estimate</th>
<th>SE</th>
<th>F ratio</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1</td>
<td>15.4432</td>
<td>0.0537</td>
<td>/</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>HearingStatus</td>
<td>1</td>
<td>0.0427</td>
<td>0.0537</td>
<td>0.63</td>
<td>0.4281</td>
</tr>
<tr>
<td>UtteranceSyllableLength</td>
<td>1</td>
<td>0.1798</td>
<td>0.0160</td>
<td>125.68</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>UtteranceSyllableLength x</td>
<td>1</td>
<td>0.0754</td>
<td>0.0160</td>
<td>22.11</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>HearingStatus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WordComplexity</td>
<td>1</td>
<td>0.1643</td>
<td>0.0143</td>
<td>131.44</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>WordComplexity x HearingStatus</td>
<td>1</td>
<td>0.0502</td>
<td>0.0143</td>
<td>12.29</td>
<td>0.0005</td>
</tr>
</tbody>
</table>

df = degrees of freedom  
SE = Standard error  
NH is the reference category

Figure 4. Interaction effects of HearingStatus with UtteranceSyllableLength and WordComplexity on LD (age 4;0, predicted values)

Fixed effect results and the estimates of the fixed effect results at age 5;0 are presented in Table 6. No effect of HearingStatus is found, but UtteranceSyllableLength as well as WordComplexity influence LD significantly. These effects are shown in Figure 5. In addition, no significant interaction between HearingStatus and UtteranceSyllableLength and between HearingStatus and
Levenshtein distance

WordComplexity are found. This means that WordComplexity and UtteranceSyllableLength influence LD in a comparative way in both groups of children.

Table 6. Fixed effect results and parameter estimates of the fixed effects at age 5;0

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>Estimate</th>
<th>SE</th>
<th>F ratio</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1</td>
<td>15.3585</td>
<td>0.0337</td>
<td>/</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>HearingStatus</td>
<td>1</td>
<td>-0.0381</td>
<td>0.0337</td>
<td>1.27</td>
<td>0.2612</td>
</tr>
<tr>
<td>UtteranceSyllableLength</td>
<td>1</td>
<td>0.0682</td>
<td>0.0110</td>
<td>38.70</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>UtteranceSyllableLength x HearingStatus</td>
<td>1</td>
<td>0.0122</td>
<td>0.0110</td>
<td>1.23</td>
<td>0.2675</td>
</tr>
<tr>
<td>WordComplexity</td>
<td>1</td>
<td>0.0381</td>
<td>0.0097</td>
<td>15.39</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>WordComplexity x HearingStatus</td>
<td>1</td>
<td>-0.0053</td>
<td>0.0097</td>
<td>0.30</td>
<td>0.5864</td>
</tr>
</tbody>
</table>

df = degrees of freedom
SE = Standard error
NH is the reference category

Figure 5. Interaction effects of HearingStatus with UtteranceSyllableLength and WordComplexity on LD (age 5;0, predicted values)
3.3.3 Influence of length of implant use and age at CI activation

In order to study the influence of age at implant activation and length of implant use the (longitudinal) data of the children with CI were analyzed separately, with HearingAge and age at CI activation as predictors in the model (Table 7). LD is plotted relative to HearingAge in Figure 6. There is a quadratic, but overall decreasing effect of HearingAge on LD. Next, LD is higher in later implanted children, but this difference is not statistically significant, indicating that the effect of age at CI activation on LD can safely be attributed to sampling error. No interaction effect between HearingAge and CI activation is found, indicating that the decrease of LD in children with CI implanted at a later age is not significantly different from the decrease of LD in earlier implanted children.

UtteranceSyllableLength and WordComplexity significantly affect LD: LD increases with increasing UtteranceSyllableLength and increasing WordComplexity, as presented in Figure 7. For both variables, interactions with CI activation and HearingAge are found. The increase of LD with increasing WordComplexity is less steep in children with later CI activation. The interaction of WordComplexity and CI activation is significant, while the interaction of UtteranceSyllableLength and CI activation is not. Next, the impact of both UtteranceSyllableLength and WordComplexity on LD decreases significantly with HearingAge.
Levenshtein distance

**Table 7.** Fixed effect results and parameter estimates of the fixed effects for children with CI

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>Estimate</th>
<th>SE</th>
<th>F ratio</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1</td>
<td>16.4720</td>
<td>0.7527</td>
<td>/</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>HearingAge</td>
<td>1</td>
<td>-0.0535</td>
<td>0.0131</td>
<td>16.69</td>
<td>0.0039</td>
</tr>
<tr>
<td>HearingAge x HearingAge</td>
<td>1</td>
<td>0.0010</td>
<td>0.0001</td>
<td>1267.55</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>CIactivation</td>
<td>1</td>
<td>0.0154</td>
<td>0.0531</td>
<td>0.08</td>
<td>0.7797</td>
</tr>
<tr>
<td>CIactivation x HearingAge</td>
<td>1</td>
<td>-0.0007</td>
<td>0.0009</td>
<td>0.68</td>
<td>0.4363</td>
</tr>
<tr>
<td>UtteranceSyllableLength</td>
<td>1</td>
<td>0.9946</td>
<td>0.0402</td>
<td>611.17</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>UtteranceSyllableLength x HearingAge</td>
<td>1</td>
<td>-0.0171</td>
<td>0.0007</td>
<td>640.03</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>UtteranceSyllableLength x CIactivation</td>
<td>1</td>
<td>-0.0031</td>
<td>0.0024</td>
<td>1.75</td>
<td>0.1860</td>
</tr>
<tr>
<td>WordComplexity</td>
<td>1</td>
<td>0.9502</td>
<td>0.0359</td>
<td>699.47</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>WordComplexity x HearingAge</td>
<td>1</td>
<td>-0.0203</td>
<td>0.0006</td>
<td>1339.48</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>WordComplexity x CIactivation</td>
<td>1</td>
<td>-0.0043</td>
<td>0.0020</td>
<td>4.42</td>
<td>0.0354</td>
</tr>
</tbody>
</table>

df = degrees of freedom
SE = Standard error

---

**Figure 6.** LD development of children with CI with HearingAge up to age 5;0

(Predicted values)
3.4 Discussion

The aim of this study was to examine phonemic accuracy of Dutch-speaking children with CI in comparison to age-mates with NH in order to determine if phonemic accuracy of children with CI is age appropriate. Accuracy development of early implanted children with CI was examined immediately after cochlear implantation up to age 2;0 and also the longer term evolution up to age 5;0 were scrutinized. Phonemic accuracy was operationalized as the dynamic Levenshtein distance (LD) between the child’s spontaneous word productions and the attempted target words. In addition, for the children with CI the effect of length of implant use (or hearing age) and age of CI activation on LD were studied.

3.4.1 Development of accuracy in children with NH and children with CI

How does phonemic accuracy develop in children with CI in comparison to children with NH? The main result of this study, based on naturalistic longitudinal data, is that at the onset of word production, children with CI’s renditions of adult words are
Levenshtein distance

consistently less accurate than those of age mates with NH: however, this difference is statistically not significant. This finding contrasts with reports in the literature: The phonemic accuracy of young children with CI has been shown to be significantly lower than the accuracy of children with NH (Schauwers et al., 2008b; van den Berg, 2012; Warner-Czyz, 2005). Note that Warner-Czyz (2005) compared children with NH and children with CI on their lexical age, i.e. up to six months after the onset of meaningful speech, while this study compared children with NH and children with CI on chronological age. A second main finding of the present study relates to the development of phonemic accuracy. At later ages and with prolonged device use production accuracy has been found to be lower in children with CI as compared to children with NH (Eriks-Brophy et al., 2013; Ertmer & Goffman, 2011; Ertmer et al., 2012). Unlike these findings, no main effect of hearing status was found in the cross-sectional data analyzed in the present study at ages 3;0, 4;0 and 5;0, indicating that there is no statistically significant difference in phonemic accuracy of children with CI and age-matched children with NH. Thus the outcomes of the present study agree with previous reports in the sense that phonemic accuracy is found to be lower in children with CI in comparison to age mates with NH. But our findings are conflicting with the current literature in the sense that the difference in accuracy is not found to be statistically significant in the present study.

How can this discrepancy be explained? Obviously there are quite a few factors, which influence the outcome of speech and language development after pediatric cochlear implantation, including child related factors (e.g., gender, the etiology of deafness, additional disabilities), audiological factors (e.g., bilateral auditory stimulation, either with a second CI or contralateral hearing aid) and environmental factors (e.g., communication mode, parental involvement in the rehabilitation), in addition to the variability among children, which is intrinsic in the speech and language development process (Boons, De Raeve, Langereis, Peeraer, Wouters & Van Wieringen, 2013).

The difference between our results up to age 5;0 and the results of Eriks-Brophy et al. (2013), Ertmer and Goffman (2011) and Ertmer et al. (2012) may be due to a difference in mean age at implantation of the children with CI. Whereas the mean age
at implantation was 2;0 (SD = 1;1) in Eriks-Brophy et al. (2013) and 1;6 (SD = 0;6) in Ertmer and Goffman (2011) and Ertmer et al. (2012), the mean age at implantation is considerably lower in the present study (1;0, SD = 0;5). Furthermore, Eriks-Brophy et al. (2013) compared children with NH to a group of both children with CI and children with hearing aids. These methodological differences possibly explain why Eriks-Brophy et al. (2013), Ertmer and Goffman (2011) and Ertmer et al. (2012) found a main effect of hearing status up to age 5;0 whereas this study did not. However, as the same participants with CI were studied in Schauwers et al. (2008b), van den Berg (2012) and the present study, age at implantation cannot explain the discrepancy in findings.

In addition, there are several other important methodological differences between the present study and the studies reported in the literature. First of all, the speech samples analyzed in order to establish phonemic accuracy vary from rigidly elicited speech to completely unrestrained spontaneous speech. In the present study, spontaneous speech is used. Eriks-Brophy et al. (2013), Ertmer and Goffman (2011) and Ertmer et al. (2012) studied elicited speech using standardized tests. It is unclear what the exact impact of the speech sampling context is for phonemic accuracy in young children, be it children with NH or children with CI. Our study reveals that in spontaneous speech samples, the children with CI are less accurate than their age-mates with NH, however the difference is statistically not significant and hence may be due to “sampling error”. However the studies of Ertmer and colleagues seem to indicate that (in a clinical context) children with CI are outperformed by children with NH in elicited speech. This apparent opposition requires further clarification.

The most important methodological difference between our results and the results of Eriks-Brophy et al. (2013), Ertmer and Goffman (2011), Ertmer et al. (2012), Schauwers et al. (2008b), van den Berg (2012) and Warner-Czyz (2005) is the applied measure of accuracy. In most cases a standardized accuracy measure was used. Eriks-Brophy et al. (2013) used the GFTA-2. Ertmer and Goffman (2011) as well as Ertmer et al. (2012) calculated PCC, and Schauwers et al. (2008b) applied pMLU. As indicated in the introduction PCC and pMLU are no optimal measures of accuracy for several reasons. Warner-Czyz (2005) evaluated consonants based on their agreement
Levenshtein distance

in manner and place, but not in voice and evaluated accuracy of vowels based on their agreement in height and backness. van den Berg (2012) used a restrictive accuracy measure: A word was considered to be incorrect even if only one phoneme deviated from the target. Hence her measure was overly conservative in computing phonemic accuracy. In contrast, the present study used Levenshtein distance, which is a much more fine-grained measure: Unlike pMLU and PCC, LD penalizes all types of speech errors, including insertions. In addition, LD considers the frequency of speech errors or deviations. Therefore, frequent and less severe speech errors are penalized to a smaller extent than infrequent and severe speech errors. The use of this more refined measure of accuracy arguably influences the accuracy results of children with CI and children with NH.

Even though the same participants with CI were studied in Schauwers et al. (2008b), van den Berg (2012) and the present study, different results in accuracy up to age 2;0 are found. In addition, our results up to age 5;0 also differ from Eriks-Brophy et al. (2013), Ertmer and Goffman (2011) and Ertmer et al. (2012). In all mentioned studies, differences between children with NH and children with CI are found, while in our study, no main effect of hearing status is found. This discrepancy can be explained by methodological differences, including the language studied, the type of speech studied, age at implantation of children with CI, statistical methods used to map variation, the number of children with NH in the control groups and the period studied and the applied measure of accuracy.

Our results have to be considered with some care, as relatively limited speech samples were available for a relatively limited number of children with CI. Even though the use of multilevel modeling takes into account variation in the data, such as the unequal number of children in each group and the unequal number of word tokens. Obviously, more word tokens and more children with CI would increase the accuracy of the estimates. In addition, a phonemic broad transcription is applied in the present study. Our results indicate that at the phonemic level children with CI become as accurate as children with normal hearing. But a broad phonemic transcription does not take into account fine phonetic variation and articulatory distortions, such as lisps. Adding such phonetic detail to the transcriptions to compute the LD measure would
add a phonetic layer to the present study, which was not the current aim. A narrow phonetic transcription could fine-tune the LD calculations. But, a narrow transcription is evidently much more time consuming than a broad transcription. This implies that in the same amount of transcription time, a smaller number of productions can be transcribed, which in turn reduces the precision of the estimates in the statistical analyses. Striking the balance between time investment and amount of material to be analyzed is a delicate matter in this respect. In the present study broad phonemic transcriptions were used, leaving the unexplored phonetic variation on our research agenda.

3.4.2 Effect of target word syllable length and target word complexity

Our results suggest effects of both target word complexity and target word syllable length in the children with NH and children with CI. Target word complexity and target word syllable length influence accuracy significantly: Phonemic accuracy decreases with increasing target word syllable length and increasing target word complexity. Thus, the more syllables in the target word, the less accurate the child’s rendition, and, the more complex the target word, for instance caused by the presence of consonant clusters, the less accurate the child’s rendition of that word. This is in line with the literature on children’s repetitions of nonwords, both in children with NH and in children with CI (Burkholder-Juhasz et al., 2007; Gathercole et al., 1991; Macrae, 2013; Nittroer et al., 2014; Von Mentzer et al., 2015), showing that accuracy decreases with increasing syllable length and with increasing phonological complexity of the target nonwords. In contrast to these studies, our results are not based on nonword repetition but on spontaneous speech. Nevertheless, our results are highly similar. The length of the target word is decisive, as phonemic accuracy decreases with increasing number of syllables in the target word. The same is true for target word complexity: Accuracy decreases when word complexity increases. Our results show that the impact of target word complexity and target word syllable length decreases in children with NH. A longitudinal analysis of children with CI showed comparable results. The impact of target word complexity and target word syllable length decreases when children with CI have more hearing experience. Thus, initially, target word complexity and target word syllable length have a significant, negative
Influence on accuracy. Even though this influence is still significant at age 5;0, our results suggest that the influence diminishes.

Interestingly, the syllable length of target words and their complexity influence phonemic accuracy more dramatically in children with CI as compared to children with NH. Consequently, the phonemic accuracy of children with CI is lower than the phonemic accuracy of children with NH when producing longer and more complex words. Thus children with CI are found to be less accurate than children with NH. However, our results suggest that this difference is subtler than a main effect of hearing status as reported in Eriks-Brophy et al. (2013), Ertmer et al. (2012), Schauwers et al. (2008b) and van den Berg (2012). Our analyses indicate that not hearing status as such, but the interaction between hearing status and target word complexity and target word length is fundamental to understand the difference between children with NH and children with CI. The difference between the two groups is to be found in the production of longer and more complex words and not so much in short and simple words.

Between the ages of 1;2 and 2;0, the decrease of phonemic accuracy with increasing word complexity is more outspoken in children with CI as compared to children with NH. This complements the findings of Schauwers et al. (2008b) that children with CI target less complex words than children with NH: The pMLU of the target words is significantly lower in children with CI. In a similar vein, van den Berg (2012) showed that children with CI use relatively more monosyllabic words than children with NH in their second year of life. Thus they acquire relatively simpler and less complex words, and with increasing complexity their accuracy is significantly lower than the accuracy of children with NH.

At ages 3;0 and 4;0, the interactions between hearing status and target word length and complexity remain. Phonemic accuracy decreases more dramatically in children with CI as compared to children with NH with increasing target word complexity and increasing target word syllable length respectively. Phonological processing of more complex words is thus more difficult for children with CI up to age 4;0. However the difference between the two groups has disappeared by age 5;0. By that age children with CI appear to have caught up with their peers with NH. This striking conclusion
holds for phonemic accuracy and is highly similar to Nicholas and Geers (2007) who concluded that children with CI have caught up with their peers with NH at age 4;6 for measures pertaining to for vocabulary, sentence complexity and morphology. For Dutch children with CI, this is also true for phonemic accuracy.

3.4.3 Effect of length of implant use and age at implantation

Speech and language development of children with CI is often reported to depend on the age at implantation (Nicholas & Geers, 2007; Schauwers et al., 2008b; van den Berg, 2012) and length of device use (Blamey et al., 2001; Tomblin et al., 2008). Our results suggest that the age at implant activation has no significant effect on phonemic accuracy. This might be due to the fact that all participants in this study were implanted at a very young age, i.e. before age 1;8. This finding is in line with Szagun and Stumper (2012) who concluded that language development of children with CI implanted before the end of what they call “the sensitive period” is not significantly affected by the age of implantation. In contrast, length of implant use influences language development in early implanted children with CI (Szagun & Stumper, 2012). In agreement with Szagun and Stumper (2012), this study suggests that length of implant use influences the development of phonemic accuracy significantly. This is also in line with the focus on length of implant use in other long-term effect studies (Blamey et al., 2001; Tomblin et al., 2008), even though the participants in these studies were implanted after the age of 2;0. Thus, for young implanted children with CI, accuracy is susceptible to the length of implant use and not so much to the age of implant activation. But in the group of children with CI investigated in the present study (age at activation between ages 0;6 and 1;8), the influence of word complexity is smaller in children with CI with later implant activation.

3.4.4 Conclusion: clinical and theoretical implications

Our results suggest that for children with CI the accurate production of more complex and longer words are especially problematic in comparison with their age-mates with NH up to age 4;0. Children with CI are still found to be delayed with respect to the phonemic accuracy of long and complex words. Their accuracy is comparable to peers with NH but only for short words. This suggests a focus for speech and language
therapy: with more hearing experience, children with CI appear to master shorter words well at a segmental level, but longer, more complex words remain problematic. Hence, clinical intervention might take up this finding and integrate more materials varying in length and complexity, provided that is not the case yet. Rehabilitation programs should also focus on the underlying factors accounting for the observation that phonemic accuracy deteriorates with increasing word length and complexity. In the literature some possible causes have been suggested which may be an additional focus of rehabilitation programs. First of all, Houston and Bergeson (2014) showed that children with CI are less attentive to speech sounds in the ambient language. Moreover they have a delay in integrating audiovisual information (Houston, Stewart, Moberly, Hollich & Miyamoto, 2012). As a result, they are found to have poorer phonological awareness and less precise phonological representations (Lund, Werfel & Schuele, 2015). Consequently, children with CI may benefit from an enhancement of their attention to speech and to speech related information, including audiovisual correlates, as this may result in better phonological awareness and representations. Secondly, children with CI are typically found to have poorer phonological short-term working memory, which results in poorer performance in more complex tasks, such as the repetition of an increasing number of digits (Kronenberger, Beer, Castellanos, Pisoni & Miyamoto, 2014; Pisoni & Cleary, 2004; Pisoni, Kronenberger, Roman & Geers, 2010). A similar effect is found in the present study. Children with CI produce longer and more complex words less accurately. In other words, enhancing the phonological working memory skills of children with CI as part of their rehabilitation may be beneficial for their processing of longer and more complex words.

To conclude, the present study has shown that the accuracy of children with CI is lower as compared to peers with NH, but only in interaction with target word complexity and target word syllable length. This reveals the importance of including explanatory variables when comparing both groups of children. In other words, it seems relevant to compare children with NH and children with CI not only as such, but it seems also relevant to include interactions with other variables. Moreover, target word complexity and target word length are also shown to influence accuracy of children with NH. This suggests that similar factors influence accuracy in children
with NH and children with CI, even though the effect is more outspoken in children with CI. Thus, target word complexity and syllable length of the target word are decisive factors in accurate producing words. Less complex and shorter words are produced more accurately than more complex and longer words. This is in line with other studies on factors influencing accuracy. As shown, those factors influence the accuracy of children with CI more severely as compared to peers with NH. Furthermore, the results showed that accuracy is more susceptible to hearing age than to age of implant activation in early implanted children with CI. Thus, length of implant use can be seen as a good predictor variable of accuracy development in children with CI.
Chapter 4
4 Intraword variability in children with cochlear implants: comparison to peers with normal hearing

ABSTRACT

Purpose: The present study evaluates intraword or token-to-token variability in the spontaneous speech of Dutch-speaking children with cochlear implants (CI) longitudinally up to five years of age in comparison with intraword variability in normally hearing (NH) age-matched peers.

Method: Spontaneous speech samples of 9 children with CI were collected longitudinally up to age five. The data of the NH control group consisted of cross-sectional recordings. Children’s word productions were categorized into 4 response types of the Variability score (consistent correct, consistent incorrect, variable with hits, variable with no hits) and the Proportion of Whole-word Variation was calculated.

Results: Intraword variability is high in both groups of children, but decreases with age. Children with CI are more variable than their peers with NH up to age four, but catch up by age five. Longer words are produced more variable in both groups of children.

Conclusions: Intraword variability is characteristic of children with CI’s spontaneous speech productions. Group differences show higher rates of intraword variability in children with CI up to age four, but children with CI catch up on their peers with NH by age five. Finally, the number of syllables seems crucial for intraword variability.

Keywords: Phonemic accuracy; children with CI; intraword variability

4.1 Introduction

A well-known characteristic of children’s early word productions is their variability: children produce specific words in different ways at different occasions. For instance, in a recording of a Dutch-speaking boy at age 1;11.08 (years;months.days), the child produced the proper name Dominiek /dominik/ in at least five different ways: [dɔmənik], [dɔnìnik], [mik], [ɔmìk], [mənik], [dənik] (Gillis, 2000). This type of
Intraword or token-to-token variability in normally hearing (henceforth: NH), typically developing children’s early word productions has received ample attention in the literature (i.a. Sosa, 2015). Whether this phenomenon is equally frequent and develops in a similar fashion in congenitally hearing impaired children who received a cochlear implant (henceforth: CI) in comparison with their normally hearing peers, has hardly been investigated. Hence, little is known about intraword variability in the productions of children with CI and the longitudinal development of this variability has not been studied yet. In the present paper, the longitudinal development of intraword variability is studied in Dutch-speaking early implanted children with CI and compared to that of their peers with NH.

In what follows, we will first focus on how to measure intraword variability, and we will briefly review the relevant literature on intraword variability in children with NH and children with CI.

4.1.1 Definition and measures of intraword variability

Intraword variability denotes the phenomenon that multiple productions of a particular adult target word (i.e. the target) differ at a phonemic level from one another in a child’s renditions (i.e. replicas) of that target word at a particular age, regardless of the accuracy of those different productions. For instance, if the adult target word /buk/ (book) is rendered once as /bu/ and once as /buk/, the child produces two replicas of the target word at that particular age. Intraword variability is frequent in the speech of children with NH from word onset onwards (Ferguson & Farwell, 1975; Holm, Crosbie & Dodd, 2007; Macrae, 2013; McLeod & Hewitt, 2008; Sosa, 2015; Sosa & Stoel-Gammon, 2006; Sosa & Stoel-Gammon, 2012; Taelman & Gillis, 2002). Even though intraword variability decreases with age (Holm et al., 2007; Sosa, 2015), it is still found in children with NH between five and ten years of age (de Castro & Wertzner, 2011).

Two measures of intraword variability are commonly used in the literature. First of all, intraword variability is assessed by the Proportion of Whole-Word Variation (PWV) measure (Ingram, 2002). PWV computes the proportion of variability per word, with 1 indicating complete variability, i.e., each instance of a particular target word is pronounced differently (all replicas differ at a phonemic level), and 0
indicating consistent production, i.e., all instances of a particular target word are pronounced in the same way (all replicas are identical at a phonemic level). In other words, the degree of variability per target word is assessed. Multiple different productions (replicas) of a particular target word receive a higher PWV score than multiple similar or phonemically identical productions (replicas) of that word.

A second method to assess intraword variability categorizes the child’s replicas per attempted target word (Holm et al., 2007; McLeod & Hewitt, 2008; Sosa, 2015), i.e. the Variability score, also referred to as Inconsistency score (Holm et al., 2007)). The Variability score categorizes each target word, i.e. the adult form, relative to the children’s renditions of that target. Four possibilities are usually distinguished: (1) all children’s productions are similar and correct (consistent correct), (2) all children’s productions are similar, but incorrect (consistent incorrect), (3) children’s productions differ, but at least one is correct (variable with hits), and (4) children’s productions differ and none of them is correct (variable with no hits). In contrast to the PWV measure, the Variability score takes the accuracy of the child’s production into account, as it distinguishes between correct and incorrect production.

### 4.1.2 Intraword variability in children with NH

Especially at younger ages, and hence in earlier stages of lexical development, children with NH exhibit elevated rates of intraword variability. For instance in binomial classifications (all replicas are phonemically similar vs. there are phonemic differences), variability scores amount to 78% for English-speaking children at a mean age of 2;04 (Macrae, 2013) and to a mean score of 74% for Dutch-speaking children between ages 1;00 and 2;09 (Taelman & Gillis, 2002). PWV scores are relatively high between one and four years of age (Sosa, 2015). For instance in spontaneous speech, PWV scores are approximately 46% at age 1;00 and 37% for the same group of children at age 2;00 (Sosa & Stoel-Gammon, 2006). Thus, the intraword variability is relatively high, but decreases with age. Similarly, Sosa and Stoel-Gammon (2012) reported that almost half of the productions of each target word are variable (49%) in children aged two to two and a half. For words with consonant clusters, the mean PWV score is about 54% between two and a half and three years of age (McLeod & Hewitt, 2008). In contrast, Holm et al. (2007) found a much lower PWV percentage.
IW variability

For instance between 3:00 and 3:05, the mean PWV score was 13% and between 6:00 and 6:11 the mean PWV was only 2.50%. Even though these figures also show a decrease with age, they are much lower than the percentages which are generally reported in the literature. In any case, we can assume that intraword variability, i.e. PWV, is high in children with NH from word onset up to approximately four years of age.

With respect to the Variability score, both McLeod and Hewitt (2008) and Sosa (2015) have shown that all 4 response types (consistent correct, consistent incorrect, variable with hits and variable with no hits) are common in typical child language up to age four. In contrast, Holm et al. (2007) indicated that the variable with no hits response type was rare in children with NH’s productions from age three up to age seven. As for PWV, the results of Holm et al. (2007) differ from other studies.

Several factors influence the likelihood of variable production of target words. A child-related factor is obviously age: intraword variability decreases with age. Another child-related factor is vocabulary size: intraword variability is lower in children with larger vocabularies (Macrae, 2013; Sosa & Stoel-Gammon, 2012). Also characteristics of the adult target word appear to influence the likelihood of variable production. For instance target words including late acquired segments (e.g. fricatives or consonant clusters) and longer target words in terms of syllable length are shown to have higher variability rates (Leonard, Rowan, Morris & Fey, 1982; Macrae, 2013; McLeod & Hewitt, 2008; Sosa, 2015; Sosa & Stoel-Gammon, 2012). Thus the phonological complexity of adult target words also influences the intraword variability in children’s production. The present paper analyses the impact of the target word’s syllable length and the effect of age on the variability in production.

Most studies of intraword variability which use PWV and the Variability score, are mainly restricted to English-speaking children with NH. But their methodologies vary from the point of view of the study design (longitudinal (McLeod & Hewitt, 2008; Sosa & Stoel-Gammon, 2006) vs. a single point in development (Holm et al., 2007; Sosa, 2015; Sosa & Stoel-Gammon, 2012)), the type of speech studied (spontaneous speech (McLeod & Hewitt, 2008; Sosa & Stoel-Gammon, 2006; Sosa & Stoel-Gammon, 2012) vs. picture-naming tasks (Holm et al., 2007; Sosa, 2015)), the target
words (determined in advance (Holm et al., 2007; McLeod & Hewitt, 2008; Sosa, 2015; Sosa & Stoel-Gammon, 2012) or not (Sosa & Stoel-Gammon, 2006)). Also the computation of the relevant measures differ: in some studies vowels are excluded from the variability counts (Sosa & Stoel-Gammon, 2006; Sosa & Stoel-Gammon, 2012), while in others vowels are included (Holm et al., 2007; McLeod & Hewitt, 2008; Sosa, 2015). Despite these methodological differences, the general conclusions are highly similar: intraword variability is high in the early stages of language development of children with NH, decreases with age, but increases with increasing (syllable) length of the adult target word.

### 4.1.3 Intraword variability in children with CI

So far, few studies have investigated intraword variability in the speech of children with CI. Ertmer and Goffman (2011) studied intraword variability in 6 English-speaking children with CI in comparison with 6 peers with NH at a mean age of 4;00, using a toy-naming task with 3 attempts for each of the 60 toys per child. The children with CI were implanted before the age of 3;00 and had at least two years of hearing experience at the time of testing. The children with NH were matched on chronological age. All target words were mono- and disyllables – except for one word with three syllables banana – and the target words varied in phonemic complexity of the onset. In other words, target words differed in the manner of articulation of the first consonant, considering target words with for instance a word initial fricative as more complex than those with an initial stop. Results showed that the PWV score was higher for children with CI than for their peers with NH and that both groups of children had higher PWV scores in target words with word initial fricatives, affricates and liquids (Ertmer & Goffman, 2011). Thus, the influence of phonological complexity is similar in both groups of children.

Moreno-Torres (2014) investigated intraword variability in the spontaneous speech of 8 Spanish-speaking children with CI and 3 peers with NH. The children with CI were followed longitudinally as part of a larger study, the data of the children with NH were extracted from the CHILDES database. Only one speech sample of each CI and NH child was selected for analysis of intraword variability, viz. the sample in which the MLU was closest to 1.2. The mean chronological age of the children with CI was
IW variability

2;09 (SD = 0;04) and the mean age at implant activation was 1;04 (SD = 0;03), hence approximately a year and a half of device experience. A total of 159 target words were analysed of children with CI and 72 of children with NH. The first two replicas of these target words were selected. Thus, the total number of analysed tokens is relatively low (159 x 2, 72 x 2). The results show that the PWV is considerably higher in children with CI (0.63) than in children with NH (0.15). Moreover, there are differences between both groups of children regarding the response types of the Variability score. The children with CI mainly had variable with no hits and consistent incorrect responses, while children with NH mainly had consistent incorrect and consistent correct responses. Thus, children with NH’s productions were more consistent and less variable. Furthermore, the number of syllables in the target words had an effect on intraword variability: variability was slightly higher in longer words. This effect was found in the NH group, but not in the CI group. In this respect, Moreno-Torres (2014)’s findings contrast with those reported by Ertmer and Goffman (2011).

Even though there is no information in the literature concerning intraword variability of Dutch-speaking children with CI, Faes, Gillis and Gillis (2016) reported that the syllable length of target words affects whole-word accuracy more in Dutch-speaking children with CI than in children with NH. Schauwers, Taelman, Gillis and Govaerts (2008) have shown that between 2;00 and 2;06 children with CI target shorter words, which may explain the difference between children with NH and children with CI regarding the effect of a target word’s syllable length in Moreno-Torres (2014). Moreno-Torres (2014) found that longer target words are more variable in children with NH. But, if the general target word’s syllable length is shorter – as in children with CI (Schauwers et al., 2008), it is quite possible that the effect is not visible in children with CI (yet). In addition, Moreno-Torres (2014) included only segmental differences in the variability measure and did not consider differences in the number of syllables of the children’s production. This may have influenced the results in that his measure did not account for the effect of syllable length in word productions.
Thus, Ertmer and Goffman (2011) and Moreno-Torres (2014) found higher rates of intraword variability in children with CI than in children with NH. In addition, Ertmer and Goffman (2011) showed that the phonological complexity of speech sounds affects intraword variability not only in children with NH but also in children with CI. In contrast, Moreno-Torres (2014) found an effect of target word syllable length on intraword variability in children with NH, but not in children with CI. However, those results may be influenced by the fact that syllable omissions were not considered. In the present study the effect of syllable length and differences in the number of syllables of target words will be studied in children with NH and children with CI acquiring Dutch.

4.1.4 The present study

In the present paper, intraword variability is analysed longitudinally in the spontaneous speech of Dutch-speaking children with CI in comparison with NH age-matched peers. So far, no information is available on intraword variability of Dutch-speaking children with CI nor has the development of intraword variability been studied longitudinally in children with CI. We aim to analyse the spontaneous speech of children with CI longitudinally from their first word productions up to the age of five. The literature on children with NH there is a substantial intraword variability from word onset up to four years of age (Sosa, 2015) and even up to the older ages (de Castro & Wertzner, 2011). Therefore, we aim to study the intraword variability of children with CI longitudinally between word onset and five years of age, instead of describing only one point in development as in Ertmer and Goffman (2011) and in Moreno-Torres (2014). Moreover, the age at implant activation is included into the longitudinal analysis of children with CI, as it is often shown that earlier implanted children with CI have better language outcomes than their later implanted peers (e.g. Niparko, Tobey, Thal, Eisenberg, Wang, Quittner & Fink, 2010). In addition, we also aim to compare the development of intraword variability in children with CI to that of peers with NH.

In contrast to Moreno-Torres (2014), the number of target words is considerably larger in the present study (14,510) and our variability measures penalize also syllable omissions. In other words, more target words are studied and each difference in
IW variability

production (also syllable omissions and vowel differences) is considered as a variable production. In addition, the present study does not only compare the PWV scores (like Ertmer & Goffman, 2011; Moreno-Torres, 2014) of children with CI and children with NH, but also the likelihood of the different response types of the Variability score (consistent correct, consistent incorrect, variable with hits, variable with no hits). To the best of our knowledge, this has never been studied in children with CI.

4.2 Method

4.2.1 Participants

A longitudinal design was set up to assess language development of children with CI (S1 – S9): 9 children with CI were followed from word onset up to five years of age (Schauwers, 2006). The children were followed monthly from the moment of implant activation up to 2;6 years after activation. Additionally, yearly data were collected between ages 3;0 and 5;0. Here, the data will be analyzed from the appearance of first spoken words (median = 1;6, range 1;3 – 1;11). In chapter 2, more detailed information about the children with CI can be found.

The control group consisted of the first 10 children from the NHlongitudinal corpus at 24 months of age (S10-S19) and the NHcross-sectional corpus between ages 3;0 and 5;0. With respect to the NHlongitudinal corpus, the data will be examined from the appearance of the first spoken words (median = 1;2, range 1;2 – 1;5). With respect to the NHcross-sectional corpus, a total of 32 children with NH participated: 9 three year olds (mean = 3;0, SD = 0;1), 12 four year olds (mean = 4;0 months, SD = 0;1), 11 five year olds (mean = 5;0, SD = 0;1). More information is given in chapter 2.

4.2.2 Procedure

The data used in the current study are part of a larger corpus, namely the CLiPS Child Language Corpus (CCLC). Only samples in which lexical items appeared are analyzed. Detailed descriptions of the data annotation procedure and the reliability assessment of the transcriptions are reported in the method section (chapter 2).

For each child at each age, the replicas of each target word token were listed. A total of 24,942 target word tokens were available (NH: 9,301; CI: 15,641). Target
word tokens with only one child replica at a particular age were excluded as no intraword variability score can be derived. This resulted in a total of 16,640 target word tokens. For children with CI, 9,698 target word tokens were available and for children with NH 6,942 target word tokens. The number of word tokens per target word was restricted to 10. If a target word token had more than 10 child realizations (replicas), the target word was excluded from the data analyses. This eventually resulted in a total of 14,383 target word tokens (CI: 8,378 and NH: 6,005).

4.2.3 Variability measures

For each child at each age, the number of child replicas, the number of different child replicas and the number of correct replicas for each adult word token, i.e. adult target word attempted by the child, are calculated. Two measures of intraword variability were computed: the proportion of whole-word variation (PWV) and the Variability score with 4 response types. In calculating these measures, all segmental information was included, i.e. consonantal and vocalic differences were considered. Differences in syllable number in children’s productions were considered as well.

The first measure, PWV, was calculated as the ratio of the number of distinct child replicas of a particular adult target and the number of child tokens (Ingram, 2002). In (1) some hypothetical examples are provided to exemplify the computation of PWV. For instance, if a child produces the adult target /brur/ broer (Eng. brother) once as /bur/ and once as /bu/, the child produces two different replicas of the target /brur/. Suppose the child produces both of these replicas only once, then there are two tokens of the attempted adult word. In this case, PWV equals 2/2 = 1.0. Suppose the child attempts to produce the adult target /pus/ poes (Eng. cat) six times, three times as /pupu/, twice as /pu/, and once as /pu/. In this example the adult target /pus/ results in three distinct replicas (/pupu/, /pu/ and /pus/), and the total number of child tokens equals six. PWV consequently equals 3/6 = 0.5. With respect to the third example in Table 1, Ingram (2002) argues that if the child uses the adult target word /buk/ boek (Eng. book) each time as /bu/, there is only one child replica and, hence, zero different replicas. This results in a PWV of 0/5 = 0. This indicates that the child produces that word consistently. It can readily be seen that the resulting PWV scores range between 0 and 1.
IW variability

Table 1. Examples of the PWV calculations

<table>
<thead>
<tr>
<th>Adult target</th>
<th>Child’s replicas</th>
<th>Number tokens replica</th>
<th>Number of distinct replicas</th>
<th>Number of child tokens</th>
<th>PWV</th>
</tr>
</thead>
<tbody>
<tr>
<td>/brur/ (Eng. brother)</td>
<td>/bur/</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>/bu/</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/pus/ (Eng. cat)</td>
<td>/pupu/</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>/pu/</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>/pus/</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/buk/ (Eng. Book)</td>
<td>/bu/</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>0</td>
</tr>
</tbody>
</table>

For each target word, the Variability score was determined as well. The Variability score includes 4 categorical response types:

- Consistent correct: the children’s replicas are phonemically identical and match the target
- Consistent incorrect: the children’s replicas are phonemically identical, but do not match the target
- Variable with hits: the children’s replicas are phonemically variable, but at least one matches the target
- Variable with no hits: the children’s replicas are phonemically variable, with none matching the target

Table 2 gives an example of the four response types for the target word /buk/ (boek, Eng. book).

Table 2. Examples of the Variability score

<table>
<thead>
<tr>
<th>Target</th>
<th>Replica</th>
<th>Variability score</th>
</tr>
</thead>
<tbody>
<tr>
<td>/buk/</td>
<td>/buk/, /buk/, /buk/</td>
<td>Consistent correct</td>
</tr>
<tr>
<td>/buk/</td>
<td>/bup/, /bup/, /bup/</td>
<td>Consistent incorrect</td>
</tr>
<tr>
<td>/buk/</td>
<td>/bup/, /buk, /pu/</td>
<td>Variable with hits</td>
</tr>
<tr>
<td>/buk/</td>
<td>/bup/, /bu/, /pu/</td>
<td>Variable with no hits</td>
</tr>
</tbody>
</table>
4.2.4 Statistical analyses

Statistical analyses are performed in R by means of multilevel models (R Core Team, 2013). Multilevel models include two parts: a random part and a fixed part. The random part considers the variation in the data caused by i.a. the nesting of variables (Baayen, 2008; Woltman, Feldstain, MacKay & Rocchi, 2012). The fixed part includes the predicting variables.

Two different measures of variability were examined. First of all, the PWV score was analyzed in a multilevel model and was scaled between 0 and 1, with 1 being completely variable and 0 being completely consistent. Secondly, the Variability score was analyzed using a logistic regression in a multilevel model. In other words, the likelihood of each response type (as compared to all other response types) was examined. These results were expressed in logits. For each analysis, only those target words were included that had between 2 and 10 child productions.

For each variability measure (PWV, Variability score), there were two parts in the analyses. Firstly, the development of intraword variability in children with CI was traced longitudinally. Therefore, the monthly data from word onset (median = 1;06, range 1;03 – 1;11) up to age 2;06 and the yearly data at ages 3;00, 4;00 and 5;00 were included in the analyses. Secondly, cross-sectional comparisons of children with CI and age-matched children with NH were performed at ages 2;00, 3;00, 4;00 and 5;00. Age ranges for the children with CI were 1;11 – 2;01, 2;10 – 3;04, 3;09 – 4;03 and 4;11 – 5;03 respectively. For the longitudinal analyses, the intercept was set at age 1;10 as we have the largest amount of data at this age point. In each analysis, the fixed effects were age (Age), the number of syllables (Syllable length) and the age at implant activation (CIactivation). Interactions between those variables as well as quadratic and cubic age effects were tested and included if they improved the model fit. For the cross-sectional analyses, the fixed effects were hearing status (Hearing status) and the number of syllables (Syllable length), and interactions between Hearing status and Syllable length were included if a better model fit was obtained. In order to model the variation in the data, we allowed random intercepts and slopes in the longitudinal analyses and random intercepts for the cross-sectional analyses. A significance level of p<0.05 was set.
IW variability

4.3 Results

4.3.1 Longitudinal analysis of children with CI

The first part of the analyses presents the results of the longitudinal analyses of children with CI for PWV and the Variability score respectively.

4.3.1.1 PWV

A first measure of intraword variability is PWV, which varies between 0 and 1. Table 3 specifies the fixed effect results for the longitudinal analyses of children with CI. The results are also plotted in Figure 1. At the intercept, the estimated average PWV score equals 0.27, which is well above zero (p<0.001), but decreases significantly with Age (p<0.001). In other words, there is intraword variability in children with CI, but it decreases with age (Figure 1). In addition, there is a significant effect of Syllable length (p<0.01, Table 3), which is such that the PWV is higher for longer target words. Nevertheless, the significant interaction effect between Syllable length and Age (p<0.01) indicates that this rising of PWV with increasing syllable length of the target word becomes less pronounced over time. Finally, the effect of CIActivation is not significant (p>0.05, Table 3), which means that the PWV scores are similar for children with CI with earlier and later implant activation. There was no interaction effect between CIActivation and Age. Since it did not improve the model fit, it was not included in the model reported here.

Table 3. Fixed effect results of the longitudinal analysis of children with CI – PWV

<table>
<thead>
<tr>
<th></th>
<th>Estimate (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.27 (0.04) ***</td>
</tr>
<tr>
<td>Age</td>
<td>-0.01 (&lt;0.01)***</td>
</tr>
<tr>
<td>Syllable length</td>
<td>0.02 (0.01) **</td>
</tr>
<tr>
<td>Syllable length x Age</td>
<td>&lt;-0.01 (&lt;0.01)**</td>
</tr>
<tr>
<td>CIactivation</td>
<td>&lt;0.01 (&lt;0.01)</td>
</tr>
</tbody>
</table>

* p≤0.05, ** p≤0.01, *** p ≤0.001
4.3.1.2 The Variability score

A second measure of intraword variability is the Variability score. PWV did not consider the degree of correctness of children’s productions. In this second part of the analyses, correctness is taken into account by examining the likelihood of each Variability score, i.e. each of the four response types (consistent correct, consistent incorrect, variable with hits and variable with no hits). The likelihood of each response type was estimated relative to all other response types. The results are shown in Table 4 and in Figure 2. The results reported in Table 4 and Figure 2 are expressed in logits, but in the discussion of these results, probabilities are provided.

First, the consistent correct responses are not significantly different form a logit 0, i.e. 50% at the intercept (p>0.05). This indicates that about half of the responses were consistent and correct at the intercept (67.48%). Figure 1 suggests that there is an increase of the likelihood of consistent correct responses with age, but the effect of Age is only borderline non-significant (p=0.06). The quadratic effect of age was not significant and because it did not improve the model fit, it was left out. In other words, there is a linear effect of age in the period studied. Next, Syllable length affects the likelihood of consistent correct responses significantly (p<0.001): consistent correct responses are less likely when the target word has more syllables. However, the

Figure 1. Development of PWV with age for children with CI – predicted values
IW variability

significant interaction between Syllable length and Age (p<0.05) indicates that this effect becomes less pronounced as children get older. There is a significant effect of CI activation (p<0.001), which shows that children with CI with later implant activation are less likely to have consistent correct responses at the intercept. At the intercept, children with later implant activation have i.a. a smaller vocabulary size, which may very well explain the difference. But, this effect decreases as children with CI get older, which is suggested by the significant interaction between CI activation and Age (p<0.001). In other words, children with later CI implant activation catch up on the likelihood of consistent correct responses.

Secondly, consistent incorrect responses are significantly less likely than all other response types at the intercept: 32.96% (p<0.001). Even though Figure 2 suggests a slight decrease with age, there is no significant effect of age (p>0.05) on the likelihood of consistent incorrect responses. However, the Syllable length of target words affects the likelihood of consistent incorrect responses significantly (p<0.001): this likelihood increases if the target word has more syllables. Moreover, this effect becomes more pronounced with age, as is shown by the significant interaction between Syllable length and Age (p<0.01). The effect of CI activation is significant as well (p<0.01), indicating that children with CI with later implant activation are more likely to have consistent incorrect responses. The interaction effect of CI activation and Age is marginally significant (p=0.054). This indicates that the difference between children with CI with earlier and later implant activation becomes smaller over time, or, in other words, that children with CI with later implant activation catch up for the likelihood of consistent incorrect responses.

Thirdly, the variable with hits responses do not differ significantly from a likelihood of 50% at the intercept (p>0.05). The estimate at the intercept is 30.15%. There is a quadratic effect of age, showing an inverse u-shaped development (Figure 2): the likelihood of consistent incorrect responses first increases, and subsequently decreases. Next, there is a significant effect of Syllable length (p<0.001), indicating that the likelihood of variable with hits responses decreases for longer target words. Thus, longer words are less often variable with hits than shorter words. In addition, the significant interaction between Syllable length and Age (p<0.05) shows that this
effect – decrease of variable with hits responses with increasing target word syllable length – becomes more pronounced over time. The effect of CI activation is significant as well (p<0.05), showing that children with CI with later implant activation are less likely to have variable with hits responses. There was no significant interaction between CI activation and Age (p>0.05), meaning that the difference between children with CI earlier and later implant activation remains stable with age. Nevertheless, this effect improved the model fit significantly and was therefore included in the final model reported here.

Finally, the likelihood of variable with no hits responses is significantly lower than that of all other response types at the intercept (p<0.001): 11.61%. Even though Figure 2 shows a decrease with age, this effect is not significant (p>0.05), indicating that the likelihood of variable with no hits responses remains low over time. The significant effect of Syllable length (p<0.001) points out that the likelihood of variable with no hits responses increases when target words have more syllables. This effect remains stable over time, as there is no significant interaction between Syllable length and Age (p>0.05). Next, the effect of CI activation is significant (p<0.001), showing that the likelihood of variable with no hits responses is higher in children with CI with later implant activation. However, those children with CI seem to catch up for the likelihood of variable with no hits responses, as shown by the negative, significant interaction effect between CI activation and Age (p<0.05).

Table 4. Fixed effect results of the longitudinal analysis of children with CI – The Variability score

<table>
<thead>
<tr>
<th></th>
<th>Consistent correct</th>
<th>Consistent incorrect</th>
<th>Variable with hits</th>
<th>Variable with no hits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>Estimate 0.61</td>
<td>-0.73 ***</td>
<td>-0.71 ***</td>
<td>-0.84 0.62</td>
</tr>
<tr>
<td>Age</td>
<td>Estimate 0.02</td>
<td>-0.02</td>
<td>0.10 **</td>
<td>0.04 -0.04</td>
</tr>
<tr>
<td>Age x Age</td>
<td></td>
<td></td>
<td></td>
<td>-0.01 *** &lt;0.01</td>
</tr>
<tr>
<td>Syllable length</td>
<td>Estimate 0.05</td>
<td>0.36 ***</td>
<td>-0.74 ***</td>
<td>0.33 *** 0.05</td>
</tr>
</tbody>
</table>
IW variability

| Syllable length x Age | Estimate | SE  | 0.01 * | 0.01 ** | -0.01 * | <0.01 |
| CIactivation x Age  | Estimate | SE  | -0.14 ** | 0.04 ** | -0.09 * | 0.10 *** |

* p≤0.05, ** p≤0.01, *** p ≤0.001

Figure 2. Development of the different response types of the Variability score with age for children with CI – predicted values in logits

4.3.2 Cross-sectional comparisons between children with NH and children with CI

The second part of the analysis concerns the cross-sectional comparisons of children with CI and age-matched children with NH. First, the results with respect to PWV are reported, and then those with respect to the Variability score.

4.3.2.1 PWV

Table 5 summarizes the fixed effect results of the cross-sectional comparisons
between children with NH and children with CI. These results are also presented in Figure 3. At age 2;00, Table 5 and Figure 3 show a significant effect of Hearing status (p<0.001): the PWV is higher in children with NH (0.60) than in children with CI (0.45). Moreover, the Syllable length of the target word affects the PWV significantly (p<0.001): PWV scores increase when the syllable length of target words increases. This effect is similar in both groups of children, as the interaction effect of Syllable length and Hearing status is not significant (p>0.05). The higher PWV scores in children with NH might be due to their target word characteristics. Additional analyses revealed that the target words of children with CI are significantly shorter in syllable length than those of children with NH (estimates: 0.46 in CI, 0.61 in NH (SE=0.06), p<0.01). By age 3;00, children with CI have caught up and are targeting words of similar length as compared to children with NH (estimate for both groups of children at age 3;00: 0.56 (SE=0.03), p>0.05).

The analyses at ages 3;00, 4;00 and 5;00 (Table 5 and Figure 3) show a further decrease of the PWV in children with NH (p<0.001): from 0.38 at age 3;00 to 0.22 at age 5;00. Table 5 and Figure 3 further display some effects of Hearing status: at age 3;00, children with CI are similar to their peers with NH (p>0.05). Nevertheless, at age 4;00, they are more variable than their peers with NH (p<0.01) and – by contrast – less variable than their peers with NH at age 5;00 (p<0.05). Next, there are some effects of Syllable length (Table 5). At age 3;00, longer words are more variable (p<0.05) and this is similar in both groups of children as indicated by a non-significant interaction between Syllable length and Hearing status (p>0.05). At age 4;00, the Syllable length has no effect on intraword variability in children with NH (p>0.05). However, a significant interaction between Syllable length and Hearing status (p<0.01) shows that PWV increases with increasing word length in children with CI at this age. At age 5;00, no effect of Syllable length is found (p>0.05) and this is similar in both groups of children (p>0.05).

Table 5. Fixed effect results of PWV scores – cross-sectional comparisons between NH and CI

<table>
<thead>
<tr>
<th></th>
<th>2;00</th>
<th>3;00</th>
<th>4;00</th>
<th>5;00</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CI</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
IW variability

<table>
<thead>
<tr>
<th></th>
<th>Estimate (SE)</th>
<th>Estimate (SE)</th>
<th>Estimate (SE)</th>
<th>Estimate (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.60 (0.03)***</td>
<td>0.38 (0.04)***</td>
<td>0.24 (0.03)***</td>
<td>0.22 (0.04)***</td>
</tr>
<tr>
<td>Hearing status CI</td>
<td>-0.15 (0.05)**</td>
<td>0.06 (0.05)</td>
<td>0.12 (0.05)**</td>
<td>-0.17 (0.06)**</td>
</tr>
<tr>
<td>Syllable length</td>
<td>0.12 (0.02)***</td>
<td>0.06 (0.03)*</td>
<td>0.02 (0.02)</td>
<td>-0.02 (0.02)</td>
</tr>
<tr>
<td>Syllable length x Hearing status CI</td>
<td>0.02 (0.05)</td>
<td>0.04 (0.04)</td>
<td>0.08 (0.03)**</td>
<td>0.02 (0.02)</td>
</tr>
</tbody>
</table>

*p≤0.05, **p≤0.01, ***p ≤0.001

NH is the reference category

Figure 3. Development of PWV with age for children with CI and children with NH – predicted values

4.3.2.2 The Variability score

In Tables 6 – 9, the fixed effect results of the likelihood of each response type of the Variability score are presented. Table 6 shows the results at age 2;00, Table 7 at age 3;00, Table 8 at age 4;00 and Table 9 at age 5;00. The observed values of each response type are also plotted in Figure 4.

At age 2;00 (Table 6), none of the response types is more likely than the other ones in children with NH, as is shown by significant negative intercepts (p<0.01 for
consistent correct and variable with hits, \( p<0.001 \) for consistent incorrect and variable with no hits). Estimated likelihoods are 11.01% for consistent correct responses, 12.90% for consistent incorrect responses, 37.99% for variable with hits responses and 34.98% for variable with no hits responses. With respect to the effect of Hearing status, Table 6 points out that there is no significant difference between both groups of children regarding consistent correct and the variable with no hits responses (\( p>0.05 \)). However, there is a significant effect of Hearing status regarding consistent incorrect and variable with hits responses (\( p<0.01 \)). These effects indicate that the likelihood of consistent incorrect responses is significantly higher in children with CI (25.35%) as compared to children with NH (12.90%) and the likelihood of variable with hits responses significantly lower in children with CI (21.42%) as compared to children with NH (37.99%). Apart from the group differences, Table 6 shows that also the syllable length of the target words affects the likelihood of the different response types, except for the consistent incorrect responses (\( p>0.05 \)). Consistent correct responses are less likely when the target word is larger in terms of syllable length (\( p<0.001 \)) and this effect is similar in both groups of children as the interaction between Syllable length and Hearing status is not significant (\( p>0.05 \)). Variable with hits responses are less likely with increasing syllable length (\( p<0.001 \)), but the effect is less pronounced in children with CI as compared to children with NH (interaction effect, \( p<0.05 \)). The likelihood of variable with no hits responses increases when the syllable length of the target word increases (\( p<0.001 \)), but this effect is less pronounced in children with CI as well (interaction effect \( p<0.05 \)).

Table 6. Fixed effect results for the Variability scores (expressed in logits) at age 2;00

<table>
<thead>
<tr>
<th></th>
<th>Consistent correct</th>
<th>Consistent incorrect</th>
<th>Variable with hits</th>
<th>Variable with no hits</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intercept</strong></td>
<td>Estimate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-2.09 **</td>
<td>-1.91 ***</td>
<td>-0.49 **</td>
<td>-0.62 ***</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>0.27</td>
<td>0.18</td>
<td>0.17</td>
</tr>
<tr>
<td><strong>Hearing status CI</strong></td>
<td>Estimate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.32</td>
<td>0.83 **</td>
<td>-0.81 **</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>0.48</td>
<td>0.31</td>
<td>0.33</td>
</tr>
<tr>
<td><strong>Syllable length</strong></td>
<td>Estimate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.77 ***</td>
<td>-0.34</td>
<td>-0.93 ***</td>
<td>1.16 ***</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>0.23</td>
<td>0.18</td>
<td>0.15</td>
</tr>
<tr>
<td><strong>Syllable</strong></td>
<td>Estimate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.81</td>
<td>-0.26</td>
<td>0.82 **</td>
<td>-0.70 *</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
IW variability

<table>
<thead>
<tr>
<th>length x Hearing status CI</th>
<th>SE</th>
<th>0.42</th>
<th>0.38</th>
<th>0.31</th>
<th>0.28</th>
</tr>
</thead>
</table>

* p<0.05, ** p<0.01, *** p<0.001
NH is the reference category

At age 3;00 (Table 7), about half of the children with NH’s productions are variable with hits, which is shown by a non-significant intercept (p>0.05). For the other response types, the likelihood is 10.07% for consistent incorrect responses, 10.34% for variable with no hits responses and 32.74% for consistent correct responses. Regarding children with CI, Table 7 shows a significant effect of Hearing status. As compared to children with NH, children with CI are significantly less likely to have consistent correct responses (18.24% p>0.01) and variable with hits responses (34.52%, p<0.05), whereas they are more likely to have variable with no hits responses (26.31%, p<0.001). Children with CI are also more likely to have consistent incorrect responses (16.38%), but this effect is only on the edge of significance (p=0.055). Finally, the effect of Syllable length is not significant for consistent correct and consistent incorrect responses (p>0.05). However, the Syllable length of the target word does affect the likelihood of the variable response types. Variable with hits responses are less likely with increasing syllable length (p<0.05), whereas variable with no hits responses are more likely with increasing syllable length (p<0.01). In other words, longer words are less likely to be variable with hits as compared to shorter words. These effects are similar in both groups of children, as is shown by non-significant interactions between Hearing status and Syllable length (p>0.05).

Table 7. Fixed effect results for the Variability scores (expressed in logits) at age 3;00
In Table 8, the fixed effect results at age 4:00 are presented. In children with NH, about half of the responses are consistent correct (likelihood of 51.50%) (non-significant intercept, p>0.05). The likelihood of the other response types is significantly lower than 50% (p<0.001). The consistent incorrect response type has a likelihood of 9.11%, the variable with hits response has a likelihood of 35.43%, and the likelihood of variable with no hits response type has a likelihood of 4.03%. With respect to children with CI, there are significant effects of Hearing status for the consistent correct, consistent incorrect and variable with no hits responses, whereas there is no significant effect of Hearing status for the variable with hits responses. Thus, children with CI are equally likely to have variable with hits responses as compared to their peers with NH (p>0.05). But, they are significantly less likely to have consistent correct responses (30.15%, p<0.01) and significantly more likely to have consistent incorrect responses (14.06%, p>0.01) and variable with no hits responses (17.80%, p<0.001). Finally, Table 8 shows some effects of Syllable length. There is no significant effect of Syllable length for consistent correct responses in children with NH (p>0.05), but the effect of Syllable length is significant in children with CI (interaction effect, p<0.05), showing that children with CI are less likely to have consistent correct responses in longer target words. Consistent incorrect responses are more likely with increasing syllable length (p<0.05) and this effect is similar in both groups of children (no interaction effect, p>0.05). In contrast, variable with hits responses are less likely with increasing syllable length (p<0.05) and this effect is similar in both groups of children (no interaction effect, p>0.05). The likelihood of variable with no hits responses is not affected by the syllable length of the target word (p>0.05) in none of the children (no interaction effect, p>0.05). Thus, longer target words are more likely to be consistent incorrect and less likely to be variable with hits in both groups of children, and in children with CI, longer target words are also less likely to be consistent correct.
Table 8. Fixed effect results for the Variability scores (expressed in logits) at age 4;00

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>SE</th>
<th>Estimate</th>
<th>SE</th>
<th>Estimate</th>
<th>SE</th>
<th>Estimate</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consistent correct</td>
<td>0.06</td>
<td>0.18</td>
<td>-2.30 ***</td>
<td>0.13</td>
<td>-0.60 ***</td>
<td>0.09</td>
<td>-3.17 ***</td>
<td>0.26</td>
</tr>
<tr>
<td>Consistent incorrect</td>
<td>-2.30 ***</td>
<td>0.13</td>
<td>0.09</td>
<td>0.09</td>
<td>-3.17 ***</td>
<td>0.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variable with hits</td>
<td>-0.60 ***</td>
<td>0.09</td>
<td>-3.17 ***</td>
<td>0.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variable with no hits</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hearing status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consistent correct</td>
<td>-0.90 **</td>
<td>0.28</td>
<td>0.49 **</td>
<td>0.17</td>
<td>-0.02</td>
<td>0.13</td>
<td>1.64 ***</td>
<td>0.36</td>
</tr>
<tr>
<td>Consistent incorrect</td>
<td>0.49 **</td>
<td>0.17</td>
<td>0.13</td>
<td>0.13</td>
<td>-0.02</td>
<td>0.13</td>
<td>1.64 ***</td>
<td>0.36</td>
</tr>
<tr>
<td>Variable with hits</td>
<td>-0.02</td>
<td>0.17</td>
<td>0.13</td>
<td>0.13</td>
<td>-0.02</td>
<td>0.13</td>
<td>1.64 ***</td>
<td>0.36</td>
</tr>
<tr>
<td>Variable with no hits</td>
<td>0.12</td>
<td>0.13</td>
<td>0.13</td>
<td>0.13</td>
<td>-0.02</td>
<td>0.13</td>
<td>1.64 ***</td>
<td>0.36</td>
</tr>
<tr>
<td>Syllable length</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consistent correct</td>
<td>-0.02</td>
<td>0.10</td>
<td>0.36 *</td>
<td>0.15</td>
<td>-0.26 *</td>
<td>0.12</td>
<td>0.33</td>
<td>0.20</td>
</tr>
<tr>
<td>Consistent incorrect</td>
<td>0.36 *</td>
<td>0.15</td>
<td>0.12</td>
<td>0.12</td>
<td>0.33</td>
<td>0.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variable with hits</td>
<td>-0.26 *</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.33</td>
<td>0.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variable with no hits</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.33</td>
<td>0.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Syllable length x Hearing status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consistent correct</td>
<td>-0.39 *</td>
<td>0.16</td>
<td>-0.17</td>
<td>0.20</td>
<td>-0.24</td>
<td>0.17</td>
<td>0.39</td>
<td>0.23</td>
</tr>
<tr>
<td>Consistent incorrect</td>
<td>-0.17</td>
<td>0.20</td>
<td>0.17</td>
<td>0.17</td>
<td>0.39</td>
<td>0.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variable with hits</td>
<td>-0.24</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.39</td>
<td>0.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variable with no hits</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.39</td>
<td>0.23</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001
NH is the reference category

Table 9 shows the fixed effect results at age 5;00. Similarly as at age 4;00, the likelihood of consistent correct approximates 50% in children with NH (p>0.05 at the intercept), but at this age, this is also the case for variable with hits responses in children with NH (p>0.05 at the intercept). The other response types are less likely: 9.20% for consistent incorrect responses and 2.77% for variable with no hits responses. There is a significant effect of Hearing status for the consistent correct, the variable with hits and the variable with no hits responses (p<0.001), showing that children with CI’s productions are less likely to be variable with hits (CI: 0.53%, NH: 33.18%) and variable with no hits (CI: 0.29%, NH: 2.77%), and more likely to be consistent correct (CI: 89.93%, NH: 52.00%). Finally, Syllable length affects only the likelihood of the variable response types: with increasing syllable length, variable with hits responses are less likely (p<0.01) and variable with no hits are more likely (p<0.01). These effects are similar in both groups of children, as the interactions between Syllable length and Hearing status are not significant (p>0.05).

Table 9. Fixed effect results for the Variability scores (expressed in logits) at age 5;00

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>SE</th>
<th>Estimate</th>
<th>SE</th>
<th>Estimate</th>
<th>SE</th>
<th>Estimate</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consistent correct</td>
<td>0.08</td>
<td></td>
<td>-2.29 ***</td>
<td>0.10</td>
<td>-0.70</td>
<td>0.17</td>
<td>-3.56 ***</td>
<td>0.23</td>
</tr>
<tr>
<td>Consistent incorrect</td>
<td>-2.29 ***</td>
<td>0.10</td>
<td>-0.70</td>
<td>0.17</td>
<td>-3.56 ***</td>
<td>0.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variable with hits</td>
<td>-0.70</td>
<td>0.17</td>
<td>-3.56 ***</td>
<td>0.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variable with no hits</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Chapter 4

<table>
<thead>
<tr>
<th>Hearing status</th>
<th>CI</th>
<th>Estimate SE</th>
<th>0.27</th>
<th>0.16</th>
<th>0.47</th>
<th>0.50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syllable length</td>
<td>CI</td>
<td>Estimate SE</td>
<td>2.11 ***</td>
<td>-0.38</td>
<td>-4.54 ***</td>
<td>-2.29 *</td>
</tr>
<tr>
<td>Syllable length x Hearing status</td>
<td>CI</td>
<td>Estimate SE</td>
<td>0.16</td>
<td>0.7</td>
<td>0.15</td>
<td>0.51 *</td>
</tr>
</tbody>
</table>

* p≤0.05, ** p≤0.01, *** p ≤0.001
NH is the reference category

---

Figure 4. Development of the different response types of the Variability score with age for children with CI and children with NH – observed values

#### 4.4 Discussion

The present paper evaluates the intraword variability in Dutch-speaking children with CI longitudinally, and compared the variability of their target word productions with that of peers with NH. Spontaneous speech samples were analyzed up to age five using PWV (Proportion of Whole-word Variability) and Variability scores (4 response types: consistent correct, consistent incorrect, variable with hits, variable with no hits). Also the impact of the target word’s syllable length was examined as well as the effect...
of the age at implant activation in the longitudinal analysis. In what follows, we first discuss the intraword variability results in the CI group, and then the differences between both groups of children.

### 4.4.1 Intraword variability in children with CI

Firstly, the results point out that age has a significant effect on intraword variability in children with CI. PWV decreases from approximately 40% to approximately 20% between word onset and five years of age. Thus, intraword variability is relatively high in children with CI. Similarly, Moreno-Torres (2014) showed that the PWV of Spanish-speaking children with CI was 63% at a mean age of 2;09. Moreover, the relatively high percentages found in the present study are within the range previously reported in other studies on intraword variability in children with NH: PWV scores are between approximately 35% and 75% up to age four (Macrae, 2013; Sosa, 2015; Sosa & Stoel-Gammon, 2006; Sosa & Stoel-Gammon, 2012). Results also showed that intraword variability decreases with age in children with CI. In the literature, no longitudinal development of children with CI’s intraword variability is available. But, this decrease with age is in line with the findings of Holm et al. (2007) and Sosa (2015) for children with NH.

With respect to the Variability score, all four response types are found in Dutch-speaking children with CI from word onset up to the age of five: consistent correct, consistent incorrect, variable with hits and variable with no hits. In other studies on children with CI, only Moreno-Torres (2014) reported on the Variability score. He indicated that Spanish-speaking children with CI had mainly incorrect responses at a mean age of 2;09, namely variable with no hits and consistent incorrect. In contrast, our results point out that Dutch-speaking children with CI’s productions fall within all four response types throughout the entire period studied. Moreover, our children with CI’s profile for the Variability score is similar to that of children with NH as reported in the literature: all response types appear in those children as well (McLeod & Hewitt, 2008; Sosa, 2015).

Second, even though age affected PWV scores, the age at CI activation did not. This may be explained by the fact that in the present study the age range of CI activation was quite small: children with CI were all implanted before 20 months of
age and the difference between the child with the earliest CI activation and that with
the latest one is only 15 months. However, when considering the intraword variability
in more detail by means of the Variability score, some effects of CI activation were
found. Our results pointed out that children with later implant activation are more
likely to have variable responses (variable with hits and variable with no hits) and less
likely to have consistent responses (consistent correct and consistent incorrect) as
compared to peers with earlier implant activation. In other words, the age at implant
activation affects the likelihood of intraword variability in children with CI, which is
in line with other studies on the benefit of early implantation in children with CI (e.g.
Niparko et al., 2010). But our results should be considered with some care, as the
range of implant activation and the group sizes were relatively small.

Thirdly, the results of the effect of target word syllable length on intraword
variability suggest that intraword variability is higher in target words with more
syllables: PWV scores increase with increasing target word syllable length, i.e., longer
words show more variability in production than shorter words, but this effect is
limited in time. As children get older, the increase of PWV with increasing syllable
length becomes lower, suggesting that the effect disappears later on in development.
Moreover, the likelihood of the different response types of the Variability score is
affected by the target word’s syllable length. Consistent response types (consistent
correct and consistent incorrect) are less likely with increasing syllable length,
whereas variable response types (variable with hits and variable with no hits) are more
likely with increasing target word length. Thus, children with CI are less consistent
and more variable as the syllable length of the target word increases. There is also an
interaction with age, which suggests that children with CI are more variable when the
target word length in syllables increases. This applies even to older children.

The length of target words affects intraword variability of children with NH
(Macrae, 2013; Sosa, 2015). To the best of our knowledge, the only study which
reported on the effect of target word length on intraword variability in children with
CI is Moreno-Torres (2014). In that study, no effect of syllable length in children with
CI was found. It should be noted that syllable omissions from target words were not
considered as variable productions in Moreno-Torres (2014). This may, at least partly,
IW variability

explain the differences between the findings of his study and those of the present investigation. Moreover, our results corroborate the finding reported in other studies that phonological complexity has an effect on intraword variability. For instance, Ertmer and Goffman (2011) found that target words with late acquired word initial consonants (e.g. fricatives) are more variable in both children with NH and children with CI. The present study found that another factor of phonological complexity, i.e. a target word’s syllable length, has a similar impact on children with CI’s variability: variability increases with increasing word length. But Ertmer and Goffman (2011) studied only one point in children’s development (i.e. age 4;00). Hence, they were no able to track a possible decline of the effect of phonological complexity. The present study traced the development of intraword variability between word onset and age 5;00 and this reveals that the impact of word length decreases in children with CI as they get older.

Thus, the length of target words seems to hinder its correct and consistent production. It is also shown to affect children with CI’s production accuracy. For instance phonemic accuracy at the word level decreases with increasing target word syllable length in children with NH (Gathercole, Willis, Emslie & Baddeley, 1991) as well as in children with CI (Burkholder-Juhasz, Levi, Dillon & Pisoni, 2007; Faes et al., 2016; Nittroer, Caldwell-Tarr, Sansom, Twersky & Lowenstein, 2014; Von Mentzer, Lyxell, Sahlén, Dahlström, Lindgren, Ors, Kallioinen, Engström & Uhlén, 2015). Furthermore, target word syllable length affects suprasegmental accuracy in children with CI: production of primary stress is less accurate and the correct number of syllables decreases in longer target words (Carter, Dillon & Pisoni, 2002). Thus, the more syllables in the target word, the less accurately and the less consistently they are produced.

4.4.2 Cross-sectional comparisons of children with CI and age-matched children with NH

For the cross-sectional comparisons between both groups of children two factors were included in the analyses: the effect of hearing status (CI vs. NH) and the effect of a
target word’s syllable length, in order to determine whether this effect is similar in both groups of children.

Is intraword variability comparable in children with NH and children with CI? Results suggest that, initially, children with CI produce words less variable as compared to their peers with NH (PWV is lower). Between ages 3;00 and 5;00, PWV further decreases in both groups of children. At age 3;00, PWV of both groups of children is similar. But, at age 4;00 the PWV of children with CI is higher than that of children with NH, whereas, at age 5;00, the opposite is true.

Thus, PWV is lower in the CI group at age 2;00 and age 5;00. The lower rate of variability at age 2;00 may be found in the difference in target word length of both groups of children. An additional analysis revealed that children with CI target shorter words in comparison to their peers with NH up to age 2;0. Thus, their target words are less complex (see also Schauwers et al. (2008)). As our results indicate that longer target words are more variable in children with CI, it is possible that children with CI are less variable because they target shorter words. The fact that children with CI target shorter words may be due to their later onset of word productions. In children with CI, the first words appear later (mean 1;06) than in their peers with NH (mean = 1;02), i.a. due to their delayed onset of hearing. However, for instance Ferguson and Farwell (1975) suggested that the first word productions are relatively accurate, and thus also relatively consistent. After an initial period of relatively accurate productions, accuracy decreases and variability goes up, but eventually children’s productions become adult-like: i.e. more accurate and consistent (u-shaped learning, see e.g. Ferguson and Farwell (1975) and Fikkert (1994)). It is possible that children with CI are still in an earlier phase of u-shaped learning, whereas children with NH have already passed this phase. As a result, the word productions of children with CI are more consistent and more accurate and the PWV is lower in children with CI. This means that variability may also be linked to lexical development. A similar observation is made by for instance Stoel-Gammon (2011): phonological development and lexical development are commensurate. For children with NH, it has already been shown that intraword variability decreases with increasing lexicon size (Macrae, 2013; Sosa & Stoel-Gammon, 2012), but not for children with CI. In further studies, it
IW variability

would be crucial to examine the effect of lexicon size on intraword variability in children with CI as well. So far, the present paper has already shown that the effect of age is similar in both groups of children. As Faes and Gillis (2016b) have shown that phonological development and lexical development are also commensurate in children with CI, it is expected that vocabulary size also affects intraword variability in children with CI.

An explanation of the lower variability in children with CI at age 5;00 is less straightforward. Yet, it remains unclear which factors contribute to this observation. One factor may be the difference in data collection method: the CI speech samples were collected longitudinally, whereas the NH speech samples were a cross-sectional data set. Consequently, the data of the children with NH at age 5;00 are different from those at e.g. age 4;00. Hence, there is little information on the previous language skills of those five-year-olds with NH. It is not clear whether the overall lower consistency (higher variability) is characteristic for all five-year-olds or only the five-year-olds in the present sample. Next to the data collection, other factors might account for some of the inconsistency such as an increased effect of word frequency, neighborhood density, and/or phonotactic probability. It has been shown that words with dense neighborhoods facilitate lexical selection and processing (Baus, Costa & Carreiras, 2008). Moreover, it has been shown that variability decreases in words with a higher word frequency and a higher phonological neighborhood density (Sosa & Stoel-Gammon, 2012). It is possible that the children with NH at age 5;00 are targeting newly acquired words with less dense neighborhoods whereas children with CI are not. This makes lexical selection and especially speech processing harder for children with NH, resulting in less consistent production. However, it remains to be established how such factors can explain the difference between both groups of children at age 5;00.

The outcomes regarding the group differences with respect to PWV should be considered with some care. The decrease of the PWV score suggests that children become increasingly more consistent in their production of adult target words: the number of different replicas of a given target decreases steadily over time. But a PWV score of 0 does not mean that children produce a target word correctly. It simply
shows that children are consistent in their production. In other words, children with CI at age 5;00 produce target words with less variability than peers with NH, but it may very well be possible that target words at this age are produced consistently incorrectly. Similarly, the more consistent production of children with CI at age 2;0 might conceal consistent incorrect production. The PWV measure provides an indication of children’s consistency in rendering target words, but does not provide any information on the accuracy of production. For this purpose a second measure of intraword variability was used: the Variability score. At age 2;00, the likelihood of consistent correct and variable with no hits responses is similar in children with NH and children with CI, but children with CI are less likely to have variable with hits responses and more likely to have consistent incorrect responses. In other words, children with CI are less likely to hit the target as compared to their peers with NH at age 2;00. Similar effects are found at ages 3;00 and 4;00: children with CI are less likely to hit the target and less likely to be consistent in comparison to peers with NH. However, children with CI catch up on their peers with NH by age 5;00. At this age, children with CI even seem to outperform peers with NH regarding correct productions: they have more consistent correct responses, whereas the children with NH have more variable with hits responses. Thus, children with CI are less correct and more variable as compared to their peers with NH up to age 4;0. By age 5;00, however, children with CI even outperform their peers with NH on consistency and are similar to their peers with NH on correctness. Similarly to the PWV results, an explanation for the children with CI’s high consistency and correctness at age 5;00 might be found in the data collection, but also in factors related to the target words of both groups of children, such as word frequency, neighborhood density, etc.

Variability increases when target words have more syllables. PWV scores are higher in longer target words and the effect is similar in both groups of children up to age 3;00. At age 4;00, the effect already disappeared in children with NH, whereas this was only the case at age 5;00 in children with CI. When considering the Variability score, the results suggest that the likelihood of consistency and correctness decreases when target words have more syllables. Longer target words seem to be both less consistent and less correct. However, at age 2;00, these effects are less
pronounced in children with CI. As for the results on PWV, this may be due to children with CI targeting shorter words at age 2;00. From age 3;00 onwards, the effect of a target word’s syllable length on the likelihood of the different response types is similar in both groups of children, except for consistent correct responses at age 4;00. Those responses are less likely in longer target words in children with CI, but not in children with NH.

4.4.3 Explanation of the results, clinical implications and future research

To summarize, intraword variability is high in children with CI, but decreases with age. Moreover, there are some group differences up to age 4;00, but intraword variability is similar in both groups of children at age 5;00. Furthermore, longer target words are more likely to be produced variable in both groups of children and to be produced incorrectly. A possible explanation can be found in the role of input in language development. Speech perception is indispensable for speech production. This also includes auditory feedback of children’s own productions, as perception of their own productions is essential to fine-tune articulation (Stoel-Gammon, 2011; Stoel-Gammon & Sosa, 2007). In children with CI, this auditory feedback starts later as a result of a period of auditory deprivation before implantation. In addition, after implantation the speech signal is still degraded and more noisy than that available in normal hearing (Drennan & Rubinstein, 2008). However, this may hinder fine-tuning of articulation and may therefore add to production inconsistency in children with CI, as hypothesized by Moreno-Torres (2014). Moreover, children with CI are found to be less attentive to the ambient language (Houston & Bergeson, 2014; Houston, Pisoni, Kirk, Ying & Miyamoto, 2003) and to have difficulties in processing speech and storage of elements in memory tasks (Kronenberger, Pisoni, Henning & Colson, 2013; Nitttrouer, Caldwell-Tarr & Lowenstein, 2013; Pisoni, Kronenberger, Roman & Geers, 2010). In other words, children with CI are more likely to miss parts of the speech signal. First of all, the speech signal is degraded in children with CI as compared to children with NH. This may disturb the input and auditory feedback in comparison to the speech signal available in normally hearing children. Secondly, children with CI seem to be less attentive to speech in the ambient language, which adds to the likelihood of children with CI missing some parts of the input. Thirdly,
even if perception were similar in both groups of children, children with CI seem to have difficulties in processing and storing the input as well as processing their own output. These factors may influence the phonological representations of target words in children with CI. If target words are perceived less detailed (input and auditory feedback), they receive less attention and are processed and stored less efficiently, it is likely that the phonological representation of those target words is less fine-grained. This may in turn explain why children with CI seem more variable and less accurate than their peers with NH: if target words have less fine-grained representations and those representations are fine-tuned less effectively due to reduced auditory feedback and speech processing skills, its production may become more variable. In addition, processing of more complex words such as target words with more syllables, is more difficult (Faes et al., 2016; Gathercole et al., 1991). That may explain why longer target words are produced more variable. Thus, more complex words are produced more variably and it is possible that their representations are less fine-grained initially.

In a similar vein, Von Mentzer et al. (2015) hypothesized that children with CI have less fine-grained representations of words with consonant clusters.

To conclude, our results show that intraword variability (a) is high in Dutch-speaking children with NH and children with CI but decreases with age, and (b) is higher in target words with more syllables. These results have important clinical implications and raise questions for further research. First of all, intraword variability is higher in children with CI, but they catch up by age five. This result is promising and underlines the positive effect of cochlear implantation. Secondly, results show that the development of intraword variability is similar in both groups of children: it decreases with age and increases in longer target words. So far, only the effect of age, the effect of age at implantation and the effect of target word syllable length on the intraword variability of children with CI have been studied. For children with NH, also other factors are reported to affect intraword variability, such as target word frequency, the phonological neighborhood density of target words, the vocabulary size of the children, etc. In future studies, it is essential to investigate the effect of those factors on intraword variability in children with CI. As such, insight in the effect of those factors on the variability of children with CI can adjust the speech and language
IW variability

therapy in children with CI and can reveal a better understanding of the role of input and auditory feedback in language development.
Chapter 5
Chapter 5

5 The effect of word frequency on phonemic accuracy in children with cochlear implants and peers with normal hearing

ABSTRACT
Word frequency influences the language acquisition of typically developing children. For instance, highly frequent words appear earliest in a child’s lexicon (Goodman, Dale & Li, 2008), and highly frequent phonemes are more accurately produced in child language (Edwards, Beckman & Munson, 2004). This study evaluates (1) whether word frequency influences word accuracy and (2) whether this is also the case for children with a history of auditory deprivation. Thus, the influence of word frequency on the phonemic accuracy of that word in children with a cochlear implant (CI) is examined and compared to children with NH, between word onset and age seven. Results show that highly frequent words are more accurate, except words with the highest frequency, predominantly closed-class words. This effect is more pronounced in children with NH as compared to children with CI. Thus, children with CI are sensitive to word frequency, but to a lesser extent than peers with NH.

Keywords: Phonemic accuracy; children with CI; Levenshtein distance; target word frequency

5.1 Introduction
Examples of frequency of occurrence effects in the language acquisition of children with normal hearing (NH) appear in all domains of language, e.g. the acquisition of inflectional morphology, the acquisition of syntax, the acquisition of single words, etc. (Ambridge, Kidd, Rowland & Theakston, 2015). These frequency effects are found in the incidence of phonemes, sequences of phonemes, words, sequences of words, etc. and all go in the same direction: phonemes, sequences of phonemes, words, sequences of words that occur frequently are acquired earlier (Ambridge et al., 2015). Less is known about such frequency effects in children with severe to profound hearing loss with partly restored hearing due to cochlear implantation (CI). If frequency effects in
children with hearing loss differ from those in children with NH, there are two possibilities: (a), it is possible that frequency effects are more pronounced in children with cochlear implants (CI) as compared to children with NH, but, (b), it is also possible that frequency effects are less pronounced children with CI as compared to children with NH. At a similar chronological age, children with CI have less language experience in comparison to their peers with NH, as they start hearing later. Moreover, their speech perception is affected by the degraded signal they receive from the implant (Drennan & Rubinstein, 2008; Wilson, 2006). For instance, it has been shown that the current devices are yet unable to accurately pass sounds at high spectrum frequencies, such as fricatives (Stelmachowicz, Pittman, Hoover, Lewis & Moeller, 2004). Due to the degraded auditory information, children with CI may either be more sensitive to frequency information (incidence of phonemes, sequences of phonemes, words) in the input (possibility a), or be less attentive to the ambient language in comparison to peers with NH (possibility b). In the literature, there is evidence for this last possibility. In a habituation phase task, for instance Houston, Pisoni, Kirk, Ying and Miyamoto (2003) found that children with NH, but not children with CI, differentiated significantly between speech sound trials and silent trials. Thus, children with CI’s attention to speech sounds is lower in comparison to peers with NH. Therefore, children with CI may well be missing some aspects of the speech directed to them, which may affect their sensitiveness to language statistics such as word frequency (Saffran, Aslin & Newport, 1996). The current study examines the influence of word form frequency in child-directed speech on phonemic accuracy in Dutch-speaking children with CI longitudinally in comparison with peers with NH. Throughout the entire manuscript, the term frequency denotes the incidence of words, unless specified else.

The effect of frequency (i.e. word frequency) on the phonemic accuracy has only been studied rarely. To the best of our knowledge, the effect of word frequency on word accuracy has not yet been documented in children with CI, but there is research on children with NH. In two-year old English-speaking children with no hearing problems, there seems to be no relationship between frequency in the input and accuracy: Sosa and Stoel-Gammon (2012) have shown that highly frequent words
productions are less variable but not more accurate. In contrast, Ota (2006) found that Japanese two-year-olds with NH are less likely to truncate words with a high frequency in the input. In other words, they are less likely to omit syllables in highly frequent words and hence are more accurate. Two questions manifest themselves: (1) Does the word frequency in the input affect word accuracy in children with NH? And (2) are children with CI similar to children with NH in this respect? In other words, do they exhibit a similar effect to the same extent as children with NH? In what follows, we first discuss frequency effects in children with NH and, afterwards, discuss the relevant literature on children with CI.

5.1.1 Frequency effects in children with NH

The effect of word frequency on word accuracy in children with NH is still a matter of debate: Sosa and Stoel-Gammon (2012) found no effect, whereas Ota (2006)’s results pointed in the opposite direction. However, the literature on children with NH reports on other frequency effects that may very well be related to word frequency and word accuracy. First, frequency at the segmental level plays a role in the accurate production and the age of acquisition of segments and sequences of segments. With respect to segmental sequences, frequency is commonly operationalized as phonotactic probability. Phonotactic probability represents the odds that a sequence of segments occurs in the target language. Children with NH produce segmental sequences with higher phonotactic probability more accurately than segmental sequences with lower phonotactic probability (for English: Edwards, Beckman & Munson, 2004; Munson, 2001). Moreover, segments that occur more often in the input appear earlier in children’s output (Lee, Davis & MacNeilage, 2010). With respect to individual segments, Stokes and Surendran (2005) have shown that the input frequency of initial consonants accounts for approximately 40% of the variation in consonant production accuracy in English-speaking and Dutch-speaking children. Moreover, Van Severen, Gillis, Molemans, Van den Berg, De Maeyer and Gillis (2013) have shown that consonants appear earlier in Dutch-speaking children’s speech if the input frequency is higher. Thus, frequency at the segmental level (PP and individual segmental frequency) affects accuracy at the segmental level and the age of acquisition of segments. Thus, children produce individual segments and sequences of
Word frequency effects

segments with a higher frequency in the input more accurately. It is possible that a similar trend appears at the word level. That would be in line with the findings of Ota (2006), but in contrast to those of Sosa and Stoel-Gammon (2012).

Secondly, there is a relation between word frequency and lexical development: frequency at the word level influences the age of acquisition of words, which is similar to the segmental frequency and segmental age of acquisition relation. Children with NH acquire highly frequent words earlier, except for words with the highest frequencies such as closed-class words. For English, Naigles and Hoff-Ginsberg (1998) have shown that higher input frequency of verbs positively influences children’s use of these verbs 10 weeks later. A similar effect of frequency on acquisition was found in other languages than English as well, e.g., in Hebrew (Ashkenazi, Ravid & Gillis, 2016). In contrast, Storkel (2009) did not find an effect of word frequency on noun acquisition in English-speaking children with NH. Goodman, Dale and Li (2008, English), Stokes (2010, English) and Stokes, Kern and Dos Santos (2012, French) have also reported that early acquired words have a relatively low frequency and, thus, that highly frequent words are missing from early vocabularies. However, highly frequent words are often closed-class words like in and some in English (Goodman et al., 2008). Such closed-class words are highly frequent in adult-directed speech (ADS) as well as in child-directed speech (CDS). However, these closed-class words are commonly absent from early vocabularies (Goodman et al., 2008; Stokes, 2010; Stokes et al., 2012). An analysis of individual lexical categories (nouns, verbs, closed-class words, etc.) reveals that early acquired words are indeed highly frequent in all categories (Goodman et al., 2008; Naigles & Hoff-Ginsberg, 1998). In other words, lexical development in children with NH is related to the word frequency in the input.

An important difference in studies on frequency effects is the source of the frequency counts (Hoff-Ginsberg, 1992). Some authors based their frequency counts on corpora of adult directed speech (ADS) (Stokes, 2010; Stokes et al., 2012), whereas Naigles and Hoff-Ginsberg (1998) have used a corpus of CDS, and Sosa and Stoel-Gammon (2012) and Storkel (2009) have based their frequency counts on the adult written language corpus of Kucera and Francis (1967). Goodman et al. (2008) have
Chapter 5

compared two corpora: The corpus of adult-directed written language of Kucera and Francis (1967) and a corpus of CDS. The use of ADS corpora may influence the effect of word frequency tremendously, as CDS differs from ADS (Bernstein Ratner, 2013; Newport, Gleitman & Gleitman, 1977). For instance, CDS has a limited vocabulary, which can affect frequency of particular words. Moreover, it is suggested that children are sensitive to CDS but not to ADS: children’s vocabulary development is only correlated with CDS and not with ADS (Weisleder & Fernald, 2013). In addition, spoken language differs from written language (Akinnaso, 1982), which may also impact the difference in outcomes. Indeed, studies which used adult written corpora (Sosa & Stoel-Gammon, 2012; Storkel, 2009) and corpora of ADS (Stokes, 2010; Stokes et al., 2012) did not find any word frequency effects. Moreover, Goodman et al. (2008) and Van Severen et al. (2013) have shown that the effect of frequency is best captured when using a corpus of CDS as opposed to an adult-directed written one. This is also in line with Stoel-Gammon (2011), who has proposed that word frequency effects should be analysed using “a variety of measures including […] [i.a.] general counts of input to children based on corpora from many children”. Furthermore, Stoel-Gammon (2011) has also proposed that frequency effects should be based on “general counts of child output based on corpora from many children”. The present paper investigates the effect of word frequency on accuracy.

5.1.2 Frequency effects in children with CI

In this section we examine whether there is an effect of whole-word frequency on whole-word accuracy in children with CI. In children with NH, highly frequent words appear earlier in children’s vocabulary (lexical acquisition) and highly frequent segments are more accurately produced. For lexical development, Han, Storkel, Lee and Yoshinaga-Itano (2015) have reported a similar frequency effect in children with CI: Early vocabularies comprise highly frequent words. This frequency effect, measured in a corpus of kindergartener’s own productions, decreases between four and seven years of age (Han et al., 2015). In contrast, Guo, McGregor and Spencer (2015) have shown that there is no correlation between word frequency, measured in CDS, and lexical development of monosyllabic nouns, verbs and adjectives one year post-implantation in children with CI, nor in children with NH.
Word frequency effects

With respect to frequency at the segmental level, Guo et al. (2015) have considered the effect of phonotactic probability (PP) in children with CI one year post-implantation. They indicated that the effect of PP on accuracy was similar in children with NH and children with CI wearing bilateral implants, but not in children with NH and children with CI wearing unilateral implants. This suggests that children with CI with unilateral implants are less sensitive to statistical effects at the word and the segmental level in the ambient language (Guo et al., 2015). The current paper investigates a longitudinal sample of children with CI’s speech. The effect of word frequency on word accuracy in children with CI is compared to that in children with NH. The children with CI in this study are a mixed population: all children used only one CI for a long period, but more than half of the children received a second implant when they were older (see method section).

To conclude, it is important to note that the CDS corpus in Goodman et al. (2008), Guo et al. (2015) and Naigles and Hoff-Ginsberg (1998) were not based on CDS addressed to the study’s participants, but on other corpora of CDS. Even though Guo et al. (2015) have recommended to do so, so far, only Van Severen et al. (2013) have adopted this suggestion in their research. In line with this, the present paper computes word frequencies on CDS to the study’s participants. Here the focus is not on lexical acquisition but on accuracy development. The effect of phonotactic probability on phonemic accuracy is similar to the effect of word frequency on lexical acquisition: Highly frequent phoneme sequences are more accurately produced and appear earlier in children’s speech, as highly frequent words are earlier acquired. Therefore, highly frequent words in CDS are expected to be more accurate in both children with CI and children with NH. In line with Guo et al. (2015), who showed that children with CI with unilateral implants are less sensitive to statistics in language, a less pronounced effect in children with CI in the present study is assumed, as those children had unilateral stimulation for a long period.
5.2 Methods

5.2.1 Participants
A longitudinal design was set up to assess language development of children with CI: 9 children with CI were followed from word onset up to the age of five (Schauwers, 2006). The children were followed monthly from the moment of implant activation up to 2;6 years after activation. Additionally, yearly data were collected between ages 3;0 and 5;0. Here, the data will be analyzed from the appearance of first spoken words (median = 1;6, range 1;3 – 1;11). In chapter 2, more detailed information on the children with CI is given.

The control group consisted of a selection of the NHlongitudinal corpus at 24 months of age (S11 – S21, i.e. the alphabetical first 11 children) and the NHcross-sectional corpus between ages 3;0 and 7;0 (S51 – S103). More information can be found in chapter 2.

5.2.2 Procedures
The data used in the current study were gathered as a part of a larger corpus, namely the CLiPS Child Language Corpus (CCLC). Only speech samples in which lexical items appeared were analyzed. Detailed transcriptions of the data collection, data annotation and the reliability of the transcriptions are provided in the method section (chapter 2).

5.2.3 Data analyses
Phonemic accuracy was measured in a dynamic cost model of Levenshtein distance (LD) as presented in Wieling, Prokic and Nerbonne (2009) and adapted by Faes, Gillis and Gillis (2016) to fit the purpose of computing phonemic accuracy in child speech (see also chapter 3). LD computes the distance between two character strings (e.g. DNA sequences and dialectal variants, but also phonemic transcription of a child’s rendition and the adult target) and is more fine-grained than other measures such as the phonological mean length of utterance (pMLU, Ingram, 2002) (for a discussion, see Faes et al., 2016). First, LD only penalises inaccurately produced phonemes and is, therefore, ideally zero. Thus the lower the LD, the higher the phonemic accuracy.
Word frequency effects

Second, in order to calculate this distance, three edit operations are assumed in transforming one string into the other: deletions, substitutions and insertions (Wieling, Margaretha & Nerbonne, 2012). In pMLU, insertions are not penalised. If no edit operations are needed to transform the child’s rendition into the adult target, i.e. if the child’s rendition is identical to the adult target, the LD equals 0, which represents complete accuracy. Third, there is a weighting of speech errors in LD, which is not found in other measures. In LD, frequent, less severe speech errors receive a smaller cost than less frequent, more severe speech errors. The LD between a child’s production and the adult equivalent is calculated in two phases: (1) an alignment between the actual adult production and the target, and (2) an alignment between the children’s speech and the target.

Thus, in the first phase, the actual adult speech of a word is aligned with the target VC-sensitive. This means that a vowel cannot be aligned with a consonant, and vice versa. Then, the pointwise mutual information (PMI) of each target – rendition pair is calculated using the formula in Eq. (1):

$$PMI(x, y) = \log_2 \frac{p(x,y)}{p(x)p(y)}$$

(1)

where $p(x, y)$ is the probability of encountering the pair $(x, y)$ in the alignment, $p(x)$ the probability of segment /x/ and $p(y)$ the probability of segment /y/. The PMI of each pair of segments is calculated for the entire transcription. Pairs of segments that co-occur frequently, receive a high PMI value. In order to convert the PMI values into a cost, the PMI value of each pair of segments is subtracted from the maximum PMI value. As a result, pairs of segments that frequently co-occur receive a smaller cost (or distance) than segments that rarely co-occur. When this final cost model is computed, the Levenshtein distance (LD) between an adult production and the target is the sum of the distances between the individual segments. Since word length affects LD, the final LD score per word is normalized by dividing it by the word length (in phonemes). As such, an average LD per word is calculated.

In the second phase, the children’s speech is aligned with its target. The LD is the sum of the distances calculated in the first phase. If this alignment had appeared
earlier in the calculations, errors that are highly frequent in child speech would receive a high PMI value and therefore a low LD cost. Therefore, the costs are calculated on a model of adult speech constructed in the first phase, and the eventual distance between the child’s rendition and the adult target is calculated in this second phase.

Before statistical data analysis, outliers were identified by the interquartile rule and excluded. Next, substandard forms (e.g. /ika/ for /ik/) were excluded as well. Whole-word frequency was operationalized as follows: the entire token frequency each word in the CDS addressed to the children with NH and children with CI was tallied. No differentiation was applied to the children with NH and children with CI, as for instance Bergeson, Miller and McCune (2006), Tribushinina, Gillis and De Maeyer (2013) and Vanormelingen (2016) showed that CDS is similar in children with CI and peers with NH for prosodic aspects, the incidence of different words classes, and the amount of input plus responsiveness respectively. The absolute frequency counts were log transformed. Next, also the token frequencies in the children’s own speech were tallied and consecutively log transformed. Correlation analyses in R showed that both measures of frequency are highly correlated ($r = 0.79$). Therefore we opted to only include one frequency measure into the analyses. The level of variance explained by both variables in the longitudinal part of the data, which includes the largest amount of data, was compared. Thus, we ran a statistical model in which accuracy was predicted by log frequency in the input and another statistical model in which accuracy was predicted by log frequency in the output. Analyses showed that the statistical model including log frequency in the input explained more variance than the model including log frequency in children’s own speech. Therefore, we chose to analyse the effect of word frequency in the input on word accuracy.

5.2.4 Statistical analyses
Statistical analyses were carried out in R (R Core Team, 2013) by means of multilevel modelling. A multilevel model consists of two parts: A fixed part and a random one. In the fixed part of the model, the independent variables are included. The random part considers the nesting of variables and variation in the data (Baayen, 2008). The data were structured hierarchically: Different words are nested within different children (at different ages). This variation between children and ages was considered
Word frequency effects

in the random part of the model. The analyses were divided into two parts: (1) A longitudinal analysis of children with CI and (2) cross-sectional comparisons of children with CI and children with NH.

The first part of the analyses included the longitudinal data of all children with CI, from word onset (median = 18 months, range 15 – 23) up to the age of seven. A total of 49,652 word tokens were examined. For the longitudinal analysis of children with CI, the random effects of the multilevel model consisted of random intercepts and slopes for each child at each age. The fixed effects of the multilevel model were age (Age, intercept at age 20 months), the age at implant activation (CIactivation) and word frequency (Frequency). Quadratic (x²) and cubic (x³) effects for Frequency and Age as well as the interaction between all fixed effects were included if that yielded a better model.

In the second part of the analyses, cross-sectional analyses were performed between the age of two and seven. As no longitudinal data of children with NH were available, it would be incorrect to include all children with NH in one model. Therefore, the data of the CI corpus were split and the analyses were performed for each age separately: At age two (range 23 – 25 months), three (range 34 – 40 months), four (range 45 - 51 months), five (range 59 - 63 months), six (range 67 – 77 months) and seven (range 82 – 86 months). A total of 67,321 word tokens were examined (children with CI: 34,707; children with NH: 32,614). The distribution of the word tokens per age was as follows: 7,628 word tokens at age two (children with CI: 2,752; children with NH: 4,876), 14,807 word tokens at age three (children with CI: 12,604; children with NH: 2,203), 12,637 word tokens at age four (children with CI: 6,309; children with NH: 6,328), 11,099 word tokens at age five (children with CI: 5,834; children with NH: 5,265), 9,587 word tokens at age six (children with CI: 2,698; children with NH: 6,889) and 11,563 word tokens at age seven (children with CI: 4,510; children with NH: 7,053). For each cross-sectional analysis, the random part of the multilevel model included random intercepts for each child. The fixed effects were hearing status (Hearing status) and word frequency (Frequency). Quadratic (x2) and cubic (x³) effects for Frequency and the interaction between Hearing status and
Frequency were included if that yielded a better model. Only the best fitting models will be reported and a significance level of p<0.05 was set.

5.3 Results

5.3.1 Longitudinal analysis of children with CI

In this section, the longitudinal analysis of Levenshtein distance (LD) in children with CI is described. The effects of age, age at implant activation and frequency were included. The fixed effect results are displayed in Table 1.

In Figure 1, the development of LD with age is presented. It is clear from Figure 1 that the LD decreases with age. In other words, the phonemic accuracy of children with CI increases as they get older. The decrease of LD is significant, as can be derived from the effect of Age (p<0.001) in Table 1. In addition, Figure 1 shows that the decrease of LD flattens out from approximately 50 months of age. This effect is also significant, as indicated by a significant quadratic effect of Age (Age², p<0.001) in Table 1. Put differently, phonemic accuracy increases as children get older, but this increase flattens out from 50 months of age onwards.

Figure 2 plots the effect of word frequency on LD for different age ranges. Regardless the age range, the global effect of word frequency seems to follow a cubic trend and more precisely an inverted s-curve. LD first increases with increasing word frequency, then decreases when word frequency further increases, but eventually increases again for words that occur most frequently (Figure 2). In other words, accuracy first decreases, then increases and finally decreases (s-curve). This developmental trend of LD with word frequency is significant, as shown by significant effects of Frequency (p<0.01), Frequency² (p<0.001) and Frequency³ (p<0.001) in Table 1. In addition, Figure 2 indicates that the global word frequency effect is dependent on age. Up to the age of four, Figure 2 shows that more frequent words have a lower LD and are more accurate. However, there is a discontinuity. Figure 2 also points out that from the age of four onwards, LD seems to increase, and thus accuracy seems to decrease, with increasing word frequency. Moreover, this developmental trend of LD is more pronounced as children grown older. Thus, word frequency affects LD, and thus phonemic accuracy, differently at different ages and
Word frequency effects

this is significant as shown by a significant interaction between Frequency and Age (p<0.001) in Table 1.

With respect to age at implantation, Table 1 shows a significant effect of CIactivation (p<0.01), meaning that children with CI with later implant activation have higher LD values. However, the significant interaction between CIactivation and Age (p<0.01) suggests that the effect of CIactivation becomes less pronounced when children get older. Thus, initially, the accuracy of children with CI with later implant activation is lower as compared to children with earlier implant activation. But, as children get older, the difference between children with CI with earlier and later implant activation becomes smaller. In other words, children with CI with later implant activation are catching up on their earlier implanted peers at the older ages. Table 1 finally shows a significant interaction between CIactivation and Frequency (p<0.001), meaning that the effect of frequency on LD (inverted s-curve) is less pronounced in children with CI with later implant activation as compared to those with earlier implant activation. As the interaction between CIactivation, Frequency and Age was not significant and did not improve the model fit, it was not included in the best fitting model reported here. However, the lack of significance of this three-way interaction suggests that the less pronounced frequency effect on LD in children with CI with later implant activation remains stable with age. In other words, children with CI with later implant activation are not catching up on their earlier implants peers: Throughout development, the effect of word frequency on LD, and thus word accuracy, is smaller in children with CI with later implant activation.

Table 1. Fixed effect results of the longitudinal analysis of children with CI

<table>
<thead>
<tr>
<th>Estimate (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
</tr>
<tr>
<td>18.4022 (0.8753) ***</td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td>-0.1173 (0.0184) ***</td>
</tr>
<tr>
<td>Age^2</td>
</tr>
<tr>
<td>0.0012 (0.0001) ***</td>
</tr>
<tr>
<td>Frequency</td>
</tr>
<tr>
<td>0.0854 (0.0316) **</td>
</tr>
<tr>
<td>Frequency^2</td>
</tr>
<tr>
<td>-0.0753 (0.0066) ***</td>
</tr>
<tr>
<td>Frequency^3</td>
</tr>
<tr>
<td>0.0051 (0.0004) ***</td>
</tr>
<tr>
<td>Frequency x Age</td>
</tr>
<tr>
<td>0.0059 (0.0002) ***</td>
</tr>
<tr>
<td>CIactivation</td>
</tr>
<tr>
<td>0.1598 (0.0589) **</td>
</tr>
</tbody>
</table>
CI activation x Age: $-0.0036 (0.0011)$ **
Frequency x CI activation: $-0.0044 (0.0008)$ ***

$p ≤ 0.05^*$, $p ≤ 0.01^{**}$, $p ≤ 0.001^{***}$

Figure 1. The development of Levenshtein distance (LD) in children with CI as a function of age – fitted values.

Figure 2. The development of Levenshtein distance (LD) in children with CI as a function of word frequency and age – fitted values.
5.3.2 Cross-sectional comparisons between children with CI and children with NH

Table 2 shows the fixed effect results of the cross-sectional analyses between the ages of two and seven and Figure 3 plots the effect of frequency on the LD development per age. Firstly, there are some effects of Hearing Status on LD. Table 2 and Figure 3 show that LD is significantly higher in children with CI as compared to peers with NH ($p<0.05$) up to the age of four and that LD is similar in both groups of children from the age of five onwards ($p>0.05$). In other words, word accuracy is higher in children with NH up to the age of four, but similar in both groups of children from the age of five onwards.

Secondly, Table 2 shows significant effects of Frequency on LD and significant interactions between Frequency and Hearing Status. This means that frequency has a significant effect on LD and that this effect differs to some extent in both groups of children. The precise development of LD with frequency in both groups of children at each age is visualised in Figure 3.

At the age of two and three, Table 2 points out significant cubic effects of Frequency ($p<0.001$) and significant interactions between Hearing status and Frequency ($p<0.001$). Figure 3 shows that these linear, quadratic and cubic effects follow an inverted s-curve at the age of two and three in children with NH. LD increases with increased frequency, then LD decreases, but for words in the highest frequency regions LD increases again. In other words, accuracy decreases, then increases, but decreases again in the most frequent words (s-curve). With respect to the group differences (interaction Hearing status and Frequency), Figure 3 indicates that the developmental trend is similar in children with NH and children with CI at the age of two, but that the effect is more pronounced in children with NH. However, at the age of three, Figure 3 further suggests that the effect of Frequency is different in children with CI: There seems to be a slight increase of LD initially, but then there is an overall decrease of LD with increasing frequency.

From the age of four onwards, there are significant quadratic effects of Frequency ($p<0.001$, Table 2). As Figure 3 shows the effect of frequency is similar in both groups of children, which can be inferred from the uniform curves: First LD decreases
with increasing frequency, but for the most frequent words, LD increases. Put differently, accuracy increases with increased frequency, but decreases for words with the highest frequencies (inverted u-shape). Thus, the developmental trends of Frequency on LD are similar in both groups of children from the age of four onwards (Figure 3). Nevertheless, the effect of Frequency on LD is significantly more pronounced in children with NH as compared to children with CI (see interaction effects in Table 2). In other words, the changes of LD as a function of frequency are more pronounced in children with NH than in children with CI. This is true for all comparisons, thus from the age of four onwards, but also at the age of two and three.

Table 2. Fixed effect results of the cross-sectional analyses between ages 2;00 and 7;00

<table>
<thead>
<tr>
<th></th>
<th>2;00</th>
<th>3;00</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate (SE)</td>
<td>Estimate (SE)</td>
</tr>
<tr>
<td>(Intercept)</td>
<td>19.19 (0.31)**</td>
<td>17.93 (0.19)**</td>
</tr>
<tr>
<td>Hearing status NH</td>
<td>-0.83 (0.37)*</td>
<td>-2.39 (0.28)**</td>
</tr>
<tr>
<td>Frequency</td>
<td>0.54 (0.11)**</td>
<td>0.27 (0.06)**</td>
</tr>
<tr>
<td>Frequency²</td>
<td>-0.20 (0.02)**</td>
<td>-0.10 (0.01)**</td>
</tr>
<tr>
<td>Frequency³</td>
<td>0.01 (&lt;0.01)**</td>
<td>0.01 (&lt;0.01)**</td>
</tr>
<tr>
<td>Frequency x Hearing status NH</td>
<td>0.10 (0.02)**</td>
<td>0.22 (0.02)**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>4;00</th>
<th>5;00</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate (SE)</td>
<td>Estimate (SE)</td>
</tr>
<tr>
<td>(Intercept)</td>
<td>17.46 (0.14)**</td>
<td>15.87 (0.12)**</td>
</tr>
<tr>
<td>Hearing status NH</td>
<td>-1.21 (0.18)**</td>
<td>-0.09 (0.16)</td>
</tr>
<tr>
<td>Frequency</td>
<td>-0.27 (0.02)**</td>
<td>-0.16 (0.02)**</td>
</tr>
<tr>
<td>Frequency²</td>
<td>0.02 (&lt;0.01)**</td>
<td>0.02 (&lt;0.01)**</td>
</tr>
<tr>
<td>Frequency³</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency x Hearing status NH</td>
<td>0.08 (0.01)**</td>
<td>0.07 (0.01)**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>6;00</th>
<th>7;00</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate (SE)</td>
<td>Estimate (SE)</td>
</tr>
<tr>
<td>(Intercept)</td>
<td>15.97 (0.08)**</td>
<td>15.83 (0.08)**</td>
</tr>
<tr>
<td>Hearing status NH</td>
<td>-0.02 (0.09)</td>
<td>0.05 (0.09)</td>
</tr>
<tr>
<td>Frequency</td>
<td>-0.27 (0.02)**</td>
<td>-0.22 (0.02)**</td>
</tr>
<tr>
<td>Frequency²</td>
<td>0.03 (&lt;0.01)**</td>
<td>0.02 (&lt;0.01)**</td>
</tr>
<tr>
<td>Frequency³</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency x Hearing status NH</td>
<td>0.04 (0.01)**</td>
<td>0.02 (0.01)**</td>
</tr>
</tbody>
</table>

*p ≤ 0.05*,  p ≤ 0.01**,  p ≤ 0.001***

Shaded cells indicate that including that a particular variable did not significantly improve the model fit and is therefore left out.
Word frequency effects

Figure 3. The development of Levenshtein distance (LD) in children with CI and children with typical levels of hearing (NH) as a function of word frequency – cross-sectional analyses – fitted values

5.3.3 Relationship between frequency and the type of words

In the previous sections, it was shown that accuracy increases with increasing frequency, except for the most frequent words. It remains to be examined what types of words are represented in these highest frequency regions. We identified two word types in the entire dataset: function words and content words, as defined in Bussmann (1996). Nouns, lexical verbs, adjectives and adverbs were labelled as content words, and for instance pronouns, conjunctions, auxiliary verbs, articles were labelled as function words. Table 3 displays the results of an additional multilevel analysis considering this question. Log frequency in CDS was included as dependent variable, word type as independent variable and the word utterance itself as random effect. The intercept represents the log frequency of content words, which is 1.8581 (SE=
The log frequency of function words is significantly higher than content words and equals 2.8639 (= 1.8581 + 1.0058), which is significantly higher (p<0.0001). Thus, the most frequent words are function words, whereas less frequent words are content words.

Table 3. Fixed effect results of the relationship between log frequency (in the input) and word type

<table>
<thead>
<tr>
<th>Estimate (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
</tr>
<tr>
<td>1.8581 (0.0311) ***</td>
</tr>
<tr>
<td>Function words</td>
</tr>
<tr>
<td>1.0058 (0.0804) ***</td>
</tr>
</tbody>
</table>

\[ p \leq 0.05*, \ p \leq 0.01**, \ p \leq 0.001*** \]

5.4 Discussion

The present paper set out to evaluate the effect of word frequency on word accuracy in the spontaneous speech of Dutch-speaking children with NH and children with CI. In contrast to the literature (Goodman et al., 2008; Guo et al., 2015; Han et al., 2015; Naigles & Hoff-Ginsberg, 1998; Stokes, 2010; Stokes et al., 2012; Storkel, 2009), the frequency counts in this study were based on CDS directed to the study participants. This resulted in a direct image of the effect of word frequency on children’s accuracy. It was investigated whether (1) word frequency affects word accuracy in children with NH, (2) children with CI also exhibit an increase of accuracy with increasing word frequency, and (3) word frequency affects accuracy in children with CI and peers with NH to the same extent?

5.4.1 Word frequency and word accuracy in children with NH

In children with NH accuracy increases with increasing frequency, except for words with the highest frequencies, which are function words. Word frequency affects the accuracy of Dutch-speaking children with NH’s spontaneous speech productions. This observation agrees well with two frequency effects reported in children with NH. (1) Frequency at the segmental level affects accuracy and age of acquisition at the segmental level. Segmental sequences with higher phonotactic probability (PP) are produced more accurately than segmental sequences with lower PP (Edwards et al., 2004; Munson, 2001) and more frequent segments appear earliest in children’s
Word frequency effects

productions (Van Severen et al., 2013). Even though our results consider word frequency and word accuracy, the effect is similar: more frequent words in the input are more accurate. (2) Our results are in line with the effect of word frequency on lexical acquisition in children with NH. Early vocabularies of children with NH contain highly frequent words, except for words with the highest frequencies, i.e. function words (Goodman et al., 2008; Naigles & Hoff-Ginsberg, 1998; Stokes, 2010; Stokes et al., 2012). Similarly, in this study, highly frequent words are more accurate, except for words in the highest frequency regions, i.a. function words. The present paper revealed that this effect is not only visible in lexical acquisition, but also in word accuracy. However, our results are not in agreement with the findings of Sosa and Stoel-Gammon (2012), who did not find an effect of word frequency on word accuracy in children with NH. This contradiction may be due to methodological differences. First of all, the frequency counts of Sosa and Stoel-Gammon (2012) are problematic as they are based on adult written language, which is known to differ from child-directed speech (Akinnaso, 1982; Bernstein Ratner, 2013; Newport et al., 1977). Secondly, they used a less fine-grained measure of accuracy, i.e. proportion of word proximity (PWP), while this study used Levenshtein distance (Faes et al., 2016).

In adult spontaneous speech, function words are often reduced, because they are short forms without sentence stress. Van Bael, Baayen and Strik (2007) have shown that function words have high rates of phoneme and syllable deletions in Dutch. For instance, the Dutch word natuurlijk /natyrlək/ ‘of course’, ‘naturally’ can be found in casual speech in various different forms, including [naty(rl)ək, nylək, natyk, ntyk, tylək, tylək] (Ernestus, 2000) and the closed-class word dat (‘that’) is often produced as [da] instead of [dat]. It is likely that children integrate those forms in their own speech as well. This can possibly account for the lower accuracy of very highly frequent words in children’s speech. Except for the most frequent words, the main conclusion is that higher frequency results in better accuracy.

5.4.2 Word frequency and word accuracy in children with CI

The present study has found that word frequency in the input affects word accuracy in children with CI’s output. As for children with NH, highly frequent words are more accurate in children with CI’s productions, except for the most frequent words. To
date, only word frequency effects on lexical development were considered in the literature. Han et al. (2015) have shown that frequently produced words appear earliest in vocabularies of children with CI. Those children with CI were tested at mean ages of four and seven, i.e. one and three years after implantation. Our results complement the findings of Han et al. (2015): Children with CI are sensitive to word frequency effects in lexical development. In contrast, Guo et al. (2015) have concluded that there was no correlation between word input frequency and lexical development in children with CI. This discrepancy may be due to methodological differences. Guo et al. (2015) have studied lexical development one year after implantation, while the current study traces phonemic accuracy from immediately after implant activation up to seven years of age. Moreover, the word frequency counts are based on a corpus of CDS that is not directed to study participants in Guo et al. (2015), while this study based the frequency counts on CDS to the study’s participants.

In the present study, a longitudinal analysis of children with CI indicated that the frequency effect was mediated by two factors: (1) Chronological age, and (2) age at implant activation. First, (1), there is a discontinuity with age in children with CI’s speech. Initially, accuracy increases with word frequency, except for words in the highest frequency regions. At the older ages, however, children with CI seem less accurate in their production of more frequent words. One possible explanation may be found in the high correlation between input and output frequency (see method section, \( r = 0.79 \)). This means that words with high frequency in the input are also those with high frequency in the output, i.e. children’s own productions. Therefore, what applies to word productions in the input may very well apply to those in children’s own production as well. In adult speech (input), highly frequent words are articulated more quickly (Ellis, 2002). Such a higher speech rate results in more reduction in speech production (Ernestus, 2000). Thus, more quickly articulated words are more often reduced in the input. As a result, such words are also less accurate. It is possible that a similar effect is present in children’s own productions, because highly frequent words in children’s own productions are also those that occur frequently in the input. Thus, it is possible that children’s own productions that occur the most frequently in their speech are produced more quickly, and as a result, therefore also less accurately. In
Word frequency effects

older children, this was exactly the case: high frequent words in the input, which occur
thus also highly frequently in their output, were less accurate than words with lower
frequency in the input (and thus output). This may suggest that children with CI are
approaching an adult model. In adults as well, more frequent words are produced less
accurately.

Secondly, (2), the effect of word input frequency on word accuracy was less
pronounced in children with CI with later implant activation throughout the entire
period studied. In other words, children with CI with later implant activation are
sensitive to the frequency of words in their input, but to a lesser extent than children
with CI with earlier implant activation. A similar advantage of earlier implant
activation has been widely reported. Better outcomes are found for children with CI
with earlier implantation in morphological development (Boons, De Raeve, Langereis,
Peeraer, Wouters & Van Wieringen, 2013; Nicholas & Geers, 2007; Nikolopoulos,
Dyar, Archbold & O'Donoghue, 2004), speech production accuracy (Connor, Hieber,
Arts & Zwolan, 2000; Leigh, Detmman, Dowell & Briggs, 2013; Schauwers,
Taelman, Gillis & Govaerts, 2008; van den Berg, 2012) and lexical development
(Boons et al., 2013; Houston, Stewart, Moberly, Hollich & Miyamoto, 2012; Nicholas
& Geers, 2007). The smaller frequency effect in children with CI with later implant
activation may very well be explained by their more restricted language experience.
Children with CI with later implant activation have less language experience as
compared to earlier implanted peers. However, Weisleder and Fernald (2013) pointed
out that the amount of language experience is crucial for processing skills: Better
processing skills are found in children with more language experience. This suggests
that children with CI with later implant activation have less well-developed processing
skills as compared to earlier implanted peers. However, O'Grady (2015) proposed that
frequency effects are processing effects. In order to pick up on frequency effects,
children need to process the input adequately. Thus, good processing skills are
necessary for children’s sensitivity to language statistics. Therefore, children with
worse processing skills, i.e. children with more restricted language experience as a
result of later implant activation, may be less sensitive to word frequency in the input.
5.4.3 Differences between children with CI and children with NH

In both groups of children word frequency affects phonemic accuracy. However, the analyses have shown that the effect of word input frequency on word accuracy is more pronounced in children with NH as compared to children with CI. In other words, children with CI and children with NH are both sensitive to language statistics (i.e. word frequency), but this sensitivity is less pronounced in children with CI than in peers with NH. These results should, however, be considered with some care, due to the small sample size of the children with CI.

There are various explanations for the different degree of sensitivity to word frequency in children with CI and children with NH. A first explanation is linked to the amount of language experience of children with CI. They have less language experience than children with NH. Spoken language input starts later due to the initial auditory deprivation and after implantation the acoustic signal is still degraded as compared to the signal available in normal hearing (Wilson, 2006). For instance, Stelmachowicz et al. (2004) have shown that the degraded speech perception in children with CI mainly affects the perception of sounds produced at high spectral frequencies such as fricatives. This also has an effect on the production of these sounds in children with CI. Fricatives occur late in children with CI’s productions (Faes & Gillis, 2016; Stelmachowicz et al., 2004) and are significantly less frequent and less accurate in comparison to peers with NH (Faes & Gillis, 2016). In addition, children with CI are less attentive to speech sounds in the ambient language (Houston & Bergeson, 2014; Houston et al., 2003). The degraded signal, fewer language experience and lower attention to language may influence their sensitivity to language statistics (Saffran et al., 1996). It is possible that children with CI are only starting to discover that words differ in their frequency of occurrence (Maekawa & Storkel, 2006). In addition, the reduced language experience is related to less developed processing skills (Weisleder & Fernald, 2013). Indeed, children with CI are found to have less developed processing skills as compared to children with NH. Children with CI have problems with implicit sequence learning (Conway, Pisoni, Anaya, Karpicke & Henning, 2011) and processing and storage of information (Burkholder & Pisoni, 2003; Cleary, Pisoni & Geers, 2001; Grieco-Calub, Saffran & Litovsky, 2009;
Word frequency effects

Kronenberger, Pisoni, Henning & Colson, 2013; Nittrouer, Caldwell-Tarr & Lowenstein, 2013; Pisoni & Cleary, 2003; Pisoni & Cleary, 2004; Pisoni, Kronenberger, Roman & Geers, 2010). Thus, processing and storage skills are affected in children with CI. As O'Grady (2015) has proposed that frequency effects can be seen as processing effects, it is possible that the lower effect of input frequency is related to the worse processing skills in children with CI.

Secondly, most of the children with CI in this study had unilateral implants for a long period. Even though some children in the present study had received a second implant during data collection, the age at second implantation was highly variable and the duration of bilateral device use may have been to short to fully integrate bilateral input. For lexical development, Guo et al. (2015) have already shown that unilateral or bilateral implantation affects the sensitivity to language statistics: Children with bilateral implants are sensitive to phonotactic probability, whereas in children with unilateral implants this effect is less pronounced. It is very likely that there is a similar bilateral advantage for the word frequency effect on speech production accuracy. For instance Sarant, Harris, Bennet and Bant (2014) have shown that children with bilateral implants have better vocabulary, receptive language and expressive language outcomes and Baudonck, Van Lierde, D'haeseleer and Dhooge (2011) also have indicated better speech production skills in children with bilateral implants as compared to children with unilateral implants. These beneficial effects of bilateral implantation on speech production may have arisen from an enhanced speech perception in bilateral implanted children (Dunn, Noble, Tyler, Kordus, Gantz & Ji, 2010). To date, it remains quite unclear which information available in bilateral, but not in unilateral, stimulation is relevant for the effect of word frequency on word accuracy. Future studies should examine the effect of word frequency on word accuracy in both children with CI with unilateral and bilateral implantation.

Finally, a last explanation may be found in the speech rate of children with CI. Children with CI are slower speakers as compared to children with NH (Burkholder & Pisoni, 2003; Vanormelingen, 2016). Slower speech may result from longer pauses in children’s productions (Burkholder & Pisoni, 2003), but even when the pauses were excluded, Vanormelingen (2016) has shown that children with CI produce were
slower speakers (fewer syllables per second and longer mean syllable duration). It has already been shown that highly frequent words are produced more rapidly in adult speech (Ellis, 2002), are therefore more often reduced (Ernestus, 2000) and thus less accurate. The present paper examined the effect of frequency in the input on word accuracy in children’s output. It was also showed that input frequency is strongly correlated with output frequency. In other words, the words with the highest frequency in the input are also those with the highest frequency in children’s own speech. Thus, words with the highest frequencies in the input are most likely to be articulated faster as well. As the speech rate is slower in children with CI, it is likely that children with CI produce highly frequent words in the output less quickly. As a result, the overall speech rate of the sentence is lower and the effect of frequency on accuracy is lower as well. As input frequency and output frequency are strongly correlated, it is possible that this effect in the output is also captured by the input frequency counts.

5.4.4 Conclusion

In children with NH, sensitivity to language statistics is shown to affect segments (Edwards et al., 2004) and words (Goodman et al., 2008) children produce. The findings of the present study point out that it also affects the accuracy of known words. Moreover, the present study shows that children with CI with unilateral devices are sensitive to language statistics as well, but to a lesser extent than their peers with NH. The present paper adds to the body of knowledge of frequency effects in children with CI by considering the effect of word frequency on word accuracy. However, further research is needed to fully understand the contribution of word frequency to phonological representations and to understand the effect of word frequency on word accuracy. For instance, an analysis of the effects of children’s own word production frequency on accuracy might be relevant in the discussion about frequency effects as well. For instance Storkel (2004) has shown that early acquired words are highly frequent in children’s own productions. Thus, production frequency might be another relevant frequency effect, as perception of own production fine-tunes articulation (Moreno-Torres, 2014). In addition, own production frequency is related to word learning in children with NH: Children’s own babbling productions guide their perception, i.e. sensitivity to and focus on similar patterns in the input, and the
Word frequency effects

experience of own word productions promotes new word learning (for a discussion, see Vihman, Depaolis & Keren-Portnoy, 2014). Next, also effects of phonotactic probability (PP), neighbourhood density (ND), vocabulary size, etc. should be considered together with word frequency. Even though Sosa and Stoel-Gammon (2012) have shown that these factors are not correlated in child language, including the effects PP, ND, vocabulary size, etc. in addition to word frequency makes it possible to disentangle the different factors affecting word accuracy.
Chapter 6
Chapter 6

6 Consonant cluster production in Dutch-speaking children with cochlear implants: a comparison with peers with normal hearing

a A version of this chapter was accepted as a paper. Reference: Faes, J. & Gillis, S. (In press). Consonant cluster production in children with cochlear implants: a comparison with peers with normal hearing. First Language.

ABSTRACT
In early word productions, the same types of errors are manifest in children with cochlear implants (CI) as in their normally hearing (NH) peers with respect to consonant clusters. However, the incidence of those types and their longitudinal development has not been examined nor quantified in the literature thus far. Furthermore, studies on the spontaneous speech of Dutch-speaking children with CI are missing. Here we examine children’s use of word-initial two-consonant clusters, the frequency of each type of error and compares it in both groups of children. The spontaneous speech of 9 Dutch-speaking children with CI and an age-matched cohort of NH children was analysed from word-onset up to age seven. Results showed that accuracy and frequency of consonant clusters increases with age and that the age at implant activation is crucial in children with CI. Cross-sectional comparisons showed that some aspects of consonant cluster production in children with CI lag behind that of their NH peers, but that children with CI catch up by age five.

Keywords: Phonemic accuracy; children with CI; consonant clusters

6.1 Introduction
For a couple of decades now, cochlear implantation has offered access to spoken language for congenitally deaf children. Even though the signal provided by a cochlear implant (henceforth: CI) is still degraded compared to the signal in normal hearing (Drennan & Rubinstein, 2008), a CI enables children with severe-to-profound hearing impairment to perceive speech after a period of auditory deprivation. After
Development of consonant clusters

cochlear implantation children’s speech perception has been shown to improve (Liu, Liu, Kirk, Zhang, Ge, Zheng, Liu & Ni, 2015; Tyler, Fryauf-Bertschy, Kelsay, Gantz, Woodworth & Parkinson, 1997). As a result of (improved) speech perception, cochlear implantation is also beneficial for speech production. For instance children with CI produce the same segments as children with normal hearing (NH) (Blamey, Barry, Bow, Sarant, Paatsch & Wales, 2001; Chin, 2002; Serry & Blamey, 1999; Spencer & Guo, 2013). In addition, their articulation of vowels and singleton consonants improves after implantation (Blamey et al., 2001), although the fine phonetic details of their production remains deviant even after several years of device use (Verhoeven, Hide, De Maeyer, Gillis & Gillis, 2016). In contrast with singleton consonants, the production of consonant clusters in children with CI has hardly been studied.

The present paper traces in detail the development of word-initial two consonant (CC) clusters longitudinally in the spontaneous speech of congenitally deaf Dutch-speaking children with CI, from their first appearance up to age seven. Several aspects of the development of CC clusters are quantified: their overall accuracy, the types of errors, and the precise development of cluster reduction (for a definition, see further). Our aims are twofold. The first aim is trace the production of consonant clusters longitudinally in children with CI. Production is related to the children’s chronological age and their age at implant activation. The second aim is to statistically compare the development of consonant clusters to that of NH peers between ages two and seven.

6.1.1 Consonant cluster production in children with NH

In English-speaking NH children, word-initial (WI) consonant clusters emerge approximately around age 2;0 (McLeod, van Doorn & Reed, 2001b). Initially, these productions are inaccurate, but they gradually become more accurate with age (Phoon, Maclagan & Adbdullah, 2015). On the road to complete accuracy, three types of errors in the production of WI consonant clusters with two consonants (CC) are well attested across languages (e.g. Dutch: Fikkert, 1994; English: Greenlee, 1974; McLeod, van Doorn & Reed, 2001a; McLeod et al., 2001b): (1) both consonants are deleted (complete deletion), (2) only one consonant is produced (cluster reduction), and (3) both consonants are produced, but one or both are produced inaccurately.
(cluster simplification). Those three phenomena typically co-occur before fully accurate production (Chin, 2007; Fikkert, 1994; Jongstra, 2003; McLeod & Hewitt, 2008; McLeod et al., 2001a). In what follows, the different types of errors in NH children’s speech are discussed.

Cluster deletion

Complete deletion of a CC cluster entails the deletion of both target consonants, i.e. the adult equivalent of the child’s rendition, e.g. /blu/ (blue) produced as /u/. In the literature, complete deletion of WI consonant clusters is characterised as a rare phenomenon. For instance, in picture naming tasks involving English-speaking children (Chin & Dinnsen, 1992; Smit, 1993) and in the spontaneous speech of Dutch-speaking children (Fikkert, 1994), complete deletion of a consonant cluster is nearly absent. In contrast, cluster reduction and cluster simplification are common.

Cluster reduction

Consonant cluster reduction is the most frequently reported and attested type of errors in NH children (Dutch: Beers, 1992; English: Dodd, Holm, Hua & Crosbie, 2003; Fikkert, 1994; Chinese: Hua & Dodd, 2000; McLeod et al., 2001a; see McLeod et al., 2001b for an overview). Cluster reduction is defined as the production of a singleton consonant instead of a consonant cluster (McLeod et al., 2001b). Two reduction patterns have been observed: (1) the singleton consonant is one of the target consonants, e.g. /pl/ rendered as /p/, and (2) the singleton consonant differs from the target consonants, e.g. /pl/ rendered as /b/.

The first reduction pattern (1) is explained by the sonority hypothesis (SH). The SH predicts reduction patterns based on the sonority of segments (Geirut, 1999; Jongstra, 2003; McLeod et al., 2001a; McLeod et al., 2001b; Ohala, 1999; Wyllie-Smith, McLeod & Ball, 2006). Segments are ordered from less to more sonorous according to the Sonority Sequencing Principle (Clements, 1990; Geirut, 1999; Ohala, 1999), which ranks plosives as the least sonorous segments and vowels as the most sonorous, as shown in (1):

(1) Plosives < Fricatives < Nasals < Liquids < Glides < Vowels

Least sonorous Most sonorous
The SH holds that when a cluster is reduced to a single consonant, the least sonorous consonant is preserved, resulting in a maximal sonority distance between the onset consonant and the (vocalic) nucleus. This regularity is in agreement with the universally preferred CV syllable: across languages, the preferred CV syllable exhibits a maximum rise in sonority between onset and nucleus (Vennemann, 1988). Similarly, the SH predicts that the least sonorous consonant is preserved in initial cluster reduction, which results in a maximal contrast in sonority between the preserved consonant and the following vowel (Fikkert, 1994; Ohala, 1999). For instance, the SH predicts that when a plosive plus liquid cluster (e.g. /pl/) is reduced, the plosive (/p/) will be preserved. Even though there is ample evidence for adherence to the SH in, for instance, English-speaking NH children (Chin & Finnegan, 2002), Jongstra (2003) showed that in children acquiring Dutch there is considerable variation in the reduction patterns between children and even within the same child.

The second reduction pattern involves the production of a singleton consonant that differs from the target consonants. The new consonant often combines features of both target consonants. This is called coalescence (Chin & Dinnsen, 1992; Dyson & Paden, 1983; McLeod et al., 2001a; McLeod et al., 2001b). For instance, when spider /spaɪdər/ is produced as [faldər], the manner feature of /s/ merges with the place feature of /p/, to become the labial fricative [f]. Such instances of coalescence suggest that the child has at least some knowledge of the two target consonants (Chin & Dinnsen, 1992). This substitution pattern cannot, however, be explained by the SH (Wyllie-Smith et al., 2006).

Cluster simplification

Next to cluster reduction, cluster simplification is a common type of errors in NH children’s speech. In cluster simplification, a consonant cluster is produced, but at least one consonant deviates from the target (McLeod et al., 2001b), as in frog /frɔɡ/ produced as [fwɔɡ].
consonant cluster production, or they do not provide a longitudinal and quantified developmental picture. More specifically, some studies only analyse cluster accuracy without further analysing the types of errors (Fulcher, Baker, Purcell & Munro, 2014; Von Mentzer, Lyxell, Sahlén, Dahlström, Lindgren, Ors, Kallioinen, Engström & Uhlén, 2015). Other studies are limited to a single aspect of cluster production, such as cluster reduction (Baudonck, Dhooge, D’haeseleer & Van Lierde, 2010; Flipsen & Parker, 2008), without considering other types of errors.

A second strand of research addresses children’s rendition of clusters at one particular point in their development without considering longitudinal development (Baudonck et al., 2010; Chin & Finnegan, 2002; Fulcher et al., 2014; Von Mentzer et al., 2015). Alternatively, some studies provide an extensive qualitative overview of the types of consonant clusters produced by children with CI and frame the development theoretically (Adi-Bensaid & Ben-David, 2010; Chin & Finnegan, 2002), but do not report development quantitatively.

To date, the only study on Dutch-speaking children with CI is Baudonck et al. (2010), who report that consonant cluster reduction occurs at a mean age of 9;0 (range 5;4 – 13;7). No further information is available on Dutch-speaking children with CI thus far. The present paper provides a detailed, longitudinal and quantitative study of consonant cluster production in children with CI acquiring Dutch as their native language. Since the same types of errors have been observed across languages, we expect to find similar patterns in children acquiring Dutch, though particular language-specific effects may be evident (Yavas, 2013).

Adi-Bensaid and Ben-David (2010) and Chin and Finnegan (2002) provide detailed, qualitative analyses of children with CI acquiring Hebrew and English respectively. They show that complete deletion of consonant clusters is rare in Hebrew children between 2;0 and 4;5 (Adi-Bensaid & Ben-David, 2010) and even absent in English at age 9;9 (Chin & Finnegan, 2002). Two patterns of consonant cluster reduction are commonly reported: (1) reduction to a singleton consonant that is part of the target cluster, and (2) reduction to another singleton consonant. With respect to the first pattern (1), Hebrew-speaking children with CI preserve the second consonant, except when it is liquid (Adi-Bensaid & Ben-David, 2010), whereas English-speaking
Development of consonant clusters

children with CI adhere to the sonority hypothesis and thus the least sonorous segment is produced (Chin, 2006; Chin & Finnegan, 2002). With respect to the second reduction pattern (2), Adi-Bensaid and Ben-David (2010) claim that coalescence only rarely occurs in Hebrew-speaking children with CI. Finally, regarding cluster simplification, Chin and Finnegan (2002) show that 34% of the target clusters were simplified by English-speaking children with CI at a mean age of 9;9.

These studies give an interesting overview of consonant cluster production. However, it should be noted that in Adi-Bensaid and Ben-David (2010) and in Chin and Finnegan (2002) a picture-naming task was used. How well do these findings generalize from a fairly controlled task to spontaneous speech in naturalistic interactions? Children’s speech production has been shown to contain fewer errors in single-word tasks than in connected speech (Healy & Madison, 1987), hence the results of such studies may well overestimate children’s accuracy in naturalistic spontaneous speech. To the best of our knowledge, only Flipsen and Parker (2008) and Fulcher et al. (2014) have collected spontaneous speech samples of English-speaking children with CI in their studies of CC clusters, but they only reported on consonant cluster reduction and consonant cluster accuracy respectively. In the present study the accuracy of consonant cluster production and various error patterns in children’s spontaneous speech will be analysed.

Moreover, Adi-Bensaid and Ben-David (2010) and Chin and Finnegan (2002) present a qualitative overview of consonant cluster production in children with CI, but they do not provide information about the incidence of consonant clusters in children with CI, about the likelihood or the relative incidence of the different types of errors, and the like. Thus, even though the literature has provided extensive qualitative overviews (with examples), a quantified picture of consonant cluster production in children with CI is still lacking. Our goal is to quantify the incidence of accurate production, the incidence of the various types and subtypes of errors.

In the literature on children with CI, often only one point in development is considered (Baudonck et al., 2010; Chin & Finnegan, 2002; Fulcher et al., 2014; Von Mentzer et al., 2015). Such a snapshot leaves longer term developmental patterns
unrevealed and leaves unanswered the question of whether children with CI ultimately reach a level of accuracy comparable to their NH peers.

A notable exception is Flipsen and Parker (2008), who did collect longitudinal speech samples of English-speaking children with CI. They report that consonant cluster reduction does not decrease between the (mean) ages five and seven. Information about other aspects of consonant cluster production is not provided. For Hebrew, Adi-Bensaid and Ben-David (2010) adopted a longitudinal approach as well, but the effect of age is not quantified. Instead, the authors focussed on the different types of errors in consonant cluster production over the entire study period (word onset – the age of seven). Even though they provide an order in which particular types of errors appear, no precise information about their incidence relative to the children’s age is provided. Flipsen and Parker (2008) note that a primary question in language acquisition research is: “What occurs at what age?”. It is surprising that the precise quantitative development with age has not been considered for children with CI. The present paper expands previous work by studying the production of consonant clusters longitudinally. Age will be entered as a predictor in the statistical analysis of each aspect of consonant cluster production.

In children with CI, not only the child’s chronological age is shown to affect language development, but also the age at implant activation, which is usually one or two months after surgery took place. Providing access to sound early in life, and thus early activation and fitting of the implant, is shown to be beneficial for grammatical development (Boons, De Raeve, Langereis, Peeraer, Wouters & Van Wieringen, 2013; Nikolopoulos, Dyar, Archbold & O'Donoghue, 2004), speech production (Leigh, Detmman, Dowell & Briggs, 2013), speech production accuracy (Connor, Craig, Raudenbush, Heavner & Zwolan, 2006; Schauwers, Taelman, Gillis & Govaerts, 2008; van den Berg, 2012), and various other aspects of language development. At present, no information about the effect of age at implantation on consonant cluster production is available in the literature, and we aim to address this gap.

In the literature, inferential statistical comparisons of consonant cluster production in children with CI and NH children are only available for the accuracy of consonant cluster production. Von Mentzer et al. (2015) showed that the accuracy of consonant
Development of consonant clusters

clusters in nonword repetition is lower in Swedish-speaking children with CI as compared to NH peers at a mean age of 6;6. In contrast, Fulcher et al. (2014) found no difference in consonant cluster accuracy between English-speaking children with CI and NH peers at ages 3;0 and 4;0. This difference may be caused by the study design: Von Mentzer et al. (2015) studied only three- and four-syllable nonwords in a nonword repetition task, whereas Fulcher et al. (2014) used a picture-naming task including mono- and disyllabic words, but also studied spontaneous speech in which shorter words appear as well. As word length affects children’s speech production accuracy (Faes, Gillis & Gillis, 2016; Gathercole, Willis, Emslie & Baddeley, 1991), this may explain why Von Mentzer et al. (2015) found lower accuracy rates. The present paper compares not only the accuracy of consonant clusters in both groups of children, but also the likelihood of each type of inaccuracy and each pattern of cluster reduction in Dutch-speaking children with CI and NH children by means of inferential statistics.

6.1.3 The current study

The current study has two goals: (a) trace the development of consonant cluster production in the spontaneous speech of Dutch-speaking children with CI longitudinally, and (b) compare this development to that of NH age-matched peers. Regarding the first research goal (a), it is as yet unclear if the age at implant activation and maturation with age significantly affect consonant cluster production in children with CI. Regarding the second research goal (b), it remains to be seen if children with CI differ significantly from their NH peers on the incidence of consonant clusters, the incidence of the different types of errors, and the incidence of the different patterns in consonant cluster reduction. A detailed quantitative analysis of consonant cluster production is provided for both research goals: the likelihood of consonant clusters is considered, as well as their accuracy and the likelihood of the different types of errors, i.e. complete deletion of the consonant cluster, consonant cluster reduction and consonant cluster simplification. In addition, the different patterns of consonant cluster reduction are examined: do children reduce more often to a consonant that is one of the target consonants? If so, to what extent does the sonority hypothesis explain
which consonant is preserved? If not, to what extent does coalescence account for the non-target consonant that is produced?

6.2 Method

6.2.1 Participants

A longitudinal design was set up to assess language development of children with CI: 9 children with CI were followed from word onset up to the age of five (Schauwers, 2006). The children were followed monthly from the moment of implant activation up to 2;6 years after activation. Additionally, yearly data were collected between the age of 3;0 and 5;0. Here, the data will be analyzed from the appearance of first spoken words (median = 1;6, range 1;3 – 1;11). In chapter 2, more detailed information on the children with CI can be found.

The control group consisted of a selection of the NHlongitudinal corpus at 24 months of age (only children with more than 5% consonant clusters in their productions and then random, S11, S13, S20, S22, S28, S29, S30, S31, S33, S37 and S38) and the NHcross-sectional corpus between ages 3;0 and 7;0 (S51 – S103). More information can be found in chapter 2.

6.2.2 Procedure

The data used in the current study were gathered as a part of a larger corpus, namely the CLiPS Child Language Corpus (CCLC). Only speech samples in which lexical items appeared were analyzed. Detailed transcriptions of the data collection, data annotation and the reliability of the transcriptions are provided in the method section (chapter 2).

6.2.3 Amount of data

Table 1 displays an overview of the total number of word tokens in the children’s speech as well as the incidence of consonant clusters in the targets, i.e. the adult equivalents of the child’s renditions. The total number of word tokens was 109,995, i.e. words with empty, singleton and complex onsets (CI: 64,035 word tokens, NH: 45,960 word tokens). In 7.87% of all word tokens, a word initial consonant cluster appeared in the adult form of the child’s actual production: 8.13% in the CI corpus
Development of consonant clusters

and 7.50% in the NH corpus. Three-element (CCC) clusters represented only 2.90% of all WI consonant clusters (2.61% and 3.34% for the CI group and the NH group respectively), and are not analysed.

Table 1. The distribution of consonant clusters in the targets of children’s productions

<table>
<thead>
<tr>
<th></th>
<th>All children</th>
<th>CI corpus</th>
<th>NH corpus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of word tokens</td>
<td>109,995</td>
<td>64,035</td>
<td>45,960</td>
</tr>
<tr>
<td>Number of word tokens with WI consonant cluster</td>
<td>8656</td>
<td>5209</td>
<td>3447</td>
</tr>
<tr>
<td>All consonant clusters</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC clusters</td>
<td>8405</td>
<td>5073</td>
<td>3332</td>
</tr>
<tr>
<td>CCC clusters</td>
<td>251</td>
<td>136</td>
<td>115</td>
</tr>
</tbody>
</table>

In Dutch WI CC clusters, two combinations of consonants are possible: obstruent plus obstruent and obstruent plus sonorant (Booij, 1995). In (2) examples of the combinatorial subtypes are presented:

(2)

Obstruent + obstruent

(2a) fricative+stop  stop /stop/ (Eng.: stop), spelen /spelən/ (to play)
(2b) fricative+fricative school /sχol/ (school), sfeer /sfer/ (ambiance)
(2c) stop+fricative  pseudo /psœdo/ (pseudo)

Obstruent + sonorant

(2d) stop+nasal  knippen /knIpən/ (to cut)
(2e) stop+liquid  broer /brur/ (brother), klas /klɑs/ (classroom)
(2f) stop+glide  kwaad /kwat/ (angry)
(2g) fricative+nasal snoep /snup/ (sweets, candy)
(2h) fricative+liquid slang /slaŋ/ (snake), vriend /vrint/ (friend)
(2i) fricative+glide  zwart /zwart/ (black)

6.2.4 Data analyses

Seven aspects of CC cluster production were studied:
(1) In order to compute the incidence of CC clusters, a list of all children’s word productions and their targets, i.e. adult equivalents, was retrieved from the corpus. The onset of each word production of a child, irrespective of the number of consonants in the target word, was identified as (a) empty, (b) singleton consonant, (c) CC cluster. The likelihood of CC clusters in children’s productions was estimated relative to empty and singleton onsets.

(2) The likelihood of accurately produced CC clusters was estimated. In this analysis, only those word productions with CC clusters in the onset of the adult form were considered. The child’s rendition was compared to the target CC cluster and classified as correct/incorrect. After this second layer in the analyses, inaccurately produced CC clusters were further analysed.

(3) The likelihood of complete deletion of the CC cluster within the subset of inaccurately produced CC clusters was estimated.

(4) The likelihood of CC cluster reduction was estimated and compared to the likelihood of CC cluster simplification. In this analysis, only inaccurately produced CC clusters that were not entirely deleted were considered. Children’s renditions that comprised only one consonant were labelled \textit{CC cluster reduction}, renditions with 2 consonants \textit{CC cluster simplification}. After this fourth aspect, we further investigated those CC clusters that were labelled \textit{CC cluster reduction}.

(5) The likelihood of cluster reduction to one of the target consonants was estimated. If the singleton rendition was a consonant that also occurs in the target cluster, the rendition was labelled \textit{reduction to one of the target consonants}, otherwise it was labelled \textit{no reduction to one of the target consonants}.

(6) The likelihood of adherence to the sonority hypothesis was estimated when the reduction was labelled \textit{reduction to one of the target consonants}.

(7) The likelihood of coalescence was estimated when the reduction was labelled \textit{no reduction to one of the target consonants}.

\textbf{6.2.5 Statistical analyses}

All statistical analyses were done in R (R Core Team, 2013) by means of logistic regressions in multilevel models (Baayen, 2008; Woltman, Feldstain, MacKay & Rocchi, 2012). Our data are hierarchically structured: utterances are nested in
Development of consonant clusters

individual children at different ages. Multilevel models take this variation into account. Multilevel models consist of two parts: a random part and a fixed part. The random part of the model considers the nesting of variables in the data, whereas the fixed part includes the predicting variables. In R, the estimates and standard errors (SE) of logistic regressions are computed in logits. Logits can easily be converted to probabilities in two steps. First, logits are converted to odds using an exponential function (step 1 in equation (A)). Second, the odds are converted to probabilities using the formula in step 2 of equation (A):

(A) Step 1: \[ \text{odds} = e^{\text{logit}} \]

Step 2: \[ p = \frac{\text{odds}}{1 + \text{odds}} \]

The analyses are presented in two steps: (1) a longitudinal analysis of children with CI and (2) cross-sectional comparisons between NH children and children with CI. For the longitudinal analysis of children with CI (1), all available data of children with CI were used: the monthly data from word onset (median = 1;6, range 1;3–1;11) up to age 2;6 and the yearly data of those children between ages 3;0 and 7;0. For these analyses, the random part of the multilevel models of each analysis included random intercepts and random slopes to model the variation between children and ages. The fixed part, i.e. the predicting variables, was composed in an incremental way. A particular predictor was added to the model and an ANOVA was used to test whether the resulting model had a significantly better model fit compared to the previous model, i.e., the model without that particular predictor. If the resulting model had a better model fit, the predicting variable was preserved in the model – even if the variable itself was not significant. The predicting variables included in each of the longitudinal analyses of children with CI were chronological age in months (Age, centred at 24 months of age) and the age at implant activation (CIactivation). Quadratic and cubic age effects (Age2, Age3) and interactions between Age and CIactivation were tested as well. Only the best fitting model is reported.

The cross-sectional analyses (2) were performed in order to compare the development of consonant clusters in children with CI and their NH age-matched
peers. For these cross-sectional comparisons, the data of the CI group were split, as no longitudinal data were available for the NH children. It would be incorrect to include the cross-sectional data of all NH children in one model with the longitudinal data of the CI group. NH children and children with CI were matched and compared at ages 2;0 (range 1;11–2;1), 3;0 (2;10–3;4), 4;0 (3;9–4;3), 5;0 (4;11–5;3), 6;0 (5;7–6;6) and 7;0 (6;10–7;4). In each analysis, a random intercept was included in the multilevel model, taking the variation between children into account in the random part of the multilevel model, and the predicting variable (the fixed part of the multilevel model) was HearingStatus (NH vs. CI).

6.3 Results

6.3.1 Longitudinal analysis of children with CI

This section presents the longitudinal development of WI CC clusters of children with CI relative to their chronological age (Age) and the age at implant activation (CIactivation). The tables represent the best fitting models expressed in logits.

6.3.1.1 The likelihood of CC clusters and the likelihood of accurate CC clusters

Table 2 presents the fixed effect results of the best fitting models for the likelihood of CC clusters and the likelihood of accurate CC clusters in the productions of children with CI.

The likelihood of words with CC clusters is significantly lower than the likelihood of words with empty and singleton onsets (intercept \( p<0.001 \)). At the intercept (i.e. 24 months of age), the likelihood of a CC cluster is 2.08%. Table 2 also displays a significant effect of Age \( (p<0.001) \), showing that the likelihood of CC clusters increases as children get older. The quadratic effect of age \( (\text{Age}^2) \) is also significant \( (p<0.001) \). In Figure 1, the effects of Age and \( \text{Age}^2 \) are plotted: there is an increase with age (Age), but this increase becomes less steep from approximately 36 months of age \( (\text{Age}^2) \) onwards. Finally, Table 2 shows that the effect of CIactivation is not significant \( (p>0.05) \), indicating that the likelihood of CC clusters is not statistically different in children with CI with different ages at implantation.
Development of consonant clusters

Table 2 also shows the likelihood of accurate CC clusters as compared to that of inaccurate CC clusters. The intercept is not significant (p>0.05), which means that the likelihood of accurately produced CC clusters is not significantly lower or higher as compared to that of inaccurately produced CC clusters. At 24 months of age, the likelihood of an accurately produced CC cluster is 46.01%. There is a significant effect of Age (p<0.01): as shown in Figure 2, the likelihood of accurately produced CC clusters increases with age. There is no significant quadratic effect of Age², nor did this effect improve the model fit, therefore it is not included in the model. Next, there is a significant effect of CI activation (p<0.001), indicating that the likelihood of an accurately produced CC cluster is lower in children with CI with later implant activation. In addition, the significant interaction between CI activation and Age (p<0.001) shows that the increase of the likelihood of accurately produced CC clusters is steeper in children with CI with later implant activation. Thus, they are catching up on the likelihood of accurate CC clusters.

Table 2. Fixed effect results of the likelihood of CC clusters and the likelihood of accurate CC clusters

<table>
<thead>
<tr>
<th></th>
<th>Likelihood of consonant clusters</th>
<th>Likelihood of accurate consonant clusters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate (SE)</td>
<td>Estimate (SE)</td>
</tr>
<tr>
<td>Intercept</td>
<td>-3.85 (0.30) ***</td>
<td>-0.16 (1.11)</td>
</tr>
<tr>
<td>Age</td>
<td>0.12 (0.01) ***</td>
<td>0.09 (0.03) **</td>
</tr>
<tr>
<td>Age²</td>
<td>-0.01 (&lt;0.01) ***</td>
<td></td>
</tr>
<tr>
<td>CI activation</td>
<td>-0.03 (0.02)</td>
<td>-0.27 (0.08) ***</td>
</tr>
<tr>
<td>CI activation x Age</td>
<td></td>
<td>0.01 (&lt;0.01) ***</td>
</tr>
</tbody>
</table>

p≤0.05*, p≤0.01**, p≤0.001***
6.3.1.2 Types of inaccuracies: the likelihood of complete deletion, reduction and simplification

This section reports on the development of inaccurately produced CC clusters. The fixed effect results of the best fitting models are displayed in Table 3, expressed in logits. In figure 5, the development of the likelihood of each type of errors is displayed.
Development of consonant clusters

The left panel of Table 3 shows the likelihood of complete deletion of the CC clusters. The negative intercept shows that CC clusters are significantly less likely to be completely deleted than to be either reduced or simplified (p<0.001). At 24 months of age, the likelihood of complete deletion of the CC cluster is 4.15%. There is a significant decrease with age (Age and Age², p<0.001), as also shown in Figure 3. The significant effect of CIactivation (p<0.001) indicates that the likelihood of complete deletion of the CC cluster is higher in children with CI with later implant activation. In addition, the decrease of complete deletion of CC clusters with age is less steep in children with CI with later implant activation (interaction between Age and CIactivation, p<0.001).

When CC clusters are not entirely deleted, they can either be reduced or simplified. The right panel of Table 3 displays the likelihood of cluster reduction as compared to that of cluster simplification. The likelihood of cluster reduction is significantly higher than that of cluster simplification (intercept p<0.001). At 24 months of age, the likelihood of cluster reduction is 88.29%, and evidently the likelihood of cluster simplification is 11.71%. There are significant age effects (Age, Age² and Age³) that can be inferred from Figure 4: the likelihood of cluster reduction decreases with age (and thus cluster simplification increases). However, this decrease levels out by approximately 70 months of age. Finally, there is a significant effect of CIactivation (p<0.001), which indicates that the likelihood of cluster reduction is higher in children with CI with later implant activation. There was no significant interaction between Age and CIactivation nor did this interaction improve the model fit. Therefore it is left out of the best fitting model reported in Table 4. However, the lack of an interaction between Age and CIactivation shows that children with CI with later implant activation are not catching up with their CI peers with earlier implant activation as to the likelihood of cluster reduction versus cluster simplification.
Table 3. Fixed effect results of the likelihood of CC cluster deletion and CC cluster reduction

<table>
<thead>
<tr>
<th></th>
<th>Likelihood of complete cluster deletion</th>
<th>Likelihood of cluster reduction (vs. simplification)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate (SE)</td>
<td>Estimate (SE)</td>
</tr>
<tr>
<td>Intercept</td>
<td>-3.14 (&lt;0.01) ***</td>
<td>2.02 (0.01) ***</td>
</tr>
<tr>
<td>Age</td>
<td>-0.20 (0.06) ***</td>
<td>-0.20 (0.01) ***</td>
</tr>
<tr>
<td>Age^2</td>
<td>-0.01 (&lt;0.01) ***</td>
<td>&lt;-0.01 (0.01) **</td>
</tr>
<tr>
<td>Age^3</td>
<td>&lt;0.01 (&lt;0.01) ***</td>
<td>&lt;0.01 (&lt;0.01) ***</td>
</tr>
<tr>
<td>Clactivation</td>
<td>0.05 (&lt;0.01) ***</td>
<td>0.13 (0.01) ***</td>
</tr>
<tr>
<td>Clactivation x Age</td>
<td>0.01 (&lt;0.01) ***</td>
<td></td>
</tr>
</tbody>
</table>

p≤0.05*, p≤0.01**, p≤0.001***

Figure 3. Development of the likelihood of CC cluster deletion (expressed in logits)
Development of consonant clusters

6.3.1.3 Reduction processes
This section presents the development of cluster reduction in more detail. The fixed effect results of the best fitting models – expressed in logits – are shown in Table 4.

We discuss, first, the likelihood of reduction to one of the target consonants, secondly,
the likelihood of adherence to the sonority hypothesis and, thirdly, the likelihood of coalescence.

Table 4 shows the likelihood of reduction to one of the target consonants. The intercept is not statistically significant (p>0.05), meaning that the likelihood of reducing a cluster to a target consonant is not significantly more or less likely than that of reducing it to another consonant (62.48%). However, the significant effect of Age (p<0.05) suggests that the likelihood of reducing a CC cluster to one of the target consonants increases. Figure 6 shows that this increase is quite small. Moreover, there is a significant effect of CIactivation (p<0.05), indicating that the likelihood of reduction to one of the target consonants is lower in children with CI with later implant activation. Including the interaction between Age and CIactivation did not improve the model fit and is therefore left out. The lack of a significant interaction indicates, however, that children with CI with later implant activation are not catching up with their earlier implanted peers.

When CC clusters are reduced to a consonant that is part of the target cluster, the sonority hypothesis (SH) predicts that children preserve the least sonorous consonant in word onset position. Table 4 shows the likelihood of adherence to the SH as compared to no adherence. At 24 months of age, the likelihood of adherence to the SH is 95.35% (intercept p<0.001), which is significantly higher than the likelihood of non-adherence (4.65%). No effect of Age (p>0.05) nor CIactivation (p>0.05) are found, showing that the likelihood of adherence to the SH remains stable over time (Figure 7) and is similar in all children with CI, regardless of their age at implant activation.

When CC clusters are not reduced to a consonant that is part of the target cluster, the consonant that is produced often combines features of both target consonants. Table 4 shows the analysis of the likelihood of coalescence. The intercept is significant (p<0.05): at 24 months of age, the likelihood of coalescence is 86.99%, which approaches a ceiling level. There is no statistically significant effect of Age (p<0.05), indicating that the likelihood of coalescence remains relatively stable, as can be seen in Figure 8. Finally, the effect of CIactivation is significant (p<0.05), showing
Development of consonant clusters

that the likelihood of coalescence is significantly lower in children with CI with later implant activation.

Table 4. Fixed effect results of the likelihood of the different reduction processes

<table>
<thead>
<tr>
<th></th>
<th>Likelihood of reduction to one of the target consonants</th>
<th>Likelihood of adherence to the sonority hypothesis</th>
<th>Likelihood of coalescence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.51 (0.53)</td>
<td>3.02 (0.83) ***</td>
<td>1.90 (0.88) *</td>
</tr>
<tr>
<td>Age</td>
<td>0.06 (0.03)</td>
<td>-0.06 (0.04) *</td>
<td>0.04 (0.05)</td>
</tr>
<tr>
<td>CI activation</td>
<td>-0.07 (0.03) *</td>
<td>-0.05 (0.05)</td>
<td>-0.16 (0.06) *</td>
</tr>
</tbody>
</table>

p ≤ 0.05*, p ≤ 0.01**, p ≤ 0.001***

Figure 6. Development of the likelihood of CC cluster reduction: reduction to one of the target consonants (expressed in logits)
6.3.2 Comparisons between children with CI and children with NH

This section provides the results of the cross-sectional comparisons between children with CI and NH children, between ages 2;0 and 7;0. The tables exhibit the fixed effect results for the predicting variable HearingStatus, expressed in logits. In these cross-sectional comparisons the effect of Hearing status is considered at each age. This
Development of consonant clusters implies that no interactions between Hearing status and Age can be considered, as each analysis gives an indication for a particular age.

6.3.2.1 The likelihood of CC clusters and the likelihood of accurate CC clusters

Table 5 presents the cross-sectional comparisons between NH children and children with CI for the likelihood of CC clusters and the likelihood of accurate CC clusters.

Table 5 and Figure 9 show no significant differences between both groups of children (p>0.05) as to the likelihood of CC clusters. However, the likelihood of accurate CC clusters reveals significant differences between the two groups, as can be inferred from Table 5 and Figure 10. The effect of HearingStatus is significant at ages 2;0, 3;0 and 4;0 (p<0.001, p<0.05 and p<0.001), meaning that the likelihood of producing a CC cluster accurately is significantly higher in NH children than in children with CI at these ages. At age 2;0, the likelihood is 8.71% in NH children and only 3.52% in children with CI, at age 3;0, the likelihood is 73.31% and 33.84% respectively and at age 4;0, the likelihood is 91.68% and 67.48% respectively. From age 5;0 onwards, there are no longer significant group differences, because accuracy reaches ceiling percentages in both groups of children (99.55%).

Table 5. Fixed effect results of the cross-sectional comparisons between NH and CI: likelihood of CC clusters and the likelihood of accurate CC clusters

<table>
<thead>
<tr>
<th></th>
<th>2;0</th>
<th>3;0</th>
<th>4;0</th>
<th>5;0</th>
<th>6;0</th>
<th>7;0</th>
</tr>
</thead>
<tbody>
<tr>
<td>The likelihood of consonant clusters</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-3.97 (0.35)***</td>
<td>-2.99 (0.16)***</td>
<td>-2.53 (0.11)***</td>
<td>-2.56 (0.16)***</td>
<td>-2.74 (0.10)***</td>
<td>-2.84 (0.11)***</td>
</tr>
<tr>
<td>HearingStatus</td>
<td>0.56 (0.45)</td>
<td>0.06 (0.24)</td>
<td>-0.14 (0.15)</td>
<td>0.11 (0.22)</td>
<td>0.10 (0.13)</td>
<td>0.08 (0.14)</td>
</tr>
<tr>
<td>[NH]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The likelihood of accurate consonant clusters</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-3.31 (&lt;0.01)***</td>
<td>-0.67 (0.39)</td>
<td>0.73 (0.35)*</td>
<td>5.39 (0.94)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No inaccurate clusters in CI.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5.99 (1.10)***</td>
<td></td>
</tr>
<tr>
<td>Intercept NH: 3.73 (0.27)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
HearingStatus   | 0.96  | 1.63 | 1.67 | -1.87 | 0.10 | -2.07  \\
[NH]           | (<0.01)*** | (0.59)* | (0.48)*** | (0.99) | (0.13) | (1.14)  \\

*p≤0.05*, *p≤0.01**, *p≤0.001***

CI is the reference category

**Figure 9.** Probability of CC clusters in NH children and children with CI (predicted values)
6.3.2.2 Types of inaccuracies: the likelihood of complete deletion, reduction and simplification

From this section onwards, the cross-sectional comparisons are restricted to ages 2;0, 3;0 and 4;0, as accuracy reaches ceiling percentages from age 5;0. In Table 6, the fixed effect results of the cross-sectional comparisons between NH children and children with CI are presented for the different types of errors. As Table 6 and Figure 11 show, the likelihood of complete cluster deletion is similar in both groups of children at age 2;0 (p>0.05) and could not be estimated after that age, as there were no more deletions from age 3;0 onwards. For cluster reduction, Table 6 and Figure 12 show a significant effect of HearingStatus at age 2;0 (p<0.001): the likelihood of cluster reduction is significantly lower in NH children as compared to children with CI. Hence, cluster simplification is more likely in NH children than in children with CI. From age 3;0 onwards, no significant effects of HearingStatus are found (p>0.05): the difference found at 2;0 has rapidly faded out.
Table 6. Fixed effect results of the cross-sectional comparisons between NH and CI: the likelihood of complete deletion, reduction and simplification

<table>
<thead>
<tr>
<th></th>
<th>2:0</th>
<th>3:0</th>
<th>4:0</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Likelihood of complete cluster deletion</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-3.51 (0.57)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HearingStatus [NH]</td>
<td>1.04 (0.68)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No more deletions occurring in the dataset</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Likelihood of cluster reduction (vs. cluster simplification)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>3.13 (&lt;0.01)***</td>
<td>0.92 (0.56)</td>
<td>-1.85 (0.43)*</td>
</tr>
<tr>
<td>HearingStatus [NH]</td>
<td>-0.62 (&lt;0.01)***</td>
<td>-0.61 (0.92)</td>
<td>0.20 (0.66)</td>
</tr>
</tbody>
</table>

*p ≤ 0.05*, *p ≤ 0.01**, *p ≤ 0.001***

CI is the reference category

Figure 11. Probability of complete cluster deletion in NH children and children with CI (predicted values)
Development of consonant clusters

Figure 12. Probability of cluster reduction (vs. cluster simplification) in NH children and children with CI (predicted values)

6.3.2.3 Reduction processes

The last part of the results presents the cross-sectional comparisons between NH children and children with CI regarding the reduction processes. The statistical models are displayed in Table 7.

The likelihood of reduction to one of the target consonants as opposed to reduction to another consonant is similar at age 2;0 (intercept p>0.05). In addition, there is no difference between both groups of children at this age, as the effect of HearingStatus was not significant (p>0.05). By age 3;0, however, a different picture appears. In children with CI, the likelihood of reduction to one of the target consonants is still not statistically significantly different from that of reduction to another consonant (intercept p>0.05). The NH children, however, differ significantly from children with CI (p<0.05). At age 3;0, the likelihood of reduction to a target consonant is 71.10% in NH children, whereas this is only 47.25% in children with CI. At age 4;0, the effect of HearingStatus was no longer statistically significant (p>0.05). Moreover, at this age,
all children are more likely to reduce to one of the target consonants than to another consonant, as shown by the significant effect at the intercept (p<0.001). The development of both patterns is shown in Figure 13.

When children reduce the CC cluster to one of the target consonants, the sonority hypothesis (SH) predicts that the least sonorous consonant is preserved word initially. Table 7 shows no significant effects of HearingStatus regarding the likelihood of adherence to the SH. This means that children with CI are equally likely to adhere to the SH as NH children at all ages (2;0, 3;0 and 4;0). The development of adherence to the SH is shown in Figure 14.

When children do not reduce a CC cluster to one of the target consonants, the new consonant often comprises features of the target consonants, i.e. coalescence. No significant effects of HearingStatus regarding the likelihood of coalescence are found in Table 8, indicating that coalescence is equally likely in both groups of children at all ages (2;0, 3;0 and 4;0). The development of coalescence is shown in Figure 15.

Table 7. Fixed effect results of the cross-sectional comparisons between NH and CI: likelihood of the different reduction processes

<table>
<thead>
<tr>
<th></th>
<th>2;0</th>
<th>3;0</th>
<th>4;0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood of reduction to one of the target consonants</td>
<td>Intercept</td>
<td>-0.07 (0.36)</td>
<td>-0.11 (0.16)</td>
</tr>
<tr>
<td></td>
<td>HearingStatus [NH]</td>
<td>0.25 (0.46)</td>
<td>1.01 (0.42)*</td>
</tr>
<tr>
<td>Likelihood of adherence to the sonority hypothesis</td>
<td>Intercept</td>
<td>4.99 (1.40)***</td>
<td>1.45 (0.32)***</td>
</tr>
<tr>
<td></td>
<td>HearingStatus [NH]</td>
<td>-2.28 (1.42)</td>
<td>1.92 (1.15)</td>
</tr>
<tr>
<td>Likelihood of coalescence</td>
<td>Intercept</td>
<td>0.61 (0.45)</td>
<td>-0.07 (0.43)</td>
</tr>
<tr>
<td></td>
<td>HearingStatus [NH]</td>
<td>-0.56 (0.58)</td>
<td>2.04 (1.12)</td>
</tr>
</tbody>
</table>

p≤0.05*, p≤0.01**, p≤0.001***
CI is the reference category
Development of consonant clusters

Figure 13. Probability of reduction to one of the target consonants in NH children and children with CI (predicted values)
6.4 Discussion

The present paper has traced the longitudinal development of word initial CC consonant clusters in the spontaneous speech of Dutch-speaking children with CI. Their use of CC clusters was compared in cross-sectional analyses with a group of age-matched NH children.

6.4.1 Longitudinal analysis of children with CI

The longitudinal analyses of children with CI revealed two main conclusions, the first with respect to development with age and the second with respect to age at implant activation.
Development of consonant clusters

6.4.1.1. Development with age

Our results show that as Dutch-speaking children with CI get older, consonant clusters become more frequent in their spontaneous speech. Their likelihood of producing consonant clusters accurately also increases with age. Thus, with age, they produce more consonant clusters and their consonant cluster productions become more accurate, which is in line with Chin and Finnegan (2002)’s findings for English-speaking children with CI and those of Phoon et al. (2015) for English-speaking NH children.

With respect to errors, our results show that complete deletion, reduction and simplification of the consonant cluster co-occur in the spontaneous speech of Dutch-speaking children with CI. A similar observation has been made for NH children (Chin, 2007; Fikkert, 1994; Jongstra, 2003; McLeod & Hewitt, 2008; McLeod et al., 2001a) and for English- and Hebrew-speaking children with CI (Adi-Bensaid & Ben-David, 2010; Chin & Finnegan, 2002).

Comparisons between those three types of errors revealed that the complete deletion of the consonant cluster is rare as compared to cluster reduction and cluster simplification. Similar results were found for Hebrew-speaking children with CI up to age 4;5 (Adi-Bensaid & Ben-David, 2010). In contrast, Chin and Finnegan (2002) reported that complete deletion of consonant clusters is absent in English-speaking children with CI at age 9;9. This observation is probably due to the age of the children with CI: our results show that complete deletion of the consonant cluster is absent from age 3;0 onwards, while the children in Chin and Finnegan (2002) were older than nine. Comparisons between the other two types of errors further revealed that, initially, Dutch-speaking children with CI produce their inaccurate consonant clusters as a singleton consonant (cluster reduction), but gradually they produce more consonant clusters, albeit with at least one incorrect consonant (cluster simplification).

Cluster reduction to a target consonant is equally likely than reduction to another singleton consonant, and this remains stable over age. In the literature there is no comparable information about this phenomenon. When children with CI reduce a consonant cluster to one of the target consonants, they adhere to the sonority hypothesis (SH) in approximately 95% of the cases. This is in line with the
observations of English-speaking children with CI (Chin, 2006). However, our results show that this adherence to the SH decreases with age. In addition, when a consonant cluster is reduced to another singleton consonant, almost 90% of the productions merge features of the target consonants, i.e. coalescence. In contrast, Adi-Bensaid and Ben-David (2010) found only a few instances of coalescence in Hebrew-speaking children with CI. Our results show that coalescence is frequent in Dutch-speaking children with CI, which indicates that they have at least some knowledge of the two target consonants (Chin & Dinnsen, 1992). The difference between Hebrew and Dutch might be due to the language-specific input (Yavas, 2013).

6.4.1.2. Age at implant activation
The age at implant activation affects the production of consonant clusters. Even though the range of ages at implant activation was relatively small in the present study (6 – 21 months), earlier implant activation leads to better outcomes. This finding is in agreement with the observation in the literature that children with CI benefit from early implantation as regards speech production (Leigh et al., 2013) and speech production accuracy (Connor et al., 2006; Schauwers et al., 2008; van den Berg, 2012). Children with later implant activation are less accurate than peers with earlier implant activation, but they catch up as shown by a faster increase of their accuracy rates. Our results suggest that later implanted children are less aware of the target consonants and have poorer suprasegmental representations. For instance, they are more likely to completely delete consonant clusters and the decrease of this type of errors is slower in these children. In addition, a consonant cluster is more often reduced than simplified. Thus, only one consonant is more often produced instead of a(n incorrect) cluster. Moreover, children with CI with later implant activation reduce the consonant cluster less frequently to a target consonant. This adds to our hypothesis that they are less aware of the target consonants. A final indication is the likelihood of coalescence, which is lower in children with CI with later implant activation. As this reduction process is less likely in children with CI with later implant activation, it suggests that those children are less aware of the target consonants as compared to children with CI with earlier implant activation.
Development of consonant clusters

6.4.2 Cross-sectional comparisons between children with CI and NH children

The cross-sectional comparisons of NH children and children with CI revealed few differences: children with CI differ from their NH peers only with respect to consonant cluster accuracy, the likelihood of cluster reduction as compared to cluster simplification, and the likelihood of reduction of a consonant cluster to a singleton consonant that is one of the target consonants. The other measures are similar at all ages in both groups of children (i.e. likelihood of occurrence between ages 2;0 and 7;0 and the likelihood of complete deletion, adherence to the SH, and coalescence between ages 2;0 and 4.0).

The accuracy of the clusters of children with CI lags behind their NH peers up to age 4;0, but by age 5;0, they have caught up and accuracy reaches ceiling percentages in both groups. In contrast to our findings, Fulcher et al. (2014) report no differences between English-speaking CI and NH children at ages 3;0 and 4;0. However, our findings are in line with Von Mentzer et al. (2015), who showed that the accuracy of consonant clusters is lower in Swedish-speaking children with CI as compared to their NH peers. According to Von Mentzer et al. (2015), even at 7;4 years the lower accuracy persisted, while our results show a catch up by age 5;0. A possible explanation for these different findings may be found in the children’s task. In the present study consonant clusters were studied in spontaneous speech, whereas Von Mentzer et al. (2015) analysed consonant cluster production in a nonword repetition task that involved the repetition of nonwords after only one auditory presentation, i.e. without any context, referent or lexical representation. Articulatory stability is higher when a nonword is accompanied by a referent, e.g. a picture (Heisler, Goffman & Younger, 2010), meaning that an out-of-context nonword repetition task is more effortful than a contextualized one. In spontaneous speech, words are produced in a referential context, and those words may be assumed to have lexical representations. Therefore, we assume that the nonword repetition task in Von Mentzer et al. (2015) is more demanding than the spontaneous speech recorded in the present study. This may explain why we have found that both groups of children attain similar accuracy scores.
at age 5;0, whereas Von Mentzer et al. (2015) still found significantly lower scores for children with CI at approximately age 7;0.

Children with CI are also found to lag behind their age-matched NH peers with respect to the likelihood of cluster reduction (production of a singleton consonant) as compared to that of cluster simplification (production of 2 consonants, at least one of which is incorrect). At age 2;0, children with CI are more likely than their NH peers to produce a reduced cluster, whereas cluster simplification is more likely in NH children at this age. From age 3;0 onwards, differences between the groups have disappeared. Next, children with CI also lag behind their NH peers concerning the reduction of the consonant cluster to one of the target consonants. At age 2;0, both groups of children reduce a consonant cluster to a singleton that is part of the target cluster in approximately half of their cluster reductions. The switch to more often reducing a consonant cluster to one of the target consonants is already present at age 3;0 in NH children, but is not present until age 4;0 in children with CI.

Children with CI catch up earlier on the number of consonants they produce than on accuracy. The decline of cluster reduction (CV) in favour of the increase of cluster simplification (CCV) concerns a suprasegmental development (i.e. the prosodic structure of consonants clusters), whereas accuracy involves segmental properties (i.e. accurate production of each consonant). Our results show that children with CI catch up earlier on the suprasegmental properties, than on the segmental properties. This is in accordance with the literature on overall phoneme production (thus not only consonant clusters). For instance Willstedt-Svensson, Löfqvist, Almqvist and Sahlén (2004) showed that suprasegmental performance in nonword repetition is better than segmental performance in Swedish-speaking children with CI. Similar outcomes in nonword repetition tasks are found for children with CI acquiring English (Carter, Dillon & Pisoni, 2002; Dillon, Cleary, Pisoni & Carter, 2004) and Spanish (Moreno-Torres & Moruno-Lopez, 2014) and for instance in NH Swedish-speaking children (Sundström, Samuelsson & Lyxell, 2014). Our results apply to spontaneous speech productions of children with CI acquiring Dutch.
Development of consonant clusters

6.4.3 Conclusions

The present paper has traced the longitudinal development of consonant cluster production in the spontaneous speech of Dutch-speaking children with CI. Accuracy is shown to increase with age and, hence, the different types of errors decrease with age. In addition, our results suggest that children with CI benefit from earlier implantation and earlier implant activation.

Cross-sectional comparisons between children with CI and age-matched NH peers revealed some group differences. Children with CI lag behind their age-matched NH peers with regard to accuracy up to age 5;0. In addition, inaccurately produced consonant clusters are more often reduced than simplified in children with CI as compared to NH peers at age 2;0. Furthermore, with respect to cluster reduction, children with CI produce less often a singleton consonant that is part of the target consonants of the consonant clusters as compared to their NH peers up to age 3;0. Nevertheless, children with CI do catch up. They catch up by age 3;0 on the likelihood of cluster reduction and cluster simplification, by 4;0 on cluster reduction to one of the target consonants, and by 5;0, on accuracy. Thus, children with CI seem to catch up earlier on the number of consonants they produce than on accuracy.
Chapter 7
7. Word initial fricative production in children with cochlear implants and their peers with normal hearing matched on lexicon size


ABSTRACT

Introduction: Fricative production is reported to lag behind in children with cochlear implants (CI) as compared to children with normal hearing (NH) in other languages (e.g. English), but not yet for Dutch. Studies have made comparisons between children with CI and age-matched children with NH. However, phonological development is more closely related to lexicon size than to chronological age. Therefore, we also compare children with equal lexicon sizes.

Method: Word-initial fricative development of 10 children with CI and 30 children with NH was compared up to 30 months of age, both on chronological age and lexicon size (i.e. lexicon size).

Results: Fricative production in children with CI is delayed as compared to that of age-matched peers with NH. However, the differences between both groups disappeared when they were matched on lexicon size. Thus, the phonological development in children with CI is similar to that of their peers with NH with equal lexicon sizes.

Keywords: Phonemic accuracy; children with CI; word initial fricatives; chronological age; lexicon size

7.1 Introduction

The present study aims to compare the early development of word-initial fricatives in Dutch-speaking congenitally deaf children with cochlear implants (CI) with that of children with normal hearing (NH). Children with CI have a history of auditory
deprivation. After cochlear implantation their speech perception improves (Calmels, Saliba, Wanna, Coehard, Fillauw, Deguine & Fraysse, 2004; Leigh, Detmman, Dowell & Briggs, 2013; Liu, Liu, Kirk, Zhang, Ge, Zheng, Liu & Ni, 2015; Tyler, Fryauf-Bertsch, Kelsay, Gantz, Woodworth & Parkinson, 1997), which has been shown to have a beneficial effect on their speech production as well (Blamey, Barry, Bow, Sarant, Paatsch & Wales, 2001; Bouchard, Le Normand & Cohen, 2007; Eriks-Brophy, Gibson & Tucker, 2013; Faes, Gillis & Gillis, 2016; Spencer & Guo, 2013). Despite the positive outcomes of cochlear implantation, the auditory information provided by the cochlear implant is still degraded as compared to the auditory information available in normal hearing, especially in the higher frequency regions (Drennan & Rubinstein, 2008). In this respect, fricatives appear to be particularly vulnerable: ‘fricatives are difficult to perceive for people with high-frequency hearing loss due to difficulty extracting high-frequency acoustic cues’ (Van Lierde, Vinck, Baudonck, De Vel & Dhooge, 2005). Consequently the acquisition of fricatives may be especially challenging for children with CI and may therefore be a more protracted process in comparison with children with normal hearing. Finally, Sotto, Redle, Bandaranayake, Neils-Strunjas and Creaghead (2014) have recently reported that the presence of a single fricative at 18 months of age is related to better expressive language development in vocabulary, sentence length and grammar. The presence and frequency of fricatives is thus important as it is related to language development.

7.1.1 Measures of comparison

In the literature, the language development of children with CI has been compared to that of children with NH in various ways: chronological age (Caselli, Rinaldi, Varuzza, Giuliani & Burdo, 2012; Eriks-Brophy et al., 2013; Ertmer, Kloiber, Jung, Kirleis & Bradford, 2012; Guo, Spencer & Tomblin, 2013; Nicholas & Geers, 2007; Salas-Provance, Spencer, Nicholas & Tobey, 2013; Schramm, Bohnert & Keilmann, 2010; Von Mentzer, Lyxell, Sahlén, Dahlström, Lindgren, Ors, Kallioinen, Engström & Uhlén, 2015), hearing age (Caselli et al., 2012; Ertmer & Goffman, 2011; Schramm et al., 2010), and language internal measures (MLU: Szagun, 2001; Szagun, 2002; Szagun, 2004; Expressive vocabulary: Warner-Czyz, 2005) have been used as a basis for comparison. Even though there is a growing tendency to match children with CI
and children with NH on more than one variable (e.g. chronological age and hearing age: Fagan & Pisoni, 2010; chronological age and vocabulary size: Lund & Schuele, 2014; Walker & McGregor, 2013), children with NH and children with CI are often matched on chronological age in order to assess whether those with CI’s linguistic functioning is age appropriate (Caselli et al., 2012; Eriks-Brophy et al., 2013; Ertmer et al., 2012; Guo et al., 2013; Nicholas & Geers, 2007; Salas-Provance et al., 2013; Schramm et al., 2010; Von Mentzer et al., 2015). Chronological age is indeed an intuitively appealing measure: a critical question is whether children with CI eventually attain speech and language skills comparable to their normally hearing age mates.

There are numerous factors which affect language development of children with CI (Boons, Brokx, Dhooge, Frijns, Peeraer, Vermeulen, Wouters & Van Wieringen, 2012; Szagun & Schramm, 2016). CI-specific factors include the age at implantation (Colletti, Carner, Miorelli, Guida, Colletti & Fiorino, 2005; Connor, Craig, Raudenbush, Heavner & Zwolan, 2006; Connor, Hieber, Arts & Zwolan, 2000; Levine, Stother-Garcia, Golinkhoff & Hirsh-Pasek, 2016; Schorr, Fox, van Wassenhove & Knudsen, 2005), the residual hearing thresholds before implantation (Houston, Stewart, Moberly, Hollich & Miyamoto, 2012; Niparko, Tobey, Thal, Eisenberg, Wang, Quittner & Fink, 2010; Szagun & Stumper, 2012), the time of implant use (Blamey et al., 2001; Eriks-Brophy et al., 2013; Schauwers, 2006), etc. In general, it can be assumed that earlier implantation, better hearing thresholds before implantation and a longer CI use are beneficial for language development in children with CI. The participants in the present study were young implanted children, i.e. before the age of two. But given children with CI’s initial auditory deprivation and later (better) access to speech sounds, it is still to be expected that their onset of spoken language development is delayed as compared to children with NH (Svirsky, Robbins, Kirk, Pisoni & Miyamoto, 2000). Therefore, it seems reasonable to assume that chronological age is a good basis for comparison. Moreover, age-matched children show a large amount of inter-subject variability (Leonard, Newhoff & Mesalam, 1980; Vihman, Ferguson & Elbert, 1986), which is even more pronounced in children with CI (Duchesne, Sutton & Bergeron, 2009; Svirsky et al., 2000). Thus,
Word initial fricative production

it looks as if that children of the same chronological age have highly variable language levels. As will be shown later, variability is lower when children are matched on lexicon size. Therefore, chronological age may not be the most optimal standard of comparison.

In order to avoid the methodological issues of different onsets of hearing, hearing age is another time-based measure of comparison. In this respect children with the same amount of speech and language experience are compared. In children with NH, hearing age corresponds to their chronological age, but in children with CI, hearing age refers to the length of device use. Length of device use appears to be important for the language development of children with CI. For instance Szagun and Stumper (2012) have shown that hearing age is a significant predictor of early implanted children’s vocabulary growth and mean length of utterance (MLU). However, using hearing age as a measure of comparison is also problematic. Consider a child whose device is activated at 19 months of age. At 24 months of age, this child has a hearing age of 5 months. If this child is compared to children with NH matched on hearing age, she is matched with five-month-old peers with NH. However, the cognitive, physical and motor development, the articulatory control differ tremendously in five-month-old and twenty-four-month-old infants (Ertmer, Young & Nathani, 2007; MacNeilage, Davis, Kinney & Matyear, 2000; Snow & Ertmer, 2009). In other words, comparing children with CI matched for hearing age may be legitimate for some purposes, but comparing children with CI and those with NH on their hearing age skews the comparisons and is therefore not ideal.

Language-intrinsic measures are sometimes used as an alternative for time-based measures. The idea is to match children on “language age” and to compare particular speech and language abilities relative to that yardstick. Mean length of utterance (MLU) is sometimes used as a proxy for “language age”. Another candidate is lexicon size (operationalized as a cumulative vocabulary count). Several studies have shown that lexical and phonological development are commensurate (Santos & Sosa, 2015; Smith, McGregor & Demille, 2006; Sosa & Stoel-Gammon, 2012; Stoel-Gammon, 2011; Van den Berg, 2012). For instance, Sosa and Stoel-Gammon (2006) found that the development of intraword variability is similar in children with NH on the basis of
the same lexicon size, but not in children with the same chronological age. In a similar
vein, the frequency of syllable types in the productions of Dutch-speaking children
with NH is correlated more readily with lexicon size than with chronological age (Van
den Berg, 2012).

A similar lexical-phonological relation is found in children with CI. Vocabulary
size, and not chronological age, predicts fricative acquisition in children with CI
(Reidy, Beckman, Litovsky & Edwards, 2015). In addition, larger vocabulary sizes are
related to a more accurate sibilant fricative production in English-speaking children
with NH (Nicholson, Munson, Reidy & Edwards, 2015) and children with CI (Reidy
et al., 2015). Thus, lexicon size rather than chronological age predicts phonological
development in children with NH and children with CI. Consequently, matching
children with NH and children with CI on lexicon size, they can be expected to have
reached a similar point in their phonological development as well. In the present
study, the longitudinal development of word initial (WI) fricatives will be analysed.
We expect the accuracy and the error patterns in children’s usage of WI fricatives to
be similar in children with normal hearing and children with CI matched on lexicon
size.

The phonological skills of children with CI are mostly compared to those of age-
matched peers with NH, and not to those of peers with NH matched on lexicon size. 
Even though there seems to be a clear link between lexicon size and phonological
development, most studies on phonological development have ignored this
relationship. Moreover, the few studies that have matched children with NH and
children with CI on lexicon size, have found contrasting results. For Dutch, Van den
Berg (2012) found that the accuracy of word, syllable, segment productions were
similar in children with CI and peers with NH matched on lexicon size rather than
chronological age. But for English, Warner-Czyz (2005) found that segmental
accuracy was higher in children with NH than in children with CI, even when matched
on lexicon size. Warner-Czyz and Davis (2008) corroborated this finding: children
with NH outperformed children with CI at word onset in overall consonant accuracy,
matched on lexicon size. These results should be interpreted with some caution due to
the small sample size (4 children in each group) and the relatively brief period studied
Word initial fricative production

(6 months) (Warner-Czyz, 2005; Warner-Czyz & Davis, 2008). The present study matches both groups of children on lexicon size and studies them over a period of on average 12 months.

7.1.2 Fricative production of children with CI: comparisons with age-mates with NH

As opposed to stops, nasals and glides, fricatives are mostly lacking from children with NH’s first words (Dutch: Beers, 1995; Fikkert & Altvater-Mackensen, 2013; English: Leonard et al., 1980; Stoel-Gammon, 1985; Van Severen, 2012). Similarly to children with NH, fricatives are acquired after stops, nasals and glides in children with CI (Dutch: Schauwers, 2006; English: Serry & Blamey, 1999; Spencer & Guo, 2013; Wiggin, Sedey, Awad, Bogle & Yoshinaga-Itano, 2013). There have been no comparisons between the frequency of fricatives of children with NH and children with CI in Dutch, but they have been studied in other languages (Salas-Provance et al., 2013). Age-matched comparisons show that all fricatives have appeared in English-speaking children with NH’s segment inventories at 42 months of age, while in English-speaking children with CI only /f/, /s/, /z/ and /ʃ/ have been attested by that same age (Salas-Provance et al., 2013). Even though precise frequency differences have not been reported, it seems that there is less fricative diversity in children with CI.

With respect to fricative accuracy, Salas-Provance et al. (2013) reported that only /s/ had an accuracy rate of 50% in English-speaking children with CI and the other fricatives were produced less accurately. In children with NH, both /s/ and /ʃ/ had an accuracy rate of 50% by the age of three and a half. This suggests that fricative accuracy is significantly lower in children with CI as compared to age-matched children with NH. It should, however, be noted that the sample size was small (5 children with NH and 5 children with CI) and that there was a considerable amount of individual variation between the children with CI in Salas-Provance et al. (2013).

Inaccurate production of fricatives results either in fricative deletion or fricative substitution. Overall, substitutions are more common in Spanish-speaking children with CI as compared to peers with NH, but those children were matched at 24 months
of hearing age (Moreno-Torres & Moruno-Lopez, 2014). For English, Spencer and Guo (2013) have shown that in WI position, substitutions are more likely than deletions in children with CI between 12 and 48 months after implantation. The incidence of fricative deletions and substitutions was not quantified and no comparison of both groups of children on chronological age was provided.

Fricatives are often substituted by stops: fricative stopping. Stopping is frequently reported in both children with NH (Dutch: Altvater-Mackensen & Fikkert, 2010; English: Dodd, Holm, Hua & Crosbie, 2003; Spanish: Macken, 1978) and children with CI (Dutch: Baudonck, Dhooge, D’haeseleer & Van Lierde, 2010; French: Bouchard et al., 2007; English: Chin, 2003; Flipsen & Parker, 2008). Apart from fricative stopping, fricative-fricative substitutions are shown to be as common as stopping substitutions in English (Holliday, Reidy, Beckman & Edwards, 2015; Li, Edwards & Beckman, 2009). Neither the likelihood of fricative stopping nor that of fricative-fricative substitution has been compared between both groups of children in the literature on Dutch-speaking children thus far. This is one of the main aims of the present study.

While in the literature fricative development of children with CI is mostly compared to that of age-mates with NH, the present study considers the matching of both groups of children relative to their chronological age as well as relative to their lexicon size. Moreover, most studies on overall phonological development in children with CI compare those children to age-matched children with NH, even though studies suggest a strong link between phonological abilities and lexicon size. The present study aims to compare WI fricative development of children with CI and children with NH on two measures: chronological age and lexicon size. The outcomes of these comparisons are expected to differ. When children with CI are matched with children with NH on chronological age, children with CI are expected to lag behind their age-mates with NH. In contrast, when both groups of children are matched on lexicon size, we expect different results. In children with NH, it has been shown that phonological development and lexical development are commensurate. With respect to children with CI, this relationship has only occasionally been studied and the results were contradicting. Thus, it remains to be seen whether there also is a similar strong
relationship between lexicon size and phonological development in children with CI. If so, similar levels of WI fricative development are expected in children with CI and children with NH when matched on lexicon size. WI fricative development is studied longitudinally, whereas most studies have chosen a single point in development is considered. Four different aspects are quantified: (1) the frequency of use, (2) the accuracy of WI fricatives, (3) the error patterns (deletion or substitution), and (4) the type of substitutions (stopping and fricative-fricative substitution).

7.2 Method

7.2.1 Participants

A longitudinal approach was taken in order to compare two groups of Dutch-speaking children: children with CI and a control group of children with NH. All children were from mid-to-high socio-economic background living in Flanders, i.e. the northern part of Belgium. The parents were normally hearing monolingual speakers of Dutch.

The corpus of children with CI consisted of ten children (S1 – S10) with a congenital severe-to-profound hearing loss. The children were followed monthly from the moment of implant activation up to 2;6 years after activation. The control group consisted of the complete NH longitudinal corpus (S11 – S40), followed longitudinally and monthly between ages 0;6 and 2;0. Here, the data will be analyzed from the appearance of the first spoken words (CI: median = 1;6, range 1;3 – 1;11, NH: median = 1;2, range 1;2 – 1;5). More information can be found in chapter 2.

Children with CI are matched with children with NH in two different ways. First, they are compared on chronological age. In other words, the monthly longitudinal data of both groups of children are matched throughout the entire sample: the speech of children with the same chronological age is analyzed. In these analyses chronological age is a predicting variable. Secondly, children with CI and children with NH are matched on lexicon size. Thus, children with CI are compared to children with NH with equal lexicon sizes. Lexicon size is operationalized as the child’s cumulative vocabulary at each consecutive data point. Cumulative vocabulary is computed automatically as follows (Huttenlocher, Haight, Bryk, Seltzer & Lyons, 1991): in the first transcription, a list is made of all word types (i.e. distinct word forms) which are
produced by the child. The cumulative vocabulary at this point is the number of word types in this list. In the second transcription, each word type is compared to the first list and each new word type is added to this list. The cumulative vocabulary of the second transcription equals the number of word types in this updated list. This is an iterative procedure. The cumulative vocabulary size was limited to 300 word types.

### 7.2.2 Data collection and transcription

The corpus used in this study is part of the CLiPS Child Language Corpus (CCLC), which consists of monthly video recordings of spontaneous, unstructured interactions between the child and the (primary) caregiver at the child’s home. Only those speech samples, which contained lexical items, were analyzed. More information on the data collection, data transcription, data annotation procedures and reliability procedures can be found in chapter 2.

After the transcription process, a list of all WI singleton consonants in children’s productions with their corresponding adult target was retrieved from the transcriptions. A total of 37,368 word productions, i.e. word tokens, with initial singleton onsets were available (CI: 10,698 word tokens, NH: 26,670 word tokens), of which 5,045 word tokens with target fricatives (CI: 1,620 target fricatives, NH: 3,425 target fricatives). The children produced 2,094 word tokens with WI fricatives (CI: 642 fricatives, NH: 1,452 fricatives), pointing out that other target fricatives were deleted in children’s productions.

Agreement on phonemic transcriptions of fricatives was computed on approximately 10% of the corpus (Cucchiarini, 1996). Two reliability procedures were used: (1) intrarater reliability, i.e. retranscription by the same person after at least 3 months, and (2) interrater reliability, i.e. retranscription by one of the other people who collected the CCLC. Kappa scores were 0.75 for interrater reliability and 0.90 for intrarater reliability. These scores can be interpreted as a substantial agreement and an almost perfect agreement respectively (Landis & Koch, 1977).
Word initial fricative production

7.2.3 Dutch consonant inventory

The phoneme inventory of Dutch consists of 20 consonants of which there are nine WI fricatives, /f, v, s, z, x, ɣ, h, ʃ and ʒ/ (Booij, 1995). Table 1 gives an overview of the Dutch phonemic consonant inventory.

<table>
<thead>
<tr>
<th></th>
<th>Labial</th>
<th>Coronal</th>
<th>Dorsal</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stop</td>
<td>p, b</td>
<td>t, d</td>
<td>k, (g)</td>
<td></td>
</tr>
<tr>
<td>Fricative</td>
<td>f, v</td>
<td>s, z</td>
<td>ɣ, ʒ</td>
<td>h</td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td>n</td>
<td>ŋ</td>
<td></td>
</tr>
<tr>
<td>Liquid</td>
<td>l, r</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glide</td>
<td>w</td>
<td>j</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Phonemes in parenthesis only occur as a result of assimilation (e.g., /ʃ/ as an assimilation of /s/ and /j/ in diminutives such as musje (mus-DIM, /meʃə/, “sparrow”) and/or in loanwords such as the /ʒ/ in garage (/ɣɑraʒə/).

7.2.4 Data analyses

WI fricative development is analysed at four different levels. (1) At the first level the occurrence of WI fricatives is considered. For this purpose, all WI singleton consonants in children’s actual word productions were analysed, leaving out empty onsets. The likelihood of fricatives as WI consonants was compared to all other WI consonants. Thus, consider a sample of five WI singleton child consonant productions /b/, /p/, /f/, /b/ and /z/ (for instance for the Dutch target word productions /buk/ (boek, “book”), /pus/ (poes, “cat”), /fits/ (fiets, “bike”), /bal/ (bal, “ball”) and /ze/ (zee, “sea”)). The likelihood of WI fricatives in this sample is 2/5. (2) At a second level, the accuracy of WI fricatives was examined. For this analysis, the adult target words with a WI fricative were selected. Each target word was compared to the child’s actual production, and this comparison of the adult target and the child’s rendition yielded a score “accurate” (e.g. /f/ produced as /f/) or “inaccurate” (e.g. /f/ produced as /p/ or deleted). Thus, the likelihood of correctly produced target fricatives was examined and
compared to inaccurate productions. (3) At a third level, the child’s inaccurate productions were further analyzed. If the child did not render the WI fricative in the target word accurately, the likelihood of WI fricative deletions (e.g. /f/ was not produced, resulting in an empty onset) was compared to that of WI fricative substitutions (/f/ produced as another phoneme, e.g. /p/ or /s/). (4) Finally, at a fourth level, the likelihood of the different substitutions processes was analyzed. As our results point out that only fricative-stop substitution (e.g. /f/ produced as /p/) and fricative-fricative substitution (e.g. /f/ produced as /s/) were common in both groups of children (see further), only those two aspects of fricative substitution will be discussed. Both fricative stopping and fricative-fricative substitution were evaluated as compared to all other substitution patterns. In other words, when the child substitutes a WI fricative in an adult target, the likelihood of fricative stopping and that of fricative-fricative substitution was computed.

7.2.5 Statistical analyses

The dataset exhibits a hierarchical structure: at the lowest level individual child productions are nested within particular ages (observation sessions), which are in turn nested in individual children. Given the hierarchical structure of the data, multilevel models were selected for the statistical analyses. Statistical analyses were performed in R (R Core Team, 2013) by means of logistic regressions in a multilevel model. Multilevel models consist of two parts: a random part and a fixed part. The random part of the multilevel model considers the nesting of variables in the data (Baayen, 2008; Woltman, Feldstain, MacKay & Rocchi, 2012). We allowed random intercepts and slopes to model variation between children and ages. The fixed part represents the predicting variables.

The models were constructed in a stepwise procedure. First, two predicting variables were added as fixed effects (Model 1). The first predicting variable was Hearing status (NH vs. CI). The second one depended of the matching of groups: when matching NH and CI children on chronological age, the second predicting variable was age in months (Age), centered at 24 months of age. When matching on lexicon size, the second predicting variable was cumulative vocabulary (CumulativeVoc), centered at a cumulative vocabulary of 150 word types. Secondly, a
quadratic effect of Age/CumulativeVoc was included (Model 2). By means of an ANOVA, the model fit of Model 1 and the fit of Model 2 were tested. If the fit of Model 2 was significantly better, the quadratic effect of Age/CumulativeVoc as a predicting variable was maintained. As will be shown in the Result section, there were no quadratic Age/CumulativeVoc effects in any of the analyses. Thirdly, an interaction between Hearing status and Age/CumulativeVoc was included in the model (Model 3). Again, the model fits were compared by means of an ANOVA. If the model fit of Model 3 was significantly better, the interaction was maintained in the model. Similarly to the quadratic Age/CumulativeVoc effects, the interaction effects between Hearing status and Age/CumulativeVoc did not improve the model fits in any of the analyses. The best fitting models (Model1) are reported in the results section.

In R, the estimates and standard errors (SE) of logistic regressions are computed in logits. Logits can be converted to probabilities in two steps. First, logits are converted to odds using an exponential function (step 1 in equation (A)). Second, the odds are converted to probabilities using the formula in step 2 of equation (A). For instance, logit = 0 refers to a likelihood of 50%.

\begin{align*}
\text{(A)} \quad \text{Step 1:} \quad \text{odds} &= e^{\text{logit}} \\
\text{Step 2} \quad p &= \frac{\text{odds}}{1-\text{odds}}
\end{align*}

7.3 Results

7.3.1 Descriptive statistics on the substitution patterns

Figure 1 summarizes the likelihood of each substitution process in logits for children with NH and children with CI, relative to the children’s chronological age (age in months) and relative to their “lexical age” (cumulative vocabulary). This graph shows that two substitution processes are well above a likelihood of 10%, i.e. fricative-stop substitution (i.e. fricative stopping) and fricative-fricative substitution. Therefore, only those two substitution processes are further analyzed in the next two paragraphs.
7.3.2 Comparisons on chronological age

Table 2 shows the fixed effect results of all the analyses of WI fricative frequency in logits. In the left panes of Figures 1 – 4, the results are plotted. The shaded areas on the figures indicate the confidence interval. The figures are expressed in logits, but the results will be discussed in terms of likelihoods. We present the four levels of analyses consecutively: (1) incidence of WI fricatives, (2) accuracy of WI fricatives, (3) WI fricative deletion versus substitution, and (4) WI fricative substitution, i.e. fricative stopping and fricative-fricative substitution.

First, Table 2 and Figure 2 display the likelihood of WI fricatives as compared to other WI singleton consonants in children’s actual productions. Figure 2 shows that the incidence of fricatives is low, but increases with age. In addition, Figure 2 suggests
that the incidence of WI fricatives is lower in children with CI than in children with NH, but the development with age seems similar. Inferential statistical analyses (Table 2) reveal that the likelihood of WI fricatives is significantly lower as compared to all other consonant manners at intercept (p<0.001). There is a significant increase with Age (p<0.001), showing that the incidence of WI fricatives increases as children get older (Figure 2). In addition, the effect of Hearing status is significant (p<0.05), indicating that the likelihood of WI fricatives differs significantly in children with NH and children with CI. At intercept, 5.42% of the WI consonants are fricatives in children with NH, whereas this is only 1.42% in children with CI. No interaction between Age and Hearing status is found, and it is therefore not included in the best fitting model reported in Table 2. Nevertheless, the lack of an interaction effect shows (a) that the development with age is similar in both groups of children and (b) that children with CI are not catching up in the period studied: their use of WI fricatives remains lower in comparison with NH children.

Secondly Table 2 and Figure 3 present the fixed effect results for WI fricative accuracy as compared to inaccurate WI fricatives. Figure 3 indicates that the incidence of an accurate WI fricative is relatively low but seems to increase with age. In addition, children with CI seem to produce their WI fricatives less accurately and their development with age is similar to that of children with NH. Inferential statistics (Table 2) indicate that the likelihood of accurate WI fricatives is significantly lower than that of inaccurate WI fricatives (p<0.001): it can be derived from the negative logit value of the intercept. No significant increase with Age is found (p>0.05). However, there is a significant effect of Hearing status (p<0.05): the likelihood of an accurate WI fricative is higher in children with NH as compared to children with CI (Figure 3): at intercept, the likelihood of an accurate WI fricative is 5.73% in children with CI and 18.09% in children with NH.

Thirdly, inaccurate fricatives are either deletions or substitutions. The likelihood of WI fricative deletion as compared to that of WI fricative substitution is shown in Table 2 as well as in Figure 4. Figure 4 indicates that the incidence of WI fricative deletion is not very different from logit 0, i.e. a likelihood of 50%. Nevertheless, there seems to be a decrease with age. Figure 4 also suggests that children with CI more
often delete a WI fricative than their peers with NH, but the decrease with age is similar as compared to peers with NH. Inferential statistics (Table 2) show that the likelihood of fricative deletion is not significantly higher than the likelihood of fricative substitution in children with CI (p>0.05): at intercept 59.51% of the inaccurate fricatives are deleted in children with CI. Even though the likelihood of deletions is lower in children with NH (47.38%), the effect of Hearing status is not statistically significant (p>0.05), meaning that there are no statistically significant differences between both groups of children. Furthermore, there is a decrease of fricative deletion with Age, but this effect is not statistically significant (p>0.05).

Finally, Table 2 and Figure 1 display the likelihood of fricative stopping and that of fricative-fricative substitution as compared to all other possible substitution patterns. Figure 1 suggests that the WI fricatives are more often substituted by a stop in children with CI as compared to children with NH and that there is a decrease with age. Statistical analysis (Table 2) shows that there is a significant difference between both groups of children (Hearing status, p<0.05): the likelihood of fricative stopping is 62.25% in children with CI, but only 34.89% in children with NH at intercept. Even though there seems to be a decreasing effect of Age, this effect is only marginally significant (p=0.055). With respect to fricative-fricative substitution, Table 2 shows that the likelihood of fricative-fricative substitutions is 15.23% at the intercept. In addition, the effect of Hearing status is not significant (p>0.05), indicating that the likelihood of fricative-fricative substitution is similar in children with CI and age-matched children with NH. There is no significant effect of Age (p>0.05), suggesting that the likelihood of fricative-fricative substitution remains stable with age.

As can be inferred from Table 2, the quadratic Age effects and the interactions between Age and Hearing status did not improve the models and are therefore left out. The lack of significant interaction effects between Age and Hearing status indicates that the age effects for each predicting variable are not significantly different for both groups of children. In other words, children with CI do not seem to catch up with their age-matched peers with NH by 30 months of age. This is also clear from the figures: the lines of both groups of children are approximately parallel. Nevertheless, this also
Word initial fricative production shows that the development over time is similar in both groups of children, but occurs later in children with CI.

Table 2. Fixed effect results of the comparisons on chronological age

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>Frequency</th>
<th>Accuracy</th>
<th>Deletion</th>
<th>Stopping</th>
<th>Fricative-fricative substitution</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-4.241</td>
<td>-2.801</td>
<td>0.385</td>
<td>0.500</td>
<td>-1.717</td>
</tr>
<tr>
<td></td>
<td>(0.389)**</td>
<td>(0.522)**</td>
<td>(0.390)</td>
<td>(0.423)</td>
<td>(0.567)**</td>
</tr>
<tr>
<td>Age</td>
<td>0.133</td>
<td>0.031</td>
<td>-0.043</td>
<td>-0.113</td>
<td>0.088</td>
</tr>
<tr>
<td></td>
<td>(0.028)**</td>
<td>(0.051)</td>
<td>(0.051)</td>
<td>(0.059)</td>
<td>(0.083)</td>
</tr>
<tr>
<td>Hearing status NH</td>
<td>1.382</td>
<td>1.291</td>
<td>-0.490</td>
<td>-1.124</td>
<td>0.532</td>
</tr>
<tr>
<td></td>
<td>(0.429)**</td>
<td>(0.543)*</td>
<td>(0.20)</td>
<td>(0.493)*</td>
<td>(0.617)</td>
</tr>
</tbody>
</table>

p≤0.05*, p≤0.01**, p≤0.001***  
CI is the reference category

Figure 2. The likelihood of WI fricatives (in logits) as a function of chronological age and lexical age in children with NH and children with CI (predicted values)
Figure 3. The likelihood of accurate WI fricatives (in logits) as a function of chronological age and lexical age in children with NH and children with CI (predicted values)

Figure 4. The likelihood of WI fricative deletion (in logits) as a function of chronological age and lexical age in children with NH and children with CI (predicted values)
Word initial fricative production

7.3.3 Comparisons on lexicon size

In the analyses presented in this section, children with CI and children with NH are compared relative to their lexicon size. The fixed effect results are presented in Table 3. In the right panes of Figures 1 - 4, the results are plotted as a function of lexicon size. The shaded areas on the figures represent the confidence interval. The figures are expressed in logits, but the results will be discussed in terms of likelihood.

Table 3 and Figure 2 display the likelihood of WI fricatives as compared to all other word initial consonants. Figure 2 shows that the incidence of WI fricatives is low in both groups of children and that there is no significant increase with increasing lexicon size. In addition, children with CI do not seem to differ from children with NH. Inferential statistical analyses show a significant negative effect of the intercept (Table 3). This indicates that the likelihood of a WI fricative is significantly lower than other WI manners of articulation (p<0.001). There is no significant increase with increasing lexicon size, as the effect of CumulativeVoc is not significant (p>0.05), meaning that the incidence of WI fricatives does not change significantly over the period studied. In addition, there is no significant effect of Hearing status (p>0.05). In other words, the likelihood of WI fricatives is similar in both groups of children.

Table 3 and Figure 3 compare the likelihood of accurate WI fricatives as compared to that of inaccurate WI fricatives. Figure 3 suggests that accurate WI fricatives are infrequent in both groups of children and there seems to be only a slight increase with increasing lexicon size. Moreover, both groups of children seem to attain a similar level of accuracy. Inferential statistical analyses (Table 3) point out that the likelihood of accurate WI fricatives is significantly lower than that of inaccurate WI fricatives at intercept (p<0.001). No significant effect of CumulativeVoc is found (p>0.05), showing that the likelihood of accurately produced WI fricatives does not increase significantly with increasing lexicon size. No significant effect of Hearing status is found (p>0.05), suggesting that the accuracy of WI fricatives is similar in both groups of children when matched on lexicon size (see also Figure 3).

With respect to WI fricative deletion, Table 3 and Figure 4 display the fixed effect results of the likelihood of WI fricative deletion as compared to WI fricative substitution. Figure 4 shows that the likelihood of WI fricative deletion is around logit
0, i.e. a probability of 50%. There seems to be no important change in the incidence with increasing lexicon size and no difference between both groups of children. Inferential statistical analyses (Table 3) show that inaccurate WI fricatives are equally likely to be deleted as to be substituted (p>0.05) at intercept. No significant effect of Hearing status (p>0.05) is found, indicating that the likelihood of deletions and substitutions is similar in both groups of children. No significant effect of CumulativeVoc is found neither (p>0.05), meaning that there is no significant change of WI fricative deletion with increasing lexicon size.

Finally, Table 3 and Figure 1 summarize the fixed effect results concerning fricative substitutions, i.e. fricative stopping and fricative-fricative substitution. The likelihood of fricative stopping is compared to all other substitution patterns. Figure 1 shows that fricative stopping is relatively infrequent in both groups of children and that there is a decrease with increasing lexicon size. In addition, figure 8 suggests that fricative stopping is less frequent in children with CI. Inferential statistics (Table 3) show a significant negative effect of the intercept, which indicates that fricative stopping is significantly less likely than all other substitution processes (p<0.05). The possible substitution patterns (e.g. a fricative substituted by a stop, or by a nasal, etc.) are thus distributed more equally when children are matched on lexical age. Even though there seems to be a decrease with increasing lexicon size, this effect is not statistically significant (p>0.05). In addition, all effects are similar in both groups of children, as no significant effect of Hearing status is found (p>0.05) (see also Figure 1).

With respect to fricative-fricative substitution, Figure 1 suggests that its likelihood differs in both groups of children. Inferential statistics (Table 3) indeed showed a significant effect of Hearing status (p<0.001). At the intercept the likelihood of fricative-fricative substitution is 46.03% in children with CI, but only 20.39% in children with NH. In addition, there is an increase of fricative-fricative substitutions with increasing lexicon size (p<0.05). Thus, fricative-fricative substitutions become more frequent with increasing vocabulary size. This developmental trend is similar in both groups of children, as no interaction between lexicon size and Hearing status is found (and therefore not included in the best fitting model reported in Table 3).
Table 3. Fixed effect results of the comparisons on lexicon size

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>Frequency (SE)</th>
<th>Accuracy (SE)</th>
<th>Deletion (SE)</th>
<th>Stopping (SE)</th>
<th>Fricative-Fricative substitution (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-3.875 (0.358)**</td>
<td>-2.455 (0.415)**</td>
<td>0.276 (0.313)</td>
<td>-0.952 (0.389)*</td>
<td>-0.259 (0.369)</td>
</tr>
<tr>
<td>Cumulative eVoc</td>
<td>-0.010 (0.003)</td>
<td>0.001 (0.001)</td>
<td>0.002 (0.001)</td>
<td>-0.004 (0.002)</td>
<td>0.005</td>
</tr>
<tr>
<td>Hearing status NH</td>
<td>0.489 (0.399)</td>
<td>0.752 (0.466)</td>
<td>-0.276 (0.354)</td>
<td>0.278 (0.474)</td>
<td>-1.203 (0.446)**</td>
</tr>
</tbody>
</table>

p\leq0.05*, p\leq0.01**, p\leq0.001***
CI is the reference category

7.4 Discussion

The present study assessed WI fricative development in Dutch-speaking children with NH and young implanted children with CI. Spontaneous speech samples of both groups of children were compared from word onset up to 24 months of age for children with NH and up to 30 months of age for children with CI. Four aspects of WI fricative development were studied: frequency, accuracy, the incidence of deletions vs. substitutions and that of fricative stopping. Furthermore, children with CI were compared matched on both age and lexicon size.

7.4.1 Comparisons on chronological age

Overall, the incidence of WI fricatives is low in both groups of children, but there is an increase with age. This finding is in line with other studies suggesting that fricatives are lacking from children’s earliest word productions irrespective of their hearing status (Chin & Pisoni, 2000; Ertmer & Goffman, 2011; Salas-Provance et al., 2013; Schauwers, 2006; Serry & Blamey, 1999; Spencer & Guo, 2013; Warner-Czyz, 2005; Warner-Czyz, Davis & Morrison, 2005). Moreover, the incidence of WI fricatives is lower in children with CI than in children with NH: 1.64% vs. 5.42% at the intercept. Thus, the incidence of WI fricatives in Dutch-speaking children with CI differs significantly from that of age-mates with NH. Similar results are found for English-speaking children with CI (Salas-Provance et al., 2013). In addition, the development of fricative use is similar in both groups: WI fricatives remain less
frequent in children with CI throughout the period studied, suggesting that children with CI do not catch up with their hearing age mates, but also that the development is similar in both groups: the incidence of WI fricatives increases steadily. Thus, fricatives are less frequent in children with CI as compared to age-matched children with NH. This is an important observation, as Sotto et al. (2014) have shown that there is a relation between the presence of a fricative at 18 months of age and better expressive language development in vocabulary, sentence length and grammar in children with NH. Our results indicate that, even at 24 months of age, WI fricative probability is significantly lower in children with CI than in children with NH. According to Sotto et al. (2014), better expressive language development should be found in children with NH but not in children with CI. In other words, vocabulary, sentence length and grammar of children with CI will lag behind that of age-matched children with NH, because fricatives are almost absent from their consonant inventory. In chapter 8, this suggestion is tested with respect to sentence length and inflectional development.

The accuracy of WI fricatives is relatively low in both groups of children and remains low in the period studied. However, there is a significant difference between both groups of children: the likelihood of accurate fricative production is 18.09% in children with NH and only 5.73% in children with CI. Furthermore, there is no significant age effect in children with CI. This indicates that children with CI are not catching up on their age-mates with NH in the period studied. However, the lack of an interaction between the variables age and hearing status shows that the development of accuracy with age is similar in both groups of children. Thus, children with CI lag behind their age-mates with NH not only on the incidence of WI fricatives, but also on the phonemic accuracy of WI fricatives. Salas-Provance et al. (2013) found similar results for English-speaking children.

An explanation for the relative low incidence of fricatives and the inaccuracies in the production of fricatives can partly be found in the degraded perception of these segments, which may have consequences for their production. When perception is degraded at fricative frequencies, this may result in greater production difficulties. For instance in older adults with NH, Gluth and Hoole (2015) have shown that even mild
hearing loss affected the distinctness of sibilant fricative production: less accurate perception of /s/ and /ʃ/ caused lower distinctiveness in production. Similarly, the perception in children with CI is affected due to the degraded signal provided by the cochlear implant (Drennan & Rubinstein, 2008). In other words, children with CI perceive the higher frequencies relevant for fricatives less well and this may well be responsible for their less accurate production of these segments. While Gluth and Hoole (2015) found an effect of degraded speech perception on the acoustics of speech production in older adults with NH, this study indicates a similar effect in children with CI. They produce less contrast between spectral peaks of sibilant fricatives as compared to peers with NH so that the distinction between these segments is rendered less clearly (Todd, Edwards & Litovsky, 2011). In a similar vein, Neumeyer, Schile and Hoole (2015) concluded that children with CI shift the frequency of the first spectral moment of /s/ downwards. As a result, /s/ and /ʃ/ productions are less distinct in children with CI. Thus, there is an overlap of the typical noise frequencies of sibilant fricative productions of children with CI, which is not apparent in peers with NH (Liker, Mildner & Sindija, 2007). Neumeyer et al. (2015) hypothesized that the reduced auditory input and feedback in children with CI affects their production of sibilant fricatives. Thus, at an acoustic level, sibilant fricative production of children with CI differs from that of children with NH as a result of degraded perception at the higher frequencies. The present study shows, similarly to Salas-Provance et al. (2013), that the production of fricatives is not only less accurate acoustically, but also phonemically.

Inaccurate production of WI fricatives is represented by a deletion or a substitution of the target segment. Our results suggest that both error types (deletion and substitution) are equally likely in both groups of children and that there is no effect of age. In other words, there is no significant difference in the likelihood of both error patterns in the period studied between the two groups of children nor does the likelihood of both types of inaccuracy (deletion and substitution) develop with age. These results differ from those reported in other studies.

Substitution errors were found to be more frequent in Spanish speaking children with CI as compared to children with NH (Moreno-Torres & Moruno-Lopez, 2014).
However, these findings pertain to the substitution of all types of consonants in all word positions and are thus not restricted to WI fricatives, as is the case here. Moreover, these children were not matched on chronological age but on hearing age. More specifically, the children with NH in Moreno-Torres et al. were 24 months of age, while the children with CI were observed at a hearing age of 24 months, which corresponded to a chronological age of 41 months. This discrepancy between the children’s chronological ages may have had an impact on the results, since developmental constraints on speech production develop considerably in the period studied (Ertmer et al., 2007; MacNeilage et al., 2000; Snow & Ertmer, 2009). Substituting a segment instead of deleting it requires finer motor control, which is obviously more developed at an older chronological age. This may explain the differences between our results and those of Moreno-Torres and Moruno-Lopez (2014): the chronological ages at which our subjects’ speech was analysed, were well below the ones studied by Moreno-Torres et al. (2014).

The results also differ from those of Spencer and Guo (2013). They found that substitutions – regardless of consonant manner – are more common than deletions in English-speaking children with CI between 12 and 48 months of hearing age. The mean corresponding chronological age was between 35 and 67 months of age. In contrast, our results show that both error patterns are equally likely between word onset and 30 months of (chronological) age. The children with CI in the present study have a mean hearing age of 12 months, i.e. 12 months of CI use (SD = 5 months) at a chronological age of 24 months, i.e. the intercept of our inferential statistical analyses. Thus, the age at which our data collection ended corresponds to the age at which Spencer and Guo (2013) started their data collection. As a result, the current study and that of Spencer and Guo (2013) may have tracked two consecutive developmental stages. Up to 30 months of age, deletions and substitutions are equally likely – as indicated by our results –, but gradually substitutions become more likely than deletions – from 35 months of age onwards, according to Spencer and Guo (2013)’s results.

In addition to deletion, WI fricatives are often substituted, mostly by a stop (fricative stopping) or a fricative. Studies have shown that fricative stopping is
common in children with NH as well as children with CI (Altvater-Mackensen & Fikkert, 2015; Baudonck et al., 2010; Bouchard et al., 2007; Chin, 2003; Dodd et al., 2003; Flipsen & Parker, 2008; Macken, 1978). The present study adds to the body of knowledge by indicating that fricative stopping is significantly more likely in children with CI as compared to age-mates with NH: the likelihood of fricative stopping is 62.25% in children with CI and only 34.89% in age-mates with NH. Furthermore, the likelihood of fricative stopping seems to decrease with age and this decrease is similar in both groups of children. Even though fricative stopping is more frequent in children with CI as compared to their peers with NH, the development with age (i.e. a decreasing effect) is similar in both groups of children. Besides fricative stopping, the present study has shown that fricative-fricative substitutions are common in Dutch-speaking children with NH and children with CI, similarly to English-speaking children (Holliday et al., 2015; Li et al., 2009). In contrast to fricative stopping, results point out that the likelihood of fricative-fricative substitutions is similar in both groups of children at the intercept.

Compared to age-mates with NH, the course of development of WI fricative production in children with CI is similar: the incidence of WI fricatives and their accuracy increase with age. But, children with CI lag behind their age-mates with NH: they produce fewer WI fricatives and their productions are less accurate. The similar development with age suggests that children with CI are not readily catching up with their peers with NH in the period studied. However, Dutch-speaking children with CI have been shown to catch up with their age-mates with NH on overall phonemic accuracy (not restricted to WI fricatives) by the age of five (Faes et al., 2016).

### 7.4.2 Comparisons on lexicon size

The results of this study suggest that children with CI lag behind age-matched children with NH with respect to WI fricative production, and more specifically for frequency of use, accuracy and incidence of fricative stopping). However, when children with CI are matched with peers with NH on lexicon size rather than on chronological age, the significant differences between both groups of children disappear: WI fricatives are equally frequent in children with NH and children with CI. In addition, children with CI produce fricatives as accurately as children with NH. Furthermore, the likelihood
of fricative stopping is similar in both groups of children. In contrast, the likelihood of fricative-fricative substitutions is higher in children with CI as compared to children with NH.

The WI fricative development of children with CI is not statistically different from that of their peers with NH matched on lexicon size. This outcome corroborates the findings of Van den Berg (2012) who pointed out that (1) the correlation between syllable type frequency and lexicon size is higher than that of syllable type frequency and chronological age, and (2) the overall production accuracy of words, syllables, and segments of children with CI is similar when compared to children with NH matched on lexicon size, but not when compared to children with NH matched on chronological age. The present study found that fricative accuracy is similar in both groups of children when they are matched on lexicon size. In contrast, Warner-Czyz (2005) and Warner-Czyz and Davis (2008) have shown that overall segmental accuracy remains lower in children with CI than in children with NH when matched on lexicon size. This discrepancy between the latter studies and the present one may possibly be attributed to differences in the ranges of lexicon size studied. But information about the exact ranges of the lexicon size in Warner-Czyz’s studies were not reported, so that further comparisons cannot be made.

7.4.3 Comparison of comparisons
Our results suggest that phonological ability and lexicon size are more closely related to one another than phonological ability and chronological age. Children with NH and children with CI with equal lexicon sizes show comparable levels of phonological ability and speech accuracy. This is in line with the claim that phonological development is more closely related to lexicon size than to chronological age (Santos & Sosa, 2015; Smith et al., 2006; Sosa & Stoel-Gammon, 2006; Sosa & Stoel-Gammon, 2012; Stoel-Gammon, 2011). In the present study, fricative production of children with CI lags behind that of children with NH matched on chronological age, but not that of children with NH matched on lexicon size. Thus, children with CI are similar in their fricative productions when they are compared to peers with NH with comparable cumulative vocabulary sizes. In addition, acoustic studies on sibilant fricative production have shown that sibilant fricative production is best predicted by
Word initial fricative production

vocabulary size (Nicholson et al., 2015; Reidy et al., 2015). Thus, fricative production is more closely related to lexicon size than to chronological age in both groups of children. This observation is explained by the development of phonological representations. Accuracy of phonological representations increases with increasing vocabulary size (Ainsworth, Welbourne & Hesketh, 2016). In other words, the development of phonological representations is related to lexicon size. Accurate phonological representations are indispensable in production. If the phonological representations become more accurate with increasing lexicon size, this also affects phonological ability.

Our findings show a different developmental picture of WI fricatives in NH and CI children’s production, depending on whether the comparison is made on chronological age or lexicon size. It was shown that children with CI lag behind their peers with NH when they are matched on chronological age. However, when both groups of children are matched on lexicon size, children with CI do not lag behind their peers with NH. In other words, when matched on chronological age, children with CI differ significantly from their peers with NH. However, when matched on lexicon size, there are no significant differences between children with NH and children with CI. Other studies have already shown that phonological ability and lexicon size are commensurate in children with NH. The present study has shown that this is also the case in children with CI. In early phonological development, children with CI typically lag behind their age-matched peers with NH, but are at the same level as their peers with NH matched on lexicon size. Children with CI are able to reach similar levels of phonological development as compared to peers with NH matched on lexicon size. In other words, when considering the phonological development of children with CI, they are expected to lag behind their age-matched children with NH, at least in the age bracket studied here. However, their production of WI fricatives shows the same (or, at least, highly similar) characteristics as the WI fricatives of children with NH with a similar lexicon size. Hence, also from a clinical perspective it appears to be of utmost importance to evaluate children with CI’s performance not only relative to their chronological age, but also relative to their lexicon size. Children with CI are expected not to be at an age-appropriate level straight after CI activation,
due to i.a. their early auditory deprivation. But their performance is expected to be comparable to that of children with NH matched on lexicon size. Consequently, while an age related delay is expected, especially in early development, a delay relative to lexicon size seems more troublesome, given the outcome of the present study. Several as yet unanswered questions turn up in this respect. First of all, if a child with CI’s production is deviant from that of children with NH with the same lexicon size, what does the phonological profile of that child look like? How does it deviate from the profile of children with NH? Secondly, how should such differences be interpreted? Should this be considered as a deviant phonological profile or a delayed phonological profile? When should clinicians intervene? These points are open for discussion and further research.
Chapter 8
8 Syntagmatic and paradigmatic development of cochlear implanted children in comparison with peers with normal hearing up to age seven 


ABSTRACT

Objective: Grammatical development has been shown to be delayed in children with a cochlear implant (CI). However, the literature has focussed mainly on one aspect of grammatical development, either morphology or syntax, and on standard tests instead of spontaneous speech. The aim of the present study was to compare grammatical development in the spontaneous speech of Dutch-speaking children with CI and normally hearing peers. Both syntagmatic and paradigmatic development will be assessed and compared with each other.

Method: Nine children with cochlear implants were followed yearly between ages two and seven. There was a cross-sectional control group of 10 normally hearing peers at each age. Syntactic development was measured by means of Mean Length of Utterance (MLU), morphological development by means of Mean Size of Paradigm (MSP). This last measure was relatively new in child language research.

Results: Results show that MLU and MSP of children with cochlear implants lag behind that of their normally hearing peers up to age four and up to age six respectively. By age five, children with CI catch up on MSP and by age seven they catch up on MLU.

Conclusion: Children with CI catch up with their normally hearing peers for both measures of syntax and morphology. However, it is shown that inflection is earlier age-appropriate than sentence length in children with CI. Possible explanations for this difference in developmental pace are discussed.

Keywords: language acquisition; children with CI; MLU; inflectional development
MLU and MSP development

8.1 Introduction

The current study examines the development of grammatical language skills of Dutch-speaking congenitally deaf children who received a cochlear implant (CI) at an early age. They will be compared to peers with normal hearing (NH) up to the age of seven. Early cochlear implantation has been shown to considerably foster language development in congenitally deaf children (Blamey, Barry, Bow, Sarant, Paatsch & Wales, 2001; Nicholas & Geers, 2007). Some children with CI are found to manifest age-appropriate language skills after one to four years of device use, while others still lag behind their peers with NH even after more than four years of device use (Szagun, 2002; Wie, 2010). Consequently, language development in children with CI is subject to large interindividual variation, as only some children with CI seem to catch up with their peers with NH. However progress and acquisition rates are also dependent on the particular linguistic aspect studied. For instance receptive language skills of children with CI are age-appropriate faster than their expressive language skills (Duchesne, Sutton & Bergeron, 2009; Wie, 2010). In addition, children with CI are found to have particular difficulties with syntax and morphology, in contrast to lexical development (Duchesne et al., 2009). In other words, most children with CI are found to catch up with their peers with NH on vocabulary measures, but not on measures of productive morphology and syntax (grammatical aspects of language use).

Language development can be studied in different ways: standard tests can be used to assess children’s grammatical competence or language measures based on spontaneous speech. A frequently used standard test for grammatical development is the Reynell Developmental Language Scale (RDLS). Duchesne et al. (2009) used the RDLS and have shown that more than half of the children with CI had receptive and expressive age-appropriate language skills at the word level after six years of implant use, while less than 50% of the same group of children had receptive and expressive age-appropriate language skills at the sentence level. Even though other standardised tests were used, similar outcomes were found in for instance Caselli, Rinaldi, Varuzza, Giuliani and Burdo (2012); Geers, Nicholas and Sedey (2003); Schorr, Roth and Fox (2008); and Young and Killen (2002). In studies of spontaneous speech, a similar relative developmental pace of lexicon and grammar is found: the delay of children
with CI is shorter for lexical development than grammatical development, expressed as, for instance, mean length of utterance (MLU) (Moreno-Torres & Torres, 2008) and adjectival inflection (Tribushinina, Gillis & De Maeyer, 2013). In contrast to outcomes on standard tests, early implanted children with CI catch up with their peers with NH by approximately age 5;00 when studying their spontaneous speech (Nicholas & Geers, 2007; Tribushinina et al., 2013). Like in standard tests, most studies on spontaneous speech investigated only one aspect of grammatical development. In contrast, the present study does not only focus on spontaneous speech, but also on two specific aspects of grammatical development, viz. syntagmatic and paradigmatic complexity. Furthermore, the development of those two aspects in children with CI is compared with each other.

Children with CI seem to have specific difficulties with grammatical development. Even in children with NH, grammatical development is a slow and gradual process (Le Normand & Moreno-Torres, 2014; Le Normand, Moreno-Torres, Parisse & Dellatolas, 2013). Grammatical development is generally considered to involve syntactic development, i.e. combining words into sentences, and morphological development, i.e. combining morphemes into larger units as in, e.g. inflection, compounding and derivation. Hence grammatical development exhibits a syntagmatic dimension, i.e., how words are ordered in sentences, and a paradigmatic dimension, e.g. the different forms of a particular root or stem. Both dimensions interact as can be seen in congruence: in languages such as English and Dutch a singular subject requires a singular form of the (finite) verb (e.g. the man is working), and a plural subject requires a plural form of the verb (e.g. four men are working). In the present study, grammatical skills are analysed in both children with NH and early implanted children with CI, with a focus on syntagmatic and paradigmatic development, operationalized as Mean Length of Utterance (MLU) and Mean Size of Paradigm (MSP) respectively.

8.1.1 Mean length of utterance (MLU)

MLU is a widely used measure of general language development linked to morphology and syntax and grammar in general. Even though MLU, as presented by Brown (1973), is not a direct measure of syntactic development – for instance, it does
MLU and MSP development

not take correctness of word order into account - it provides an indication of the degree of sentence complexity (Hammer, 2010). Recently, Mimeau, Plourde, Ouellet and Dionne (2015) have shown that MLU is a valid and reliable measure of morphosyntactic complexity up to school ages. When children combine more words into longer sentences, MLU is higher, which indicates at least the knowledge of some syntagms. Therefore, MLU is considered as a measure of syntagmatic development.

MLU can be calculated in several ways: Brown (1973) suggested to divide the number of morphemes, i.e. the smallest meaningful units, words or word parts, by the number of utterances (MLU in morphemes). However, there are strong correlations between MLU in morphemes and words (Flipsen & Kangas, 2014; Hickey, 1991; Parker & Brorson, 2005; Rice, Smolik, Perpich, Thompson, Rytting & Blossom, 2010) and in syllables (Hickey, 1991) are found. More detailed information about MLU computation is given in the method section (2. Method).

MLU increases with age (Blake, Quartaro & Onorati, 1993; Rice et al., 2010) between approximately 1;06 (years;months) and 5;00. MLU is useful to detect language problems in children (Klee & Fitzgerald, 1985). For instance in children with specific language impairment (SLI), MLU is lower in comparison to typically developing peers (Hewitt, Scheffner Hammer, Yont & Tomblin, 2005; Rice et al., 2010). With respect to children with CI, Tobey and Hasenstab (1991) found no increase in MLU after one year of implantation. The mean age at implantation was 6;00 (SD = unknown), while in Blamey et al. (2001), Moreno-Torres and Torres (2008) and Schauwers (2006) found an increase of MLU with longer implant use. Participants in these studies were implanted at younger ages: mean ages at implantation are 3;09 (SD = 1;00), 1;04 (case study) and 1;00 (SD = 0;05) respectively.

Comparisons between children with NH and children with CI can reveal delays in syntagmatic development of children with CI. Table 1 shows the outcomes of some recent studies in various languages. Even though the study design (longitudinal or cross-sectional, number of CI participants) and mean ages at implantation are different across studies, Table 1 suggests that MLU of children with CI is mainly found to lag behind that of peers with NH up to approximately age 8;00. But, Nicholas and Geers
(2007) and Hammer (2010) concluded that early implanted children with CI catch up with their peers with NH by age 4;06 and 8;00 respectively.

Table 1. Literature overview MLU in children with CI and children with NH

<table>
<thead>
<tr>
<th>Authors</th>
<th>Language</th>
<th># children with CI</th>
<th>Design</th>
<th>Mean age at implantation (SD)</th>
<th>Outcome: MLU CI &lt; NH at age&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Do children with CI catch up?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ouellet, Le Normand and Cohen (2001)</td>
<td>French</td>
<td>5</td>
<td>L</td>
<td>3;09 (1;02)</td>
<td>5;02</td>
<td>Not reported</td>
</tr>
<tr>
<td>Szagun (2001) and Szagun (2002)</td>
<td>German</td>
<td>22</td>
<td>L</td>
<td>2;05 (0;08)</td>
<td>5;06</td>
<td>Not reported</td>
</tr>
<tr>
<td>Schauwers (2006)</td>
<td>Dutch</td>
<td>9</td>
<td>L</td>
<td>1;00 (0;05)</td>
<td>2;06</td>
<td>Catch up at 7;00</td>
</tr>
<tr>
<td>Nicholas and Geers (2007)</td>
<td>English</td>
<td>76</td>
<td>L</td>
<td>1;11 (unknown)</td>
<td>3;04</td>
<td>Catch up at 4;06</td>
</tr>
<tr>
<td>Hammer (2010)</td>
<td>Dutch</td>
<td>48</td>
<td>C</td>
<td>1;04 (0;09)</td>
<td>6;00</td>
<td>Catch up at 7;00</td>
</tr>
<tr>
<td>Nittrouer, Caldwell-Tarr, Sansom, Twersky and Lowensthein (2014a)</td>
<td>English</td>
<td>55</td>
<td>PL</td>
<td>1;09 (1;02)</td>
<td>8;04</td>
<td>Not reported</td>
</tr>
<tr>
<td>Nittrouer, Sansom, Low, Rice and Caldwell-Tarr (2014b)</td>
<td>English</td>
<td>55</td>
<td>PL</td>
<td>1;09 (1;02)</td>
<td>7;08</td>
<td>Not reported</td>
</tr>
</tbody>
</table>

<sup>a</sup> L = longitudinal, C = cross-sectional, PL = one data point as part of a longitudinal design

<sup>b</sup> Ages are represented in years;months

The reported delays of children with CI with respect to MLU have been explained by deficits of the short-term phonological working memory in these children (Kronenberger, Pisoni, Henning & Colson, 2013; Pisoni & Cleary, 2004; Pisoni, Kronenberger, Roman & Geers, 2010). For instance Willis and Gathercole (2001) has shown that an effect of phonological working memory capacities on sentence repetition accuracy. Working memory involves short-term storage, rehearsal and
MLU and MSP development

handling of information (Baddeley, 2003). In longer and more complex sentences, more phonological information must be stored and handled. Furthermore, the cognitive load is higher in longer sentences, which reduces the efficiency of the phonological short-term working memory (Pisoni et al., 2010). As children with CI have lower short-term phonological working memory capacities (Kronenberger et al., 2013; Pisoni & Cleary, 2004; Pisoni et al., 2010), their sentence length and complexity will be affected negatively. For instance Charest, Johnston and Small (2015) have shown a decrease in MLU with increasing load of the working memory in children with NH. Similarly, Willis and Gathercole (2001) have shown a decrease in sentence repetition accuracy with increasing sentence length and thus an increase in cognitive load. A similar process is assumed to be present in children with CI.

The present study examines MLU development in 9 early implanted Dutch-speaking children with CI up to age seven and compares those children to age-matched children with NH.

8.1.2 Mean size of paradigm (MSP)

Besides syntagmatic richness, paradigmatic richness is essential to children’s grammatical development. Paradigmatic richness can be expressed by means of MSP, as presented by Xanthos and Gillis (2010). It can be considered as a measure of paradigmatic complexity, and more specifically inflectional diversity. Inflectional diversity gives an indication of the number of different inflected word forms per lemma, i.e. root. When children have more inflected word forms for a lemma, inflectional diversity is bigger. MSP gives the average of the number of inflected word forms over all lemmas (Xanthos & Gillis, 2010). In its simplest form, MSP is calculated by dividing the number of distinct word forms, i.e. number of inflected word forms, by the number of lemmas (Xanthos & Gillis, 2010). The present study examines MSP development in children’s language. More detailed information is given in the method section (8.2.3).

MSP has been used to examine the influence of paradigmatic variation in child-directed speech on children’s speech in different languages (Laaha & Gillis, 2007; Xanthos, Laaha, Gillis, Stephany, Aksu-Koç, Christofidou, Gagarina, Hrzica, Ketrez, Kilani-Schoch, Korecky-Kröll, Kovacevic, Laalo, Palmovic, Pfeiler, Voeikova &
Dressler, 2011). These cross-linguistic studies have revealed that bigger inflectional diversity in the input leads to faster development of inflectional diversity. Paradigmatic richness in child-directed speech is thus positively correlated to development of paradigmatic richness in the child’s speech. Laaha and Gillis (2007) examined total MSP input in child-directed speech of different languages. For Dutch, MSP is 1.07 for nouns and 1.82 for verbs. By way of comparison, MSP of Turkish is 1.91 for nouns and 3.93 for verbs (Laaha & Gillis, 2007). Higher MSP is an indication of a morphologically richer inflecting language. Thus, Dutch is a language with ‘poor’ morphology and Turkish is language with ‘rich’ morphology. In the present study MSP will be used to compare the development of inflectional diversity in Dutch-speaking children with NH and children with CI.

The development of children with CI’s inflection is found to lag behind that of peers with NH. For instance for German, Szagun (2002) has shown that inflectional morphology of children with CI is less advanced compared to MLU-matched children with NH. Case and gender marking of articles and noun plurals are less accurate in children with CI. With respect to plurals, children with NH are found to make errors, but children with CI simply do not mark plurals, and therefore avoid erroneous complex morphology (Szagun, 2002). Likewise, Laaha, Blineder and Gillis (2015) have shown that Dutch and German children with CI produce significantly more singular nouns compared to age-matched peers with NH in an elicitation task of plurals. With respect to articles, the same trend as in nouns is found: children with NH erroneous use, while children with CI frequently omit articles (Szagun, 2002). Whereas children with CI have difficulties with morphology of nouns and articles, no differences with respect to verbal morphology are found (Szagun, 2002). For English, Guo, Spencer and Tomblin (2013) have shown that tense marking is less accurate in children with CI as compared to peers with NH up to five years of implant use. For Dutch Laaha et al. (2015) have found few differences in noun plural marking in 4;06 year-olds in an experimental task. However, similar to Szagun (2002), children with CI did not mark the plural at all and responded with an imitation of the singular prompt. Hammer (2010) found similar results as Guo et al. (2013): children with CI acquiring Dutch are delayed on verb morphology, subject-verb agreement and past
tense marking. Nevertheless, children with CI seem to catch up for nominal and verbal morphology by age seven (Hammer, 2010).

Inflectional development of children with CI is thus poorer as compared to peers with NH, due to reduced auditory speech input and poor perception of already low salient grammatical morphemes (Guo et al., 2013; Svirsky, Stallinfs, Ying, Lento & Leonard, 2002). For instance Svirsky et al. (2002) have shown that the most prominent grammatical markers are acquired before the less prominent ones in children with CI. Grammatical morphemes are less salient in speech perception. As a result of the limitations of the cochlear implant, children with CI are assumed to poorly perceive such low salient morphemes. Nevertheless, the perception of such morphemes is indispensable for their production, as perception is a prerequisite for production (Stoel-Gammon, 2011). A limited perception of low salient items, like grammatical morphemes, is thus assumed to negatively influence their production.

8.2 Method

8.2.1 Participants

A longitudinal design was set up to assess language development of children with CI (S1 – S9): The yearly data (between ages two and seven) of the CI corpus were used (Schauwers, 2006). The NHcross-sectional corpus was used as control group (S41 – S103). This corpus includes 10 two-year-olds (mean = 24.62 months, SD = 1.06 months), 9 three-year-olds (mean = 35.88 months, SD = 1.09 months), 12 four-year-olds (mean = 48.35 months, SD = 1.39 months), 11 five-year-olds (mean = 60.41 months, SD = 1.32), 11 six-year-olds (mean = 72.43 months, SD = 2.80) and 10 seven-years-olds (mean = 83.32 months, SD = 1.74). More detailed information is given in chapter 2.

8.2.2 Data collection and transcription

For each data point, approximately a one-hour video recording of spontaneous speech was made at children’s home. The four youngest groups of children (two- to five-year-olds) were video recorded in an unstructured parent-child interaction, while the video recordings in the two oldest groups of children (six- and seven-year-olds) were semi-
structured: those children were asked to spontaneously tell a story based on large, busy images or a picture book (“Frog, where are you?” (Mayer, 1969). A selection of approximately twenty minutes of each recording was transcribed orthographically and phonemically in CHILDES’ CLAN utility according to the CHAT conventions (MacWhinney, 2000). All word forms were automatically tagged and manually disambiguated using CLAN’s MOR utility. Each word form was lemmatized, morphologically decomposed, and received an appropriate part-of-speech tag.

Interrater reliability was computed on 65% of the data. A second investigator re-transcribed the complete twenty-minute selections orthographically and two different percentages of agreement were calculated: the agreement of the number of words per utterance and the agreement on the words themselves. The agreement on number of words per utterance is of importance with respect to the analyses of utterance length and the agreement on the words identified matters when examining inflectional development. The median agreement of number of words per utterance equalled 90.88% (range: 81.50% - 97.25%). The median agreement on the identified words themselves equalled 81.38% (range: 69.50% - 92.63%).

More detailed information can be found in the method section (chapter 2).

8.2.3 Language measures and statistical analyses

Language development was assessed by means of two language measures. Firstly, we calculate MLU in words in order to give an indication of syntagmatic richness of children’s speech. MLU in words was calculated by dividing the number of words by the number of utterances. Consider for instance the two utterances in example 1:

<table>
<thead>
<tr>
<th>Child utterance 1</th>
<th>ik boek mama</th>
<th>(I book mummy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child utterance 2</td>
<td>ik ook lezen boek</td>
<td>(I also read book)</td>
</tr>
</tbody>
</table>

For each utterance, the number of words in that utterance was counted: 3 words in utterance 1 and 4 words in utterance 2. Next, the total number of utterances was tallied, which equals two in this example. Subsequently, the mean length of utterance (MLU) is the ratio of the total number of words (7) on the total number of utterances (2) and equals 3.5. Therefore, MLU in words actually reflects mean sentence length.
MLU and MSP development

Secondly, paradigmatic richness of children was assessed by measuring MSP of each child. As expressed in equation A, MSP is the ratio of the size of the inflected lexicon $|F|$ - which includes only distinct word forms – to the size of the root lexicon $|L|$, i.e. the lemmatized word forms (Xanthos & Gillis, 2010: 180). Consider a sample of, for instance, $N = 7$ with the following English inflected word forms: am, are, were, book, books, are, car. In this sample, the inflected lexicon $F$ contains 6 distinct word forms (am, are, were, book, books, car) and the root lexicon $L$ consists of 3 lemmas (be, book, car). Consequently, the MSP equals 2.

\[ MSP = \frac{|F|}{|L|} \]

The MSP option used in the present study was the weighted entropy-based MSP in the open sourced software MSPMeter as developed by Gillis (2013), which takes into account the entropy of each paradigm as well as the relative frequency of each individual inflected form (Xanthos & Gillis, 2010). More detailed information of MSP calculation can be found in Xanthos and Gillis (2010).

Statistical analyses were done in the open source software R (R Core Team, 2013) by means of fixed occasion multilevel models (package lme4). Outliers were defined by means of the interquartile rule and were omitted. At each age (2;00, 3;00, 4;00, 5;00, 6;00 and 7;00), the dependent variable was MLU, respectively MSP, and the independent variable was the children’s hearing status (NH or CI). T-tests were used to investigate MLU and MSP in children with NH at consecutive ages. An analysis of the longitudinal data of the children with CI was done in order to examine the growth curve of MLU and MSP. Hearing age, i.e. length of device use, was also entered as independent variable. A random intercept was entered for each child separately in the mutual comparisons of children with NH and children with CI in order to consider the variation between children. In the longitudinal analyses of children with CI, random effects for each child and each hearing age, i.e. time after cochlear implantation, were
added and this resulted in random intercepts and slopes for each child at each age. A significance level of \( p<0.05 \) was set.

### 8.3 Results

#### 8.3.1 Mean length of utterance (MLU)

Table 2 displays syntagmatic richness, expressed in MLU, of children with NH and children with CI at each age. Although the data of the children with NH are cross-sectional, estimates in Table 2 and the growth curve in the top pane of Figure 1 show an increase of MLU with age. MLU values at age 2;00 are 1.6203 (SE = 0.1711) for children with NH and 1.0955 (SE = 0.1241) for children with CI. At age 7;00 MLU has increased to 5.7243 (SE = 0.5276) for both groups of children. Post hoc analyses by means of t-tests indicate that the yearly increase of MLU is significant for children with NH (in each t-test, \( p < 0.0001 \)). For children with CI, longitudinal data were available. Analyses indicated a significant linear increase of MLU up to age 7;00: MLU increases with 0.0754 (SE = 0.0052, \( p < 0.0001 \)) per year. The increase of MLU is plotted in the top pane of Figure 1.

Table 2 shows that the initial difference between children with NH and children with CI disappears at age 7;00. More specifically, our analyses indicate that MLU values of children with CI are significantly lower than MLU values of children with NH from age 2;00 up to age 6;00 (table 2). As table 2 shows, no significant differences between children with NH and children with CI are found at age 7;00 (\( p = 0.4405 \)). Differences between both groups of children are plotted in the top pane of Figure 1.
Table 2. Fixed effect results of syntagmatic richness (MLU)

<table>
<thead>
<tr>
<th>Age (years;months)</th>
<th>Fixed effects</th>
<th>Estimate (SE)</th>
<th>T-value</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>2;00</td>
<td>Intercept</td>
<td>1.0955 (0.1241)</td>
<td>8.8258</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>Hearing status NH</td>
<td>0.5248 (0.1711)</td>
<td>3.0374</td>
<td>0.0022</td>
</tr>
<tr>
<td>3;00</td>
<td>Intercept</td>
<td>2.2406 (0.2899)</td>
<td>7.7292</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>Hearing status NH</td>
<td>1.0548 (0.4100)</td>
<td>2.5730</td>
<td>0.0101</td>
</tr>
<tr>
<td>4;00</td>
<td>Intercept</td>
<td>3.2635 (0.2694)</td>
<td>12.1136</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>Hearing status NH</td>
<td>1.3091 (0.3633)</td>
<td>3.6036</td>
<td>0.0003</td>
</tr>
<tr>
<td>5;00</td>
<td>Intercept</td>
<td>3.7075 (0.3941)</td>
<td>9.4074</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>Hearing status NH</td>
<td>1.0557 (0.5314)</td>
<td>1.9867</td>
<td>0.0470</td>
</tr>
<tr>
<td>6;00</td>
<td>Intercept</td>
<td>4.6032 (0.3839)</td>
<td>11.9908</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>Hearing status NH</td>
<td>1.1359 (0.5276)</td>
<td>2.1527</td>
<td>0.0313</td>
</tr>
<tr>
<td>7;00</td>
<td>Intercept</td>
<td>5.7243 (0.5276)</td>
<td>10.8502</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>Hearing status NH</td>
<td>0.5306 (0.6879)</td>
<td>0.7714</td>
<td>0.4405</td>
</tr>
</tbody>
</table>

Figure 1. Development of MLU and MSP in CI (longitudinal) and NH (cross-sectional)

8.3.2 Mean size of paradigm (MSP)

Table 3 displays paradigmatic richness, expressed in MSP, of children with NH and children with CI at each age. Estimates range between 0.9795 and 1.5227. As the bottom part of Figure 1 indicates, MSP increases over time. Post hoc analyses by
means of t-tests indicate a significant yearly increase of MSP up to age 6;0 for children with NH (in each t-test, $p < 0.0001$). A longitudinal analysis of children with CI showed a linear increase of MSP up to age 7;00: MSP increases with 0.0064 ($SE = 0.0017$, $p = 0.0001$) per year. The development of MSP is plotted in the bottom part of Figure 1.

Next, some effects of hearing status are found, as shown in Table 3. Paradigmatic richness is significantly lower in children with CI compared to children with NH at ages 3;00 ($p < 0.0001$) and 4;00 ($p < 0.0001$). At age 2;00, no difference between both groups of children is found ($p = 0.1057$). This was expected as inflectional development starts about this age. At that age, each lemma is represented by only one word form, and hence MSP is around its bottom value. From age 5;00 onwards differences in MSP values between the two groups of children disappeared (Table 3). The development of both groups of children is also plotted in the bottom part of Figure 1.

Table 3. Fixed effect results of paradigmatic richness (MSP)

<table>
<thead>
<tr>
<th>Age (years;months)</th>
<th>Fixed effects</th>
<th>Estimate (SE)</th>
<th>T-value</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>2;00</td>
<td>Intercept</td>
<td>1.3126 (0.0460)</td>
<td>28.5248</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>Hearing status NH</td>
<td>-0.0912 (0.0564)</td>
<td>-1.6180</td>
<td>0.1057</td>
</tr>
<tr>
<td>3;00</td>
<td>Intercept</td>
<td>0.9795 (0.0805)</td>
<td>12.1736</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>Hearing status NH</td>
<td>0.7537 (0.0053)</td>
<td>143.4638</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>4;00</td>
<td>Intercept</td>
<td>1.3539 (0.0322)</td>
<td>42.0953</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>Hearing status NH</td>
<td>0.1806 (0.0059)</td>
<td>30.4393</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>5;00</td>
<td>Intercept</td>
<td>1.5102 (0.0540)</td>
<td>27.9565</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>Hearing status NH</td>
<td>-0.0054 (0.0725)</td>
<td>-0.0721</td>
<td>0.9401</td>
</tr>
<tr>
<td>6;00</td>
<td>Intercept</td>
<td>1.4561 (0.0486)</td>
<td>29.9821</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>Hearing status NH</td>
<td>0.1224 (0.0687)</td>
<td>1.7818</td>
<td>0.0748</td>
</tr>
<tr>
<td>7;00</td>
<td>Intercept</td>
<td>1.5227 (0.0413)</td>
<td>36.8927</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>Hearing status NH</td>
<td>-0.0219 (0.0538)</td>
<td>-0.4073</td>
<td>0.6838</td>
</tr>
</tbody>
</table>

8.4 Discussion

The present paper studied the grammatical development of Dutch-speaking early implanted children with CI in comparison with age-matched peers with NH. Grammatical development was assessed in two ways. First of all, syntagmatic richness
was measured by means of MLU in words, and secondly, paradigmatic richness was measured by MSP. Longitudinal yearly data of 9 children with CI between 2;0 and 7;0 were available as well as cross-sectional data of approximately 10 peers with NH at each age. Analyses revealed three important findings with respect to syntagmatic and paradigmatic development.

8.4.1 **Syntagmatic development**

The first finding concerns syntagmatic development and more specifically the development of sentence length. Our results revealed that MLU of children with CI lags behind that of children with NH up to age six, but children with CI catch up by age seven. This is in accordance with for instance Nittrouer et al. (2014a); Nittrouer et al. (2014b); Ouellet et al. (2001); Schauwers (2006); Szagun (2001); and Szagun (2002), who have found lower MLU in children with CI as compared to children with NH. However, our results show that children with CI have age-appropriate MLU scores by age seven. This is in line with Hammer (2010); and Nicholas and Geers (2007), who both have found a catch up of children with CI by age seven and age four-and a half respectively.

The delay of syntagmatic development of children with CI can be explained by the poorer phonological short-term working memory skills in children with CI. Syntagmatic development was measured by means of MLU, which reflects sentence length. In the production of a sentence (sentence planning), children need to store phonological information of sequences of words in the short-term phonological memory. However, studies have found a storage deficit in children with CI (Nittrouer, Caldwell-Tarr & Lowensthein, 2013).

Although working memory is not limited to the phonological short-term, it is particularly this part of the working memory, also called the phonological loop, which is related to sequence learning and phonological coding (Pisoni et al., 2010) and thus to the production of (long) sentences. For instance Willis and Gathercole (2001) have shown that higher phonological short-term working memory capacities result in better sentence repetition accuracy in children with NH. Thus, the capacity of the phonological short-term working memory influences sentence production. The phonological short-term working memory is often assessed by forward digit spans in
Chapter 8

children (Pisoni et al., 2010). Forward digit spans comprise the repetition of an increasing number of digits in the order as presented and involve the temporal and linear storage of sequences of items (Montag, AuBuchon, Pisoni & Kronenberger, 2014). Therefore forward digit spans reflect linguistic abilities and they have been shown to be related to for instance sentence intelligibility (Montag et al., 2014) and accuracy (Dillon, Cleary, Pisoni & Carter, 2004). Forward digit span scores are smaller in children with CI as compared to peers with NH (Burkholder & Pisoni, 2003; Cleary, Pisoni & Geers, 2001; Edwards & Anderson, 2014; Kronenberger et al., 2013; Pisoni & Cleary, 2003; Pisoni & Cleary, 2004; Pisoni et al., 2010). Shorter forward digits spans indicate that the phonological short-term working memory is poorer in children with CI. As this part of the working memory influences speech production, it is likely that sentence length lags behind in children with CI.

Furthermore, while phonological short-term working memory is mainly active in speech perception and production, the general executive of working memory is also active (Baddeley, 2003). A greater cognitive load will increase the activity of the general executive and therefore decrease the activity of the phonological short-term working memory (Pisoni et al., 2010). An increase in sentence length enlarges the cognitive load as more information must be stored and handled, as shown by for instance Willis and Gathercole (2001). Consequently, phonological short-term working memory will have fewer resources available, which results in more difficulties with longer sentences. Thus, an increase in cognitive load will reduce sentence length (Charest et al., 2015). In children with CI, longer sentences will reduce the already poorer short-term phonological memory, resulting in poorer performance. For instance Young and Killen (2002) have shown that children with CI are weak in remembering and repeating sentences of increasing length and complexity. Thus, sentence length has an influence on children with CI’s short-term phonological working memory. Long sentences are difficult in children with CI, and this may explain why MLU of children with CI lags behind on MLU of peers with NH. In a similar vein, Valian (1991) proposed that processing limitations affect sentence length in children with NH, as children simply omit, e.g. auxiliaries. When
MLU and MSP development

processing limitations decrease, omission rate will be reduced and utterances will become longer.

8.4.2 Paradigmatic development

The second finding concerns morphology and more specifically the development of inflectional diversity. Analyses have shown lower MSP in children with CI at ages three and four as compared to children with NH. While MSP as used as a measure of inflectional development in children with CI is new in the present study, noun, verb and article inflection have been studied before. For instance Guo et al. (2013); and Hammer (2010) found a delay in verb morphology in children with CI as compared to children with NH. Similarly, children with CI are also delayed in case and number marking of German nouns (Laaha et al., 2015; Szagun, 2002) and German articles (Szagun, 2002). However, the results show that MSP of children with CI is age-appropriate from age five onwards. In a similar vein, Hammer (2010) has shown that children with CI catch up on verbal morphology by age seven.

Thus, paradigmatic development of children with CI is delayed up to age four. This observation is most likely explained by two related factors, i.e. speech perception and lexical organisation. A first explanation is the poorer speech perception in children with CI. Inflected word forms contain suffixes in Dutch. Such suffixes are less salient in the acoustic speech signal. Even in children with NH, grammatical morphemes are more difficult to perceive, as they are unstressed and therefore less salient. Thus, children with NH better perceive highly salient items. Furthermore, perception is indispensable for production (Le Normand et al., 2013). In other words, perception of grammatical morphemes, even though they are low salient, is a prerequisite for their production. However, speech perception in children with CI is negatively affected by their history of auditory deprivation before implantation and limitations of the cochlear implant after implantation. For instance Bouton, Serniclaes, Bertoncini and Colé (2012) found poorer categorical precision, i.e. accuracy in feature discrimination and identification, in children with CI as compared to children with NH. In addition, children with CI are found to pay less attention to speech sounds and to the ambient language as compared to children with NH (Houston & Bergeson, 2014; Houston, Pisoni, Kirk, Ying & Miyamoto, 2003). Attention deficits are a key part of executive
functioning. Similarly, children with CI have been shown to have deficits in executive functioning (Kronenberger, Beer, Castellanos, Pisoni & Miyamoto, 2014). As a result of poor speech perception, poor attention and thus executive functioning, children with CI are expected to focus more on salient items in the speech signal and are therefore probably missing low salient grammatical morphemes. This results in poorer production of such morphemes. The focus on salient items in speech perception and the reflection on speech production is for instance shown by Svirsky et al. (2002). They found the most prominent grammatical markers to be acquired first in children with CI, while prominence is not shown to influence acquisition order in children with NH. Likewise, Guo et al. (2013), Hammer (2010) and Szagun (2002) connect poor inflectional development of children with CI to the reduced auditory input.

The second explanation, i.e. deficits in lexical organisation and retrieval, is also linked to poor speech perception in children with CI. Weak attention to the speech signal and poor speech perception in children with CI influence phonological representations in children with CI (Nittrouer et al., 2013; Pisoni et al., 2010). For instance Lund, Werfel and Schuele (2015) have shown poor phonological awareness in children with CI. However, poor phonological representations have a negative influence on word learning (Houston & Bergeson, 2014; Houston et al., 2003). In a similar vein, Wechsler-Kashi, Schwartz and Cleary (2014) reported problems in the lexical organisation of children with CI. Limited phonological representations result in poor lexical organisation and poor connections between words in children with CI. Therefore, lexical organisation is considered to be problematic for children with CI. As a result, Wechsler-Kashi et al. (2014) reported that children with CI have problems to retrieve words in a verbal fluency task in which children were asked to name as many phonologically or semantically related words as possible to a given word. In order to accomplish this task children must search the mental lexicon and switch from one subcategory to another inside the mental lexicon. For this last task, good connections between words are indispensable (Wechsler-Kashi et al., 2014). In MSP, the number of inflected word forms per lemma in spoken language is tallied. There are two possibilities, either the difficulty of children with CI lies only in retrieval or the difficulty lies in storage and retrieval. With respect to the first possibility, children
MLU and MSP development

with CI might have no problems in storage of word forms. But even though children with CI may have stored an equal number of inflected forms per lemma in their mental lexicon than children with NH, poor lexical organisation in children with CI may cause them not to be able to retrieve those words when needed. As a result, MSP is lower in children with CI as compared to children with NH. With respect to the second possibility, the difficulty in children with CI is dual. Children with CI might have stored fewer word forms in their mental lexicon, as it is storage of word forms which has been shown to be problematic in children with CI (Nittrouer et al., 2013). It is likely that the fewer words stored in the mental lexicon, the lower MSP will be. Fewer stored words mean fewer different word forms per lemma and thus lower MSP values in children with CI. The effect of lexical organisation and retrieval problems adds to the poorer storage of word forms. In the already fewer stored words, children with CI also have difficulties in retrieving them, resulting in lower MSP performances. Further research is recommended to examine this hypothesis.

8.4.3 Relationship between syntagmatic and paradigmatic development in children with CI

The third finding of the present study involves the relationship between syntagmatic and paradigmatic development in children with CI. Children with CI are found to have earlier age-appropriate paradigmatic language skills than age-appropriate syntagmatic language skills. When comparing syntagmatic and paradigmatic development of children with CI with children with NH, results suggest that children with CI catch up with their peers with NH. However, children with CI catch up with paradigmatic richness earlier than with syntagmatic richness. Thus, morphological development is delayed for a shorter period than syntactic development. A similar finding with respect to late talkers is reported by Rescorla and Turner (2015): by age five, the group of late talkers had caught up with their typically developing peers on morphology, but not on syntax.

A possible explanation of the relatively fast catch up on paradigmatic richness can be formulated. This explanation is related to the representation of inflected word forms in the mental lexicon. One of the generally accepted views is that each lexical
item has its own representation and thus its own separate entry in the mental lexicon. According to one view in the literature, inflected word forms also have unique entries in the lexicon (Lukatela, Cerello & Turvey, 1987; Lukatela, Gligorijevic, Kostic & Turvey, 1980; Sereno & Jongman, 1997; Stemberger & MacWhinney, 1986). The assumption that inflected word forms have separate lexical representations indicates that learning inflected word forms is closely related to learning new words and thus to lexical development. For instance Song, Sundara and Demuth (2009) have shown that children with larger vocabularies have more accurate third person –s productions, suggesting a close relationship between lexical and inflectional development. Studies have found age-appropriate language skills of children with CI sooner in the lexical domain as compared to syntax and morphology (Caselli et al., 2012; Duchesne et al., 2009; Geers et al., 2003; Moreno-Torres & Torres, 2008; Rinaldi, Baraffaldi, Burdo & Caselli, 2013). As children with CI are found to catch up earlier on lexical than on grammatical development, it is evident that the same trend is visible with respect to paradigmatic, measured there by MSP, and syntagmatic development, measured here by MLU.

8.5 Conclusion

Children with CI are found to lag behind on sentence length and inflectional development as compared to children with NH. However, children with CI catch up on their peers with NH for both syntax and inflection. Nevertheless, inflectional development is earlier age-appropriate as compared to sentence length. This is due to the different nature of both aspects of grammatical development.
Chapter 9
9. Conclusions

The aim of the present dissertation was twofold: (1) examine the effect of particular word-related factors influencing speech production and speech production accuracy at the word level, and (2) trace the development of more specific segments often reported as complex sounds, i.e. speech production and speech production accuracy at the segmental level. For both research aims, the development of children with CI was studied in comparison to that of peers with NH in order to disentangle the effect of auditory deprivation before implantation and degraded speech perception after implantation. The present dissertation took a long-term approach from word onset up to the age of seven and focussed on spontaneous speech production. The goal was not only to examine whether children with CI’s phonological development is delayed, but also to explore whether children with CI do catch up on their peers with NH.

9.1 The outcomes of the present dissertation

The results of the present dissertation have clearly shown a promising developmental trend. The phonological development of children with CI is delayed in comparison with age-matched children with NH. But, children with CI had caught up on nearly all of the measures by age seven. This developmental trend (first delay, but then catch up) complements other studies, which have shown that children with CI catch up on vocabulary, syntax and morphology (Boons, De Raeve, Langereis, Peeraer, Wouters & Van Wieringen, 2013; Nicholas & Geers, 2007; Szagun, 2001; Wie, 2010). The findings of the present dissertation in that they show that early implanted children with CI are also able to catch up on phonological development by age seven. Paragraph 9.2 discusses the relevance of this outcome in a methodological, scientific and clinical perspective. In what follows, a detailed overview of the different measures applied in this dissertation is presented. Paragraph 9.1.1 concerns research aim (1) and paragraph 9.1.2 research aim (2).

9.1.1 Speech production at the word level

In the present dissertation, speech production at the word level was investigated by considering whole-word phonemic accuracy (chapters 3 and 5) and intraword variability (chapter 4). Phonemic accuracy was measured using a relatively new
metric in child language research, i.e. Levenshtein distance (LD), whereas the intraword variability measures were taken from existing studies: the PWV metric of Ingram (2002) (*Proportion of Whole-word Variation*) and the *Variability Score* (e.g. Holm, Crosbie & Dodd, 2007). Effects of target word length in syllables (chapters 3 and 4), a target word complexity (chapter 3) and a target word input frequency (chapter 5) were considered.

In general, results showed that phonemic accuracy is lower in children with CI implanted before their second birthday as compared to peers with NH up to the age of four. By the age of five, however, children with CI have caught up (chapters 3 and 5). Similarly, chapter 4 revealed that intraword variability is higher in children with CI up to age four (both PWV and the *Variability score*). As for phonemic accuracy, children with CI catch up on intraword variability by the age of five as well. Thus, the phonological development at the word level of children with CI is delayed in comparison to age-mates with NH, but the difference between the two groups of children has disappeared by the age of five.

Next to this general picture of accuracy and variability in both groups of children, the analyses revealed that the target words’ characteristics affect children’s speech production accuracy and variability. In the present dissertation, three aspects of target words were considered: their syllable length, their complexity and their frequency in child-directed speech. All three were found to have a significant impact on speech production in both groups of children. Firstly, longer target words are produced with more variation and less accurately (chapters 3 and 4). Secondly, more complex target words, e.g. target words with consonant clusters, are produced less accurately as well (chapter 3). Finally, target words with a higher frequency in child-directed speech are produced more accurately, except for the most frequent target words chapter 5). Moreover, at the older ages, the results suggest a discontinuity: words, which are more frequent, are becoming less accurate, suggesting a trend towards an adult model (chapter 5).

Even though syllable length, complexity and frequency affect speech production accuracy and variability in both groups of children, the results revealed that they did so to a different extent in both groups of children. In other words, differences are
found in the effects of syllable length, complexity and input frequency on the speech production of children with CI and children with NH. Results revealed that (a) the effects of syllable length and complexity are more pronounced in children with CI up to age four (chapters 3 and 4), and (b) the effect of input frequency is less pronounced in children with CI up to age seven (chapter 5). Thus, (a), longer and more complex words are produced less accurately and more variable by both groups of children, but this decrease in accuracy and increase in variability is more dramatic in children with CI. In chapters 3 and 4, we concluded that phonological processing of more complex words seems more difficult for children with CI up to the age of four. Nevertheless, again, results showed that children with CI are catching up on their peers with NH by age five. At this age, syllable length and complexity still affect phonemic accuracy and intraword variability in both groups of children, but to a similar extent. In contrast, (b), the effect of input frequency on phonemic accuracy remains more pronounced in children with NH as compared to CI peers up to the age of seven. In other words, children with CI do no seem to catch up in this respect. The discrepancy between the catch up in (a) and the remaining delay by the age of seven in (b) may be explained by the difference in processing. All factors are likely to be processing effects. However, whereas the effect of syllable length and complexity seems to be related to phonological processing while speaking, the input frequency effect seems related to not only processing in production, but also storage and representation. In chapter 8, we already argued that storage of word forms is problematic in children with CI (see also Nittrouer, Caldwell-Tarr & Lowensthein, 2013).

Finally, analyses at the word level indicate an effect of age at implantation in children with CI. The children with CI in the present dissertation were all implanted at an early age, i.e. between the age of 5 and 20 months. But even in this small range, a benefit of earlier implantation is found. Speech production accuracy is lower in later implanted children (chapter 5) and intraword variability is higher (chapter 5, Variability score). This finding is in line with studies on early implantation (Connor, Craig, Raudenbush, Heavner & Zwolan, 2006). Similarly to the results in this dissertation, research has indicated better outcomes for earlier implanted children with respect to speech production (Leigh, Detmman, Dowell & Briggs, 2013) and speech
Conclusions

production accuracy (Connor et al., 2006). Better outcomes are also found in other domains, such as grammatical development (Boons et al., 2013; Nikolopoulos, Dyar, Archbold & O’Donoghue, 2004), speech perception (Nikolopoulos, O'Donoghue & Archbold, 1999; Svirsky, Teoh & Neuburger, 2003), receptive vocabulary (Schorr, Roth & Fox, 2008), speech intelligibility (Nikolopoulos et al., 1999) and phonological memory (Schorr et al., 2008). Thus, there seems to be a benefit of earlier implantation for overall language development. However, results also revealed that the differences between earlier and later implanted children disappeared with age (chapters 4 and 5). This is probably due to the small range in age at implantation (between the age of 5 and 20 months). Therefore, our results with respect to the effect of early implantation on phonological development in children with CI should be considered with some care.

9.1.2 Speech production at the segmental level

Chapters 3 and 4 showed that phonological complexity and syllable length affect phonemic accuracy and intraword variability, more dramatically in children with CI than in children with NH. We investigated both consonant clusters and fricatives, as those speech sounds are often reported as complex. Therefore, the second goal was to examine the development of consonant clusters and fricatives in children with CI as compared to peers with NH. Consonant clusters are acquired after singleton consonants (Saaristo-Helin, Kunnari & Savinainen-Makkonen, 2011) and fricatives appear among the latest consonants in early word production (e.g. Beers, 1995; Ferguson & Farwell, 1975; Fikkert & Altvater-Mackensen, 2013; Leonard, Newhoff & Mesalam, 1980; Schramm, Bohnert & Keilmann, 2009; Stoel-Gammon, 1985; Van Severen, 2012). In chapter 6, the development of word initial consonant clusters was examined and in chapter 7 that of word initial fricatives. In addition to overall accuracy, different types of inaccuracies are traced.

With respect to consonant cluster development, analyses are performed from word onset up to the age of seven. The likelihood of consonant clusters is similar in both groups of children at all ages. However, there are several differences with respect to accuracy and the different types of inaccuracies. In children with CI, consonant clusters are less accurate, more often reduced to a singleton consonant than substituted
by another consonant cluster, and less often the reduction to a singleton consonant preserving one of the target consonants. Nevertheless, children with CI catch up by the age of five on the accuracy of consonant clusters and by the age of four on the other two types of inaccuracies. This seems to indicate that children first catch up on suprasegmental performance and afterwards on segmental performance. This is in line with other studies on the difference between segmental and suprasegmental development (Carter, Dillon & Pisoni, 2002; Dillon, Cleary, Pisoni & Carter, 2004; Willstedt-Svensson, Löfqvist, Almqvist & Sahlén, 2004). Furthermore, this result supports the radical template phonology of Vihman and Croft (2007), suggesting that children initially have a limited amount of templates, i.e. word shapes, and they adapt their productions to those templates. As such, an initial template for word initial syllables could be CV instead of CCV, as children are found to reduce consonant clusters to singleton consonants. Gradually, children expand their range of word templates and it is shown that this expansion first applies to suprasegmental structure (from CV to CCV) and afterwards to segmental structure (accurate production of both consonants in the cluster).

In addition, an effect of age at implantation is found in consonant cluster production. This effect at the segmental level complements the findings at the word level. In earlier implanted children, consonant clusters are more accurate, less often deleted, less often reduced and less often this reduction to a singleton consonant maintains one of the target consonants. But, children with later implant activation are catching up on their earlier implanted children with age. Similarly to the results at the word level, this is probably due to the small range of ages at implantation in the present study.

In chapter 7, a slightly different approach was taken. In this chapter, the development of word initial fricatives was considered longitudinally, but only up to 30 months of age – instead of the long-term approaches up to the age of five and the age of seven in the other chapters. The reason for this shorter-term approach is straightforward: the goal was to examine whether phonological development and lexical development are commensurate in children with CI and in peers with NH. In other words, we compared the development of word initial fricatives of children with
Conclusions

CI and children with NH with similar chronological ages and with similar lexicon sizes. In order to determine lexicon size, longitudinal corpora are needed. As we only had a longitudinal corpus of children with NH up to the age of 24 months, the analyses were only performed for the NH longitudinal corpus and the CI corpus up to the age of 30 months. Results showed that word initial fricative development is delayed in children with CI: fricatives are less likely to occur, less accurate and more often substituted by a stop. However, when both groups are compared on equal lexicon size, the differences disappear. Therefore, we concluded that phonological development is related to lexical development in children with CI, similarly to peers with NH (Santos & Sosa, 2015; Smith, McGregor & Demille, 2006; Sosa & Stoel-Gammon, 2006; Sosa & Stoel-Gammon, 2012; Stoel-Gammon, 2011; Van den Berg, 2012).

Chapter 7 showed that the likelihood of WI fricatives is only 1.42% in children with CI at the age of 24 months. However, Sotto, Redle, Bandaranayake, Neilson-Strunjas and Creaghead (2014) claimed that the presence of at least one fricative at 18 months of age (regardless of word position) is related to better expressive language development in vocabulary, sentence length and grammar in children with NH. As the likelihood of fricatives is that low at the age of 24 months in children with CI, it may very well be that an effect on vocabulary, sentence length and grammar is found in children with CI at the later ages as well. We explored this hypothesis in chapter 8. Only sentence length and grammar, i.e. inflection, were investigated. Results revealed that children with CI are indeed delayed as compared to peers with NH with respect to mean size of the inflectional lexicon and the mean sentence length. Similarly to the results at the word level and those of consonant clusters at the segmental level, a catch up was also found for inflection and sentence length. Children with CI catch up on their peers with NH by the age of 5 for inflection and by the age of 7 for sentence length.

9.2 Relevance of the dissertation and further research

The conclusions of the present dissertation are methodologically, scientifically and clinically relevant.
Firstly, the present dissertation is methodologically and scientifically relevant as a relatively new measure of phonemic accuracy is introduced. Levenshtein distance (LD) is used in order to consider the whole-word phonemic accuracy. This measure is already commonly used in bio-informatics and dialectology and consists of the comparison of two character strings. Instead of comparing DNA strings of dialectal variants, a phonemic transcription of children’s utterances with adult targets was compared in the present dissertation. Chapter 3 showed that this measure of phonemic accuracy is more fine-grained than existing measures, such as the Proportion of Phonemes Correct (PPC, Shriberg & Kwiatkowski, 1982) and the Phonological Mean Length of Utterance (pMLU, Ingram, 2002).

Secondly, a longitudinal long-term approach phonological development was studied using samples of spontaneous speech. We showed that children with CI’s phonological development is delayed in comparison to age-mates with NH. However, for all measures except input frequency effects, children with CI catch up within the period studied. In other words, the results clearly showed a tendency of first a delay, but then a catch up. This outcome (first delay, but then catch up) is of scientific relevance. Thus far, research on phonological development in CI has often focused on short-term development. Even though there have been a few longitudinal longer-term studies, they mainly used standard tests such as picture naming tasks (PNT), etc. (e.g. Eriks-Brophy, Gibson & Tucker, 2013). However, children’s speech production is less erroneous in single-word tasks, such as PNT, than in spontaneous speech interactions (Healy & Madison, 1987). Therefore, the use of naturalistic data is an added value in the present dissertation. Thus, we added to the body of knowledge on the development of children with CI by tracing the phonological development in spontaneous speech and up to seven years of age. Thus far, it remained to be seen whether children with CI catch up on phonological development. We have clearly shown that a longitudinal approach is indispensable when describing children with CI’s phonological development and language development in general. Further research should focus on longer-term development using connected speech data in order to scrutinize the overall language development of children with CI.
Conclusions

Thirdly, in addition to highlighting the importance of a long-term approach – as discussed above, the *first delay, but then catch up* development has also other relevance, scientific as well as clinical. Five different observations (A – E) can be made here.

(A) The first point concerns the development of fricatives. Analyses revealed that fricative development is similar in both groups of children when compared on lexicon size, but not when compared on chronological age. Thus, phonological and lexical development are commensurate in children with CI, which corresponds to children with NH. Thus, relative to lexicon size, children with CI do not seem to be delayed. This shows that it is essential to examine children with CI’s phonological skills relative to their lexicon size in scientific research, but also for clinical intervention. It remains to be investigated whether children with CI are similar to their peers with NH with equal lexicon size on other aspects of phonological development such as phonemic accuracy, intraword variability, etc. For instance, it should be examined whether intraword variability decrease with increasing lexicon size in children with CI similarly to peers with NH. Therefore, longitudinal data of both children with NH and children with CI are needed. One of the limitations of the present study is that no longitudinal data of children with NH up to ages seven were available. In further research, it is therefore crucial that the longitudinal, long-term development of children with CI is compared to that of children with NH. In this way, the differences and similarities between both groups of children can be analysed in more detail.

(B) The second point in the *first delay, but then catch up* development is the importance for other domains of language development. It has been shown that good phonological skills are related to later reading skills (Catts, 1997; Desjardin, Ambrose & Eisenberg, 2009; Spencer, Barker & Tomblin, 2003; Spencer & Oleson, 2008). In Belgium, children learn to read when they are about five to six years old. A catch up on phonological measures was found here by the age of five. In other words, the phonological representations and phonological development seems age-appropriate by this age – at least for consonant clusters, whole-word accuracy and intraword variability. This seems beneficial for children’s reading development, but the exact
effects remain to be investigated. More research is necessary to understand the precise nature of children with CI’s phonological representations.

(C) The third point concerns the factors identified in the present dissertation as affecting phonological development. We showed that the complexity and the syllable length of target words affect speech production more substantially in children with CI then in children with NH. The opposite is true for input frequency. This suggests less well-developed processing skills in children with CI. In addition to integrating more complex words, longer words and less frequent words in speech and language therapy, children with CI may benefit from enhancing their processing skills. Moreover, children catch up earlier on suprasegmental performance of consonant clusters than on segmental accuracy of these clusters. For speech and language therapy, this highlights the need of support at the segmental level. Intervention and focus on the more problematic parts of phonological development may decrease the time needed for children with CI to catch up. As the age of catching up now coincides with the age at which children learn to read, an earlier recovery might be beneficial for reading development as well.

(D) The fourth point is the effect of good speech production on social and emotional well-being. Research has shown that psychosocial well-being and social and emotional development are related to language (Preisler, Tvingstedt & Ahlström, 2002; Wong, Ching, Leigh, Cupples, Button, Marnane, Whitfield, Gunnourie & Martin, 2016). Speech production accuracy is one of the factors which positively influences speech intelligibility (Ingram, 2002). In turn, good mutual intelligibility is important in social interactions in order to avoid communication breakdowns (Tye-Murray, 2003). Preisler et al. (2002) and Wong et al. (2016) have shown that such social (peer) interactions are of utmost importance in children’s social and emotional development: delays in language development may affect the development of social skills and might prevent those children of taking part in social interactions. However, peer interactions have a substantial part in peoples social well-being (Ginsburg, 2007). Therefore, the catch up of children with CI on speech production, speech production accuracy and speech intelligibility is promising for their conversational skills. As such, this might contribute to improved psychosocial well-being. Thus far, there is
Conclusions

very little research with respect to conversational skills of children with CI. Some studies have shown that children with CI spend more time in communication breakdown and their repair strategies are mostly initiated by others (Ibertsson, Hansson, Mäki-Torkko, Willstedt-Svensson & Sahlén, 2009; Martin, Bat-Chava, Lalwani & Waltzman, 2010; Tye-Murray, 2003). This suggests less developed conversational skills in children with CI. However, these studies are typically on older children – from the age of five onwards and teenagers. But, it remains to be investigated whether repairs also occur in younger children with CI and if there is a development with age. It is yet unclear whether children with CI repair their own speech, whether others initiate these repairs, as is found for older children with CI, or whether they also repair their own speech on their own initiative (self-repair and self-monitoring). In other words, does the degraded auditory feedback in these children with CI prevent monitoring overt speech? It would also be interesting to examine whether there is an immediate relationship between phonological proficiency and conversational skills in children with CI, also younger children with CI, and by extension, to investigate whether there is a relation between their conversational skills, the amount of their interactions and their psychosocial well-being.

(E) The fifth and final point considers the age at implantation. Results have shown that children with CI with earlier implantation have better outcomes on different measures of phonological development as compared to their later implanted peers. This result highlights the benefit of early implantation in congenitally deaf children. Nevertheless, these results should be considered with some care. The group size and the range of age at implantation were small: only 9 children with CI participated in the present study and the children with CI were all implanted between 5 and 20 months of age. Moreover, results also showed that the difference between earlier and later implanted children disappeared with age. In other words, the initial benefit is not found anymore, as children get older. This outcome should be kept in mind when considering early implantation: is the game worth the candle?
Summary – Nederlandse samenvatting
10. Summary – Nederlandse samenvatting

10.1 Summary

The present dissertation discussed the phonological development of children with cochlear implants (CI) in comparison to that of peers with normal hearing (NH). Spontaneous speech production is investigated from word onset up to seven years of age. This long-term approach aimed to identify whether children with CI are (a) delayed in comparison to their peers with NH, but (b) catching up on their peers with NH within the period studied. We looked at (1) speech production accuracy and intraword variability at the word level, including the effect of a few word-related factors on those two variables (chapters 3 – 5), and (2) the development of a few complex speech sounds, i.e. the development at the segmental level (chapters 6 and 7). Moreover, a supplementary study was performed with respect to inflection and sentence length of both groups of children (chapter 8).

As the aim was to study the phonological development of children with CI in a long-term perspective, a longitudinal corpus of 9 children with CI was used. These children were all implanted at an early age, i.e. before 20 months of age. Monthly data were available from one month before implantation up to 30 months after implantation and yearly data between the ages of three and seven. In order to make comparisons with peers with NH, two different corpora were used: a longitudinal corpus of 30 children with NH with monthly data between 6 and 24 months of age, and secondly, a cross-sectional corpus of – in total – 63 children with NH with yearly data between the ages of two and seven. A description of these corpora is given in chapter 2, including information on data collection, data transcription, transcription reliability and the amount of data available.

Chapters 3, 4 and 5 investigated the speech production of children with CI at the word level. In chapters 3 and 5, phonemic accuracy is considered, in chapter 4 intraword variability. In general, it was found that phonemic accuracy was lower and intraword variability was higher in children with CI than in peers with NH. However, by the age of five, these differences have disappeared. With respect to children with CI, also an effect of age at implantation was found, showing higher phonemic
accuracy and lower intraword variability rates in children who are implanted earlier. Nevertheless, this difference between earlier and later implanted children disappeared with age.

Chapters 3 – 5 included the effects of syllable length, complexity and word frequency in the analyses. For both groups of children, phonemic accuracy decreased when the target word was longer and more complex (chapter 3). In addition, syllable length also negatively affected intraword variability: children’s word productions were more variable in longer target words (chapter 4). In contrast, phonemic accuracy in both groups of children increased if the target word had a higher input frequency, except for the most frequent words (chapter 5). Next to this general picture, syllable length, complexity and frequency influenced children with NH’s and children with CI’s word productions to a different extent. The effect of syllable length and complexity on phonemic accuracy and on intraword variability was more pronounced in children with CI than in peers with NH (chapters 3 and 4). In contrast, the effect of frequency on phonemic accuracy was less pronounced in children with CI than in children with NH (chapter 5). Even though children with CI did not seem to catch up on the frequency effect by the age of seven, a catch up is found for the effect of syllable length and complexity by the age of five. Up to the age of four, syllable length and complexity affected phonemic accuracy and intraword variability more substantially in children with CI than in children with NH, but by the age of five, this difference had disappeared. At this age, syllable length and complexity still influenced phonemic accuracy and intraword variability in children with CI, but to the same extent as children with NH.

In chapters 6 and 7, the segmental development of more so-called more complex speech sounds was investigated. Chapter 6 examines the development of word initial consonant clusters and chapter 7 that of word initial fricatives.

The development of consonant clusters is delayed in children with CI with respect to both accuracy and the types of inaccuracies (chapter 6). Children with CI were less accurate, more often produced a singleton consonant instead of a cluster and more often used a singleton consonant which was not part of the target cluster. In similar vein as the results on phonemic accuracy and intraword variability, children with CI
are found to catch up on all aspects of consonant cluster production. First, a catch up for the suprasegmental structure is found, followed by a catch up for segmental accuracy. In addition, an effect of age at implantation is found, showing better results for earlier implanted children. Similarly to the results in chapters 3 – 5, this effect of age at implantation disappeared with age.

Chapter 7 studied the development of word initial fricative production, but a slightly different approach was taken. Instead of investigating the long-term development up to the age of seven, only the development up to the age of 30 months is examined. This methodological shift was chosen as we aimed to study the relationship between lexical and phonological development in children with CI. Therefore, longitudinal corpora of both groups of children were required, which was only available for the younger children with NH. The word initial fricative development was delayed in children with CI when they are compared to age-mates with NH: fricatives were less frequent and less accurate and were more often substituted by a stop. These differences disappeared when both groups of children are matched on lexicon size. In other words, phonological development and lexical development are commensurate in children with CI as it is in children with NH.

Chapter 8 is included in the dissertation in order to test a prediction based on the frequency on fricatives. It is shown that fricatives are nearly absent of children with CI’s production and studies suggest that this also affects the development of inflection and sentence length. The results in chapter 8 indicated that children with CI are delayed with respect to the mean size of the inflectional lexicon and the mean length of utterance. Similarly to our results on phonological development (chapters 3 – 6), children with CI caught up on inflection by the age of five and on sentence length by the age of seven.

The conclusion (chapter 9) outlines the relevance of the present results. Methodologically, Levenshtein distance is introduced as a measure of phonemic accuracy in child language research. Scientifically, the importance of long-term longitudinal research on spontaneous speech production was shown. Little was known about the long-term phonological development of children with CI thus far. This dissertation has shown that children with CI are delayed as compared to peers with
NH, but catch up with the older ages for all measures (chapters 3 – 6). In addition, the importance of comparisons on lexicon size was emphasized in chapter 7. In the conclusion, it was also argued that the catch up on phonological development is important in later language and reading development and for children’s social well-being.
10.2 Nederlandse samenvatting

Deze dissertatie onderzocht de fonologische ontwikkeling van kinderen met een cochleair implantaat (CI) in vergelijking met de fonologische ontwikkeling van normaalhorende (NH) kinderen. We analyseerden de spontane spraakproducties van beide groepen kinderen op lange termijn: vanaf het begin van de eerste woordproducties tot en met de leeftijd van zeven jaar. Op die manier werd nagegaan (a) of kinderen met een CI een achterstand hebben in vergelijking met hun NH-leeftijdsgenootjes, en (b) of deze kinderen hun achterstand goedmaken binnen de bestudeerde periode. Twee aspecten werden bestudeerd. Ten eerste werd de accuraatheid en de variabiliteit van spraakproductie op woordniveau bekeken. Daarbij werd ook gekeken naar enkele factoren die een invloed kunnen hebben op de accuraatheid en variabiliteit van woordproducties (hoofdstukken 3 – 5). Ten tweede, op segmenteel niveau, werd de ontwikkeling van twee zogenaamde meer complexe categorieën van spraakklanken bestudeerd (hoofdstukken 6 en 7). Aanvullend werden ook inflectie en gemiddelde zinslengte onderzocht (hoofdstuk 8).

Om de bovenstaande intenties te verwezenlijken, was er een longitudinaal corpus van 9 kinderen met een CI beschikbaar. Deze kinderen werden allen op vroege leeftijd geïmplanteerd, meer bepaald voor een leeftijd van 20 maanden. Het CI-corpus bestond uit maandelijkse data vanaf één maand voor implantatie tot en met 30 maanden na implantatie en jaarlijkse data tussen drie en zeven jaar. Om de vergelijkingen met NH-kinderen te kunnen maken, waren er twee corpora voorhanden. Ten eerste was er een longitudinaal corpus met 30 NH-kinderen, die maandelijkse data vanaf 6 tot 24 maanden gevolgd werden. Ten tweede was er een cross-sectioneel corpus met in totaal 63 NH-kinderen, dat jaarlijks data bevatte tussen twee en zeven jaar. In hoofdstuk 2 is een gedetailleerd overzicht te vinden over de participanten, de dataverzameling, de dataverwerking, de betrouwbaarheid van de data, etc.

In hoofdstukken 3 tot 5 werd de spraakproductie van kinderen met een CI op woordniveau geanalyseerd. Accuraatheid van de spraakproductie werd bestudeerd in hoofdstukken 3 en 5, variabiliteit van de spraakproductie in hoofdstuk 4. Algemeen toonden onze resultaten aan dat fonemische accuraatheid lager was en variabiliteit van
woordproducties hoger was bij de kinderen met een CI in vergelijking met NH-leeftijdsgenootjes. Echter, de verschillen tussen beide groepen kinderen verdwenen vanaf vijfjarige leeftijd. Er werd tevens een effect van implantatieleeftijd gevonden voor kinderen met een CI: fonemische accuraatheid was hoger en variabiliteit was lager in kinderen met een CI die eerder geïmplanteerd werden. Dit verschil verdween echter naarmate de kinderen ouder werden.

Naast dit algemene beeld werd er in hoofdstukken 3, 4 en 5 ook gekeken naar enkele factoren die accuraatheid en variabiliteit kunnen beïnvloeden, met name de syllabelengte, de complexiteit en de frequentie van de targetwoorden, i.e. de volwassen equivalent van de productie. In beide groepen kinderen daalde de accuraatheid wanneer het targetwoord meer syllabes telde en complexer was (hoofdstuk 3) en waren woordproducties variabeler wanneer de targetwoorden langer waren, i.e. meer syllabes telden (hoofdstuk 4). In tegenstelling tot de negatieve effecten van woordlengte in syllabes en complexiteit, werd er een positief effect voor woordfrequentie gevonden. Woordproducties waren accurater wanneer het targetwoord frequenter was in kindgerichte spraak, behalve wat betreft de meest frequente woorden, meestal woorden uit een gesloten woordklasse zoals lidwoorden (hoofdstuk 5). Onze analyses hebben uitgewezen dat deze effecten van syllabelengte, complexiteit en frequentie aanwezig zijn in beide groepen kinderen, maar in een verschillende gradatie. Zo was het negatieve effect van syllabelengte en complexiteit van targetwoorden op accuraatheid en variabiliteit meer uitgesproken bij kinderen met een CI dan bij NH-leeftijdsgenootjes, maar enkel tot op vierjarige leeftijd (hoofdstuk 3 en 4). Het positieve effect van frequentie, daarentegen, was geprononceerder in NH-kinderen tot en met zevenjarige leeftijd. Dus, syllabelengte, complexiteit en frequentie beïnvloeden woordproducties in verschillende mate in beide groepen kinderen. Hoewel kinderen met een CI hun achterstand niet leken goed te maken voor frequentie binnen de bestudeerde periode, werd er wel een inhaalbeweging gevonden voor wat betreft syllabelengte en complexiteit. Op vijfjarige leeftijd was de invloed van deze twee factoren gelijkaardig in beide groepen kinderen.
In hoofdstukken 6 en 7 werd de ontwikkeling van enkele zogenaamde *complexe* groepen spraakklinken bekeken op segmenteel niveau. In hoofdstuk 6 werd er dieper ingegaan op de verwerving van woordinitiële consonantenclusters, in hoofdstuk 7 op de verwerving van woordinitiële fricatieven.

Uit de resultaten in hoofdstuk 6 bleek dat de verwerving van woordinitiële consonantenclusters vertraagd verloopt bij kinderen met een CI. In vergelijking met NH-leeftijdsgenootjes waren consonantenclusters minder accuraat, werden clusters vaker gereduceerd tot een enkelvoudige consonant en was die enkelvoudige consonant minder vaak één van de consonanten uit het *target*, wat in dit geval telkens de volwassen equivalent was van de consonantenclusters. Net zoals bij de resultaten op woordniveau, werd ook hier een inhaalbeweging gevonden. De productie van consonantenclusters was gelijkenaardig in kinderen met een CI en NH-leeftijdsgenootjes vanaf vierjarige leeftijd voor wat betreft suprasegmentele aspecten (de reductie van een consonantcluster tot een enkelvoudige consonant en het type enkelvoudige consonant) en vanaf vijfjarige leeftijd voor wat betreft accuraatheid. Daarenboven werd er een effect van implantatieleeftijd gevonden dat eveneens gelijklopend is met de resultaten op woordniveau: kinderen met een CI die vroeger geïmplanteerd werden, vertoonden initiële betere resultaten dan later geïmplanteerde kinderen, maar deze verschillen verdwenen naarmate kinderen ouder werden.

In hoofdstuk 7 werd de verwerving van woordinitiële fricatieven bestudeerd. Daarvoor werd een licht ander onderzoeksopzet gebruikt. Ons doel was immers om na te gaan of lexicale en fonologische ontwikkeling bij kinderen met een CI ook gerelateerd zijn, zoals het geval is bij NH-kinderen. Derhalve werden kinderen met een CI niet enkel vergeleken met NH-leeftijdsgenootjes, maar ook met NH-kinderen gematcht op lexicongrootte. Om de evolutie in lexicongrootte te kunnen berekenen, is een longitudinaal corpus nodig voor beide groepen kinderen. Bijgevolg werd de ontwikkeling van woordinitiële fricatieven enkel bekeken in het longitudinaal corpus van NH-kinderen tot 24 maanden en een gedeelte van het longitudinaal CI-corpus, met name tot 30 maanden. Uit de resultaten bleek dat kinderen met een CI minder frequent een fricatief produceren dan NH-leeftijdsgenootjes, en dat wanneer ze een fricatief trachtten te produceren (i.e. een fricatief als *target*), die fricatief minder vaak accuraat
was en vaker vervangen werd door een plosief. Deze verschillen tussen beide groepen kinderen verdwenen echter wanneer ze niet op chronologische leeftijd, maar op lexicongrootte werden gematcht. Kinderen met een CI en NH-kinderen met een overeenkomstige lexicongrootte hadden met andere woorden eenzelfde verwerving van woordinitiële fricatieven. Daaruit kan besloten worden dat ook in kinderen met een CI fonologische ontwikkeling gerelateerd is aan lexicale ontwikkeling.

Hoofdstuk 8 werd opgenomen in deze dissertatie om de hypothese die uit hoofdstuk 7 volgde, te testen. Er werd gevonden dat fricatieve zo goed als afwezig zijn in de spraak van kinderen met een CI op 24 maanden. De literatuur suggereert dat de afwezigheid van fricatieve in de vroege spraakproductie een negatief effect heeft op de latere ontwikkeling van woordenschat, inflectie en zinslengte. Dit laatste werd getest in hoofdstuk 8 voor wat betreft inflectie en zinslengte. Uit de resultaten bleek dat kinderen met CI inderdaad een achterstand hadden voor wat betreft de gemiddelde grootte van het geïnflcheerde lexicon en de gemiddelde zinslengte. Maar net zoals voor de resultaten voor fonologische ontwikkeling, werd er ook hier een inhaalbeweging gevonden. De gemiddelde grootte van het geïnflcheerde lexicon in kinderen met een CI was gelijkwaardig aan dat van NH-kinderen vanaf vijfjarige leeftijd en de gemiddelde zinslengte vanaf zevenjarige leeftijd.

Tot slot werd de relevantie van deze dissertatie aangetoond in hoofdstuk 9. Op methodologisch vlak werd *Levenshtein distance* geïntroduceerd als een relatief nieuwe maat in onderzoek naar kindertaalverwerving. Op wetenschappelijk vlak hebben we aangetoond dat het van aanzienlijk belang is om fonologische ontwikkeling longitudinaal te onderzoeken en om spontane spraak te gebruiken. Tot dusver was er weinig onderzoek naar fonologische ontwikkeling op lange termijn in kinderen met een CI. Wij hebben laten zien dat kinderen met een CI initieel een achterstand hebben in vergelijking met hun NH-leeftijdsgenootjes, maar dat ze deze achterstand inhalen wanneer ze ouder worden. Aanvullend wezen onze resultaten uit dat het belangrijk is om niet enkel kinderen met een CI te vergelijken met leeftijdsgenootjes in zowel wetenschappelijk onderzoek als in de klinische praktijk, maar om ook groepen kinderen met overeenkomstige lexicongroottes met elkaar te vergelijken. Tot slot hebben we aangetoond dat de inhaalbeweging die kinderen met een CI maken van
belang is voor hun latere taalontwikkeling, voor de ontwikkeling van hun leesvaardigheid en voor hun psychosociaal welzijn.
11 References


Measures of digit span and verbal rehearsal speed in deaf children after more than 10 years of cochlear implantation. *Ear and Hearing*, 32(1), 60S - 74S. doi: 10.1097/AUD.0b013e3181ffd58e


Mean length of utterance levels in 6-months intervals for children 3 to 9 years with and without language impairments. *Journal of Speech, Language and Hearing Research*, 53(2), 333 - 349. doi: 10.1044/1092-4388(2009/08-0183)


The phonological mean length of utterance: methodological challenges from a crosslinguistic
perspective. *Journal of Child Language*, 33, 179 - 190. doi: 10.1017/S0305000905007294


IEEE Workshop on spontaneous speech processing and recognition, Tokyo, Japan.


Young, G., & Killen, D. (2002). Receptive and expressive language skills of children with five years of experience using a cochlear implant. *Annals of Otology,
Rhinology & Laryngology, 111, 802 - 810. doi: 10.1177/000348940211100908