Rapid learning of minimally different words in five- to six-year-old children: effects of acoustic salience and hearing impairment

MARCEL R. GIEZEN*

Laboratory for Language and Cognitive Neuroscience, San Diego State University

PAOLA ESCUDERO

MARCS Institute, University of Western Sydney

AND

ANNE E. BAKER

University of Amsterdam, Department of Linguistics

ABSTRACT

This study investigates the role of acoustic salience and hearing impairment in learning phonologically minimal pairs. Picture-matching and object-matching tasks were used to investigate the learning of consonant and vowel minimal pairs in five- to six-year-old deaf children with a cochlear implant (CI), and children of the same age with normal hearing (NH). In both tasks, the CI children showed clear difficulties with learning minimal pairs. The NH children also showed some difficulties, however, particularly in the picture-matching task. Vowel minimal pairs were learned more successfully than consonant minimal pairs, particularly in the object-matching task. These results suggest that the ability to encode phonetic detail in novel words is not fully developed at age six and is affected by task

demands and acoustic salience. CI children experience persistent difficulties with accurately mapping sound contrasts to novel meanings, but seem to benefit from the relative acoustic salience of vowel sounds.

INTRODUCTION

Previous research has shown that, while infants possess phonologically detailed representations of familiar words before their second year of life, the same level of detail is not easily processed for novel words. For example, Stager and Werker (1997) found that infants at the age of 1;2 were unable to learn to associate novel words that only differ in one sound (e.g. *bih* and *dih*) to two different objects, while they were able to auditorily discriminate the same words when word learning was not involved. Subsequent studies showed that infants' ability to learn phonologically minimal pairs appears to be mediated by general factors such as age, vocabulary size, and task demands (for discussion see, e.g. Werker & Curtin, 2005; Yoshida, Fennell, Swingley & Werker, 2009), but also more specific factors such as whether consonant or vowel contrasts are tested (e.g. Nazzi, 2005).

In contrast to the many studies on infant and toddler novel word learning, little is known about novel minimal pair learning in older children. However, a recent study found that three- to five-year-old children still have some difficulty learning phonologically minimal pairs, especially for minimal (i.e. one-feature) contrasts (Havy, Bertoncini & Nazzi, 2011). The finding that typically developing three- to five-year-olds still experience some difficulties with short-term learning of phonologically minimal pairs is in line with reports that children continue to develop their sound perception abilities well beyond five years of age (e.g. Hazan & Barrett, 2000), and that phonological and orthographic representations are still less well specified than those of adults at this age (e.g. Garlock, Walley & Metsala, 2001). Furthermore, three- and four-year-olds still perceptually confuse similar-sounding words (e.g. Gerken, Murphy & Aslin, 1995) and five- to six-year-olds are less sensitive to mispronunciations of familiar words than adults (e.g. Bowey & Hirakis, 2006).

The present study further investigated the short-term learning of phonologically minimal pairs in older children, namely five- to six-year-old children. The aim was to examine the roles of acoustic salience, hearing impairment and task demands when learning consonant and vowel minimal pairs by comparing children with normal hearing to deaf children with a cochlear implant, who have limited years of hearing experience and reduced access to acoustic detail.

311

Learning consonant and vowel minimal pairs: typically developing children

As mentioned earlier, the type of contrast that is tested affects successful learning of phonologically minimal pairs. For instance, Nazzi (2005) compared the performance of twenty-month-old French-speaking infants on several consonant (e.g. /d/–/g/, /p/–/t/) and vowel (e.g. /i/–/y/, /i/–/a/) contrasts embedded in an object categorization task. Word position of the contrast was manipulated in both consonant and vowel contrasts. In addition, phonetic distance was manipulated in the vowel contrasts (less versus more pronounced contrast). The results showed that the infants only successfully learned the novel words that differed in consonant contrasts (regardless of word position), but performed at chance on vowel contrasts.

Nazzi, Floccia, Moquet, and Butler (2009) presented additional evidence for young children's reliance on consonant information in a lexical task. When thirty-month-old French and English toddlers were given the choice between neglecting a vocalic one-feature change and a consonantal one-feature change in a word–object matching task, they chose to neglect the vowel change. In this task, children were presented with three different objects with different (novel) names, e.g. /tyde/, /pide/, and /tide/ (target object). The experimenter would pick up the target object (/tide/) and the infant was asked to match that target object with one of the other two objects, which differed from the target object either in a consonant (/pide/) or vowel (/tyde/).

A focus on consonants in early word learning might reflect a greater reliance on consonants in lexical processing, whereas vowels have more importance at prosodic and morphosyntactic levels (Nespor, Peña & Mehler, 2003). In line with this idea, Havy *et al*. (2011) found that whereas three-year-olds performed better on consonant minimal pairs, four-year-olds performed better on vowel minimal pairs, and five-year-olds showed no difference (Experiment 2). The authors suggested that this change in bias might reflect a (temporary) developmental shift in the use of lexical and morphosyntactic information in language processing between three and four years of age. This hypothesis is further supported by studies that indicate a more important role for consonants in adult word segmentation (Bonatti, Peña, Nespor & Mehler, 2005) and visual word recognition (e.g. Carreiras, Duñabeitia & Molinaro, 2009). Finally, Escudero, Mulak, and Vlach (2015) reported better performance with consonant minimal pairs (/bɔn/–/dɔn/) than vowel minimal pairs (/dit/–/dut/) in a cross-situational statistical learning study with English-speaking adults. In this paradigm, learners are presented with multiple words and objects during each trial of a learning phase and have to track

co-occurrence probabilities across learning trials to infer the word–object mappings.

Importantly, however, results from other studies seem to question the apparent early reliance on consonants. For instance, Dietrich, Swingley, and Werker (2007) found that eighteen-month-old infants were sensitive to small changes to vowels in novel words, and Mani and Plunkett (2008) showed that fourteen-month-old infants notice broad (three-feature) vowel mispronunciations in recently learned novel words. Furthermore, Curtin, Fennell, and Escudero (2009) found that fifteen-month-old infants were able to learn novel minimal pairs with a contrast in vowel height, namely /dɪt/ versus /dit/, which is at an EARLIER age than reported for consonant contrasts with the same task. However, the infants in that study failed to learn minimal pairs with contrasts in lip rounding or vowel backness (/dit/ versus /dut/ and /dɪt/ versus /dut/).

A focus on vowels in early word learning fits well with the idea that sound perception, and possibly acoustic salience, plays an important role in explaining why infants are able to learn words that form phonologically minimal pairs for some sound contrasts at an earlier age than other contrasts (Curtin *et al.*, 2009). Vowels contain more acoustically salient information than consonants, e.g. in terms of acoustic energy (spectral properties) and durational properties, which could help children to encode vowel sounds in novel words. Possibly because of this, sensitivity to native sound contrasts develops earlier for vowel than consonant contrasts in infants (Polka & Werker, 1994). Furthermore, while both consonant and vowel perception develop well into childhood, adult-like perception is achieved earlier for vowels (e.g. Gerrits, 2001). Developmental changes in consonant and vowel perception might be explained by changes in the weighting of language-specific acoustic cues, either because of changes in auditory specificity and spectral distinctiveness (e.g. Mayo & Turk, 2005) or a shift from predominantly relying on dynamic cues, such as formant transitions, to relying on more static cues, such as broad spectral patterns or durational properties (e.g. Hicks & Ohde, 2005; Nittrouer & Miller, 1997).

Sound perception and word learning by children with a CI
A cochlear implant (CI) is an electronic hearing prosthesis that is inserted into the cochlea and directly stimulates the auditory nerve at different places. Although speech processing through a cochlear implant is characterized by relatively poor spectro-temporal resolution, many children with a CI enjoy great benefits in speech perception. However, the outcomes are extremely variable, with high-performers showing near-typical speech perception in quiet listening conditions and low-performers

313

exhibiting only minimal benefits (e.g. Peterson, Pisoni & Miyamoto, 2010). Child, family, implant, and educational factors account for some, but not all of this inter-individual variation (e.g. Davidson, Geers, Blamey, Tobey & Brenner, 2011). One of the most important variables affecting levels of speech perception appears to be age at implantation, with earlier implantation being associated with better outcomes (e.g. Svirsky, Teoh & Neuburger, 2004).

Importantly, it has been shown that children with a CI perceive vowels more accurately than consonants (e.g. Kishon-Rabin, Taitelbaum, Muchnik, Gehtler, Kronenberg & Hildesheimer, 2002; Pisoni, Cleary, Geers & Tobey, 1999). This is despite the fact that spectral information, which is of special importance for the identification of vowels, is distorted relatively more by a CI than temporal and amplitude information, which are more important for the identification of consonants (e.g. Xu & Pfingst, 2008). One possible explanation is that the rapid temporal fluctuations in spectral information in formant transitions for consonants place children with a CI at a disadvantage when using spectral information in the identification of consonants compared to vowels.

Giezen, Escudero, and Baker (2010) examined the use of spectral and durational acoustic cues in the categorization of consonant and vowel contrasts in children with a CI and children with normal hearing with the same chronological age. The contrasts tested were the low vowel contrast /ɑ/–/a/ (a relatively large difference in the first two formants and in duration), the high vowel contrast /ɪ/–/i/ (a relatively small difference in the first two formants and in duration), the place of articulation contrast /f/–/s/ (a difference in intensity and in center of gravity), and the voicing contrast /b/–/p/ (a difference in Voice Onset Time). Identification accuracy for either vowel contrast did not differ between the two groups, but was significantly lower for the children with a CI for both consonant contrasts. Although the children with a CI overall showed more shallow discrimination curves for both vowel and consonant contrasts, they did not differ from their hearing peers in the relative use of spectral and durational cues to discriminate the vowel contrasts, whereas they did experience substantial difficulty with using spectral cues (center of gravity) to discriminate the consonant contrast /f/–/s/. Similarly, Bouton, Serniclaes, Bertoncini, and Colé (2012) reported lower levels of feature identification and feature discrimination in French-speaking children with a CI compared to controls matched on hearing age (years of hearing experience) for both consonant features and vowel features, with particular difficulty for the place of articulation of consonant contrasts.

If acoustic salience and sound perception play an important role in learning phonologically minimal pairs, then children with a CI are expected to perform better on vowel minimal pairs than consonant

minimal pairs, given their more accurate vowel than consonant perception. If, on the other hand, consonants are indeed more informative than vowels for lexical processing (Nespor *et al.*, 2003), then better performance on consonant minimal pairs is expected both for children with normal hearing and for children with a CI. Previous studies that tested phonologically dissimilar words have already shown that children with a CI generally score lower on rapid word learning tasks than children with normal hearing of the same age (e.g. Davidson, Geers & Nicholas, 2014; Houston, Stewart, Moberly, Hollich & Miyamoto, 2012; Willstedt-Svensson, Löfqvist, Almqvist & Sahlén, 2004). To the best of our knowledge, only one other study has looked at the learning of phonologically minimal pairs in children with a CI. Havy, Nazzi, and Bertoncini (2013) tested several less and more pronounced consonant and vowel contrasts in a novel word learning task with 46 three- to six-year-old French children with a CI (implanted between 10 and 57 months). The children with a CI only performed above chance with the more-pronounced contrasts, but no differences between consonant and vowel contrasts were observed. Hearing age (but not age at implantation or chronological age) correlated significantly with word learning scores, suggesting an important role for auditory experience in learning to encode fine phonetic detail in novel words.

The present study aims to provide further insight into minimal pair learning by children with a CI by investigating the role of acoustic salience and task demands. To that end, the stimuli in the present study included novel words that differed in the same consonant and vowel contrasts used in a previous sound perception study with children with a CI (Giezen *et al.*, 2010). Furthermore, minimal pair learning was assessed in the context of low and high task demands, and learning performance of the children with a CI was directly compared to that of children with normal hearing of the same age.

The role of task demands

In word learning studies with younger children, the nature of the task is another important factor that affects successful learning of phonologically minimal pairs. For instance, familiarization with the words and/or objects before the task, the use of naming phrases during familiarization, and the nature of the test phase have all been shown to affect learning performance by young children (Yoshida *et al.*, 2009). Furthermore, Havy *et al.* (2011) discuss the possibility that children at age four years and older may have pragmatic difficulties with conflict situations in word learning tasks. The nature of the word learning task may be especially relevant when testing children with a CI, who have been found to score lower on auditory verbal short-term and working memory measures (e.g. Pisoni, Conway,

315

Kronenberger, Horn, Karpicke & Henning, 2008). This suggests that they might be more sensitive to cognitive demands in novel word learning tasks than children with normal hearing.

To address possible effects of task demands in the present study, the learning of phonologically minimal pairs was assessed with both an on-line picture-matching task and an off-line object-matching task. In the former, pictures of novel objects and audio-recordings of novel words were used, while in the latter real objects were used and the novel words were presented by an experimenter. This increased the social dimensions of the task and, most importantly, allowed for the use of visual speech cues, which might particularly benefit the children with a CI (e.g. Bergeson, Pisoni & Davis, 2005).

METHODS

Participants

Thirteen prelingually deaf children with a CI participated in this study (CI children, mean age 5;9, range: 4;4–6;7, 3 girls). Background characteristics of the children with a CI are given in Table 1. Two other children were tested but excluded because they were unable to complete the word learning tasks for at least one consonant and one vowel contrast. All children received their implant before age 4;0, and their mean age of implantation was 1;9 (range 0;7 to 3;9). On average, they had been using their CI for four years and one month (range one year, seven months to five years, seven months). Four children had (sequential) bilateral CIs and one used an acoustic hearing aid in the non-implanted ear. None of them had known additional disabilities, and surgery was uneventful and the implants were fully inserted for all of them. They were fitted with the latest speech processing algorithm available at the time. Two children were in mainstream education at the time of testing. The others were in schools for the deaf where forms of sign support were used. Their performance was compared to twenty children without a history of speech, language, or hearing impairment of the same chronological age (NH children, mean age 5;10, range: 5;2–6;6, 14 girls). Twenty-one young Dutch adults without a history of speech, language, or hearing impairment (mean age 22;3, range: 19;0–26;3, 18 females) also completed the on-line word learning task for task validation purposes. All participants had Dutch as their native language. The adults all came from the western part of the Netherlands. The children came from the western part of the Netherlands and the northern part of Belgium (Antwerp province).[1]

---

[1] Because the high vowel contrast tends to be more pronounced in the western part of the Netherlands than in the northern part of Belgium, this contrast may have been more

316

TABLE 1. *Background characteristics of the children with a CI*

Participant	Gender	Age	Age at implant	Stimulation	Implant	Education
N7	F	5;1	0;7	CI	Clarion (Platinum)	Bilingual
X5	M	6;7	0;7	CI+HA	Cochlear (Sprint)	SimCom
A1	M	5;3	0;9	CI	Cochlear (Sprint)	Mainstream
J3	M	4;4	0;10	CI	Cochlear (Sprint)	SimCom
V4	M	6;7	0;11	Bilateral CI	Cochlear (Freedom)	SimCom
S7	M	5;2	1;2	Bilateral CI	Cochlear (Sprint) / Digisonic (SP)	SimCom
T1	M	4;10	1;11	CI	Cochlear (Sprint)	SimCom
L2	F	6;0	2;0	CI	Clarion (Platinum)	Bilingual
D8	M	6;7	2;1	Bilateral CI	Cochlear (Sprint/ Freedom)	Mainstream
K3	M	6;4	2;1	CI	Clarion (Platinum)	Bilingual
L6	F	6;7	2;9	Bilateral CI	Digisonic SP	SimCom
L4	M	6;7	3;2	CI	Cochlear (Sprint)	Bilingual
S5	M	5;4	3;9	CI	Cochlear (Freedom)	Bilingual

NOTE: Ages are in years;months.

Stimuli

The auditory stimuli for the picture-matching task were produced by a male adult native speaker of Dutch. Stimuli were recorded with a Sennheiser MKH105 T microphone on a digital TASCAM CD-RW900 recorder in a sound-attenuated room with a sampling rate of 44,100 Hz. The stimuli for the object-matching task were produced live during the experiment by the same male adult speaker. The auditory stimuli were monosyllabic nonwords and familiar words (see 'Appendix'). A set of eight monosyllabic familiar words known to typically developing six-year-old children were selected. Monosyllabic consonant–vowel–consonant words were generated with WordGen$^{©}$ (Duyck, Desmet, Verbeke & Brysbaert, 2004). The minimal nonword pairs were formed contrasting either in the words' vowel (/ɑ/–/a/ or /ɪ/–/i/) or initial consonant (/f/–/s/ or /b/–/p/), for a total of twelve pairs, three per sound contrast. These contrasts had previously been used in a sound perception study with children with a CI and children with normal hearing and were selected to represent different degrees of perceptual difficulty for the children (Giezen *et al*., 2010). Specifically, the /ɑ/–/a/ vowel contrast is more acoustically salient than the

difficult for some children than others in the present study. However, it is important to note that this was true for both the NH children and the CI children. Furthermore, as can be seen in Table 2, performance on the high vowel contrast was clearly above chance for the NH group and the standard deviation for this sound contrast was comparable to that for the other contrasts. Additionally, even if the Flemish children would not produce this contrast themselves, it is very likely that they are exposed to different variants of this sound contrast, including more distinct variants, in their linguistic environments.

317

/ɪ/–/i/ vowel contrast, and the perception of high-frequency noise spectra (as present in the /f/–/s/ consonant contrast) has been found to be particularly difficult for young children, including children with a CI (Kishon-Rabin *et al.*, 2002). Average phonological neighborhood density and neighborhood frequency (per million) were calculated using CLEARPOND (Marian, Bartolotti, Chabal & Shook, 2012) and were, respectively, 11·50 and 2·55 for the /ɑ/–/a/ words, 10·33 and 2·98 for the /ɪ/–/i/ words, 13·00 and 3·05 for /f/–/s/ words, and 17·67 and 28·93 for the /b/–/p/ words. That is, all nonwords had relatively dense neighborhoods, although relatively more so for /b/–/p/ words.

The visual stimuli in the picture-matching task were black-and-white drawings of novel and familiar objects. Pictures of novel objects were selected from a previously used database (Shatzman & McQueen, 2006). Pictures of familiar objects were taken from a publicly available database of black-and-white drawings, which was designed for reading instruction in classrooms. The visual stimuli in the object-matching task were familiar objects (e.g. a spoon, a ball), as well as unfamiliar objects that the children were unlikely to name (e.g. kitchen utensils such as a water dispenser, a scouring pad, or a fruit juice extractor).

On-line word learning task

As an on-line measure for rapid word learning, a picture-matching task was designed that consisted of a familiarization phase and a testing phase (see Figure 1). E-Prime 2·0® (Psychology Software Tools, Pittsburgh PA) was used to present the stimuli and record responses and reaction times. The experiment was divided into four blocks, corresponding to four stimulus sets of two novel words/objects (e.g. /taχ/–/taχ/ with two novel objects) and one familiar word/object (e.g. /hut/ 'hat').

During familiarization, each word/object pair was presented three times in random order embedded in the carrier phrase *Kijk, een X!* 'Look, a X!'. The object was presented in the center of the screen and remained visible for 4,000 ms while the auditory stimulus was presented, which averaged about 3,000 ms in duration. During testing, which immediately followed familiarization, one of the novel or familiar objects was asked for (*Waar is de X?* 'Where is the X?'), and participants then had to choose between two of the familiarized objects presented on the left and right side of the screen. The auditory stimuli during testing also averaged about 3,000 ms in duration. The objects remained visible until participants gave a response by pressing a left or right response key. The next testing trial immediately followed the participant's response.

The testing phase consisted of ten trials presented in random order. They were either TARGET (4) or CONTROL (6) trials. A target trial contained two
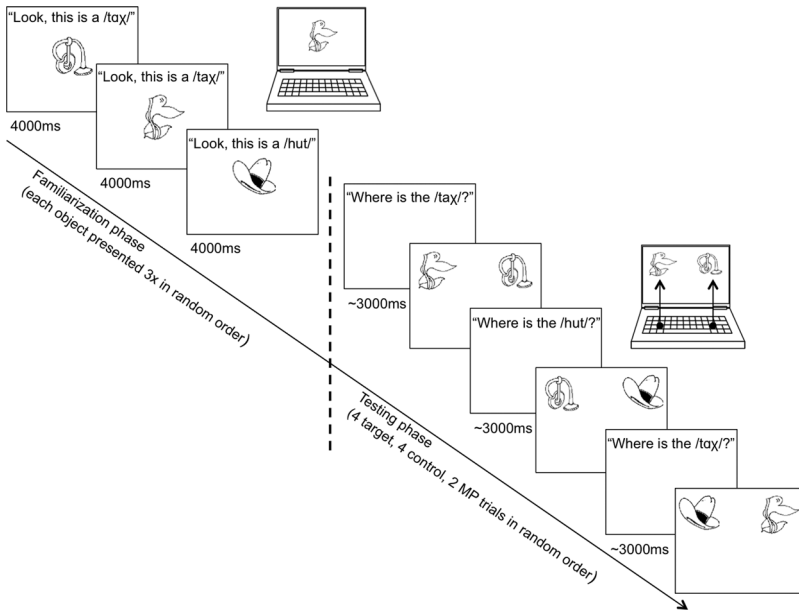
318

Fig. 1. Illustration of the familiarization and testing phases in the on-line picture-matching task.

novel objects, e.g. the /taχ/ and the /taχ/, while a control trial had a novel and the familiar object. In the four target trials, each novel object was tested twice. In the six control trials, the novel and familiar objects were each tested twice. Presentation on the screen (left or right side) was counterbalanced for both novel and familiar objects. Of the six control trials, two trials were included as MISPRONUNCIATION trials, in which the presented nonword was not the correct label for the novel object on the screen (e.g. the auditory stimulus *Waar is de* /taχ/? 'Where is the /taχ/?' followed by pictures of /taχ/ and /hut/ on the screen). These trials were included as a reaction time based measure of discrimination of the novel minimal pairs. They were based on the type of stimuli presented in the mispronunciation paradigm used in infant studies (e.g. Swingley & Aslin, 2002). In this paradigm, children are presented with correct pronunciations (CP) or mispronunciations (MP) of trained familiar or novel word–object pairings. If children have a detailed knowledge of the words, they will fixate less on the target picture in MP trials than in CP trials. Otherwise, no difference in looking times is expected. In the present study, participants might choose either the novel or the familiar object, if they notice the mispronunciation, but in either case a delay in reaction time in MP trials compared to CP trials was expected. Alternatively, if

319

they do not notice the mispronunciation, they should choose the novel object with similar reaction times in CP and MP trials. The participants were not told that some of the trials featured mispronunciations. If they overtly noticed the mispronunciation and objected that the correct answer was not on the screen (all adults, 7 NH children and 1 CI child), they were prompted by the experimenter to still choose one of the two objects.

In sum, the testing phase of the experiment consisted of four target CP trials (novel/novel pair), four control CP trials (novel/familiar pair), and two control MP trials (novel/familiar pair). Accuracy was defined as the number of trials answered correctly. The MP trials were excluded from the accuracy analysis since no errors could be made on these trials. Reaction times (RTs) were measured from the offset of the auditory stimulus until the overt response, i.e. the key press, to avoid an effect of differences between auditory stimuli (initial consonant contrasts versus non-initial vowel contrasts). Only trials answered correctly were analyzed for RTs and trials with RTs longer than 5,000 ms and RTs larger than 2·5 standard deviations above and below the mean for each participant were removed from the analysis. The difference in RTs for the MP and corresponding CP control trials was analyzed to determine whether mispronunciations were noticed. In these trials the same two novel and familiar objects were presented, but the novel object was either correctly (CP) or incorrectly (MP) asked for.

The task took approximately fifteen minutes. It was preceded by a practice block with two novel words not used in the main experiment (/wʏχ/ and /wɔt/) and a familiar word. Familiarization was identical to the experimental blocks, but testing was limited to three trials, two target trials and one control (CP) trial, presented in random order. All participants completed the practice block successfully, except for two CI children who received additional instructions before proceeding with the experiment.

Off-line word learning task

An object-matching task was designed as an off-line measure of rapid word learning to control for performance differences related to task demands by making the task shorter and more interactive. More specifically: (i) colorful tangible objects were used instead of black-and-white drawings; (ii) audiovisual cues were available to the child because the stimuli were presented live by the experimenter instead of using prerecorded audio strings, which also enhances the social-interactive nature of the task; and (iii) only three testing trials rather than ten were presented, none of which were mispronunciation trials.

The experiment was divided into four blocks corresponding to four stimulus sets of two novel words/objects. The child and the experimenter

were seated next to or opposite one another depending on the set-up of the testing room. During familiarization, one familiar and two novel objects were placed on the table in front of the child and presented twice using pointing and naming with the phrases *Kijk, een X* 'Look, a X!'. Familiarization was followed by three testing trials in which the experimenter asked for one of the objects (*Waar is de X*? 'Where is the X?') and the child had to pick up the correct object.

In the first two testing trials, the experimenter asked for one of the novel objects and the familiar object, in random order. In the final testing trial, the experimenter either asked for the remaining novel object or for the same novel object as in the first trial. This was done in order to prevent the children from guessing the answer to the final testing trial based on the two preceding trials. After completion of the first stimulus set, a new set of objects was placed on the table and the familiarization and testing procedures were repeated until all four stimulus sets had been completed.

The task took approximately ten minutes and was preceded by a short novel word learning and generalization test which assessed whether the children could successfully associate a novel word with a novel object and generalize the novel word to a new exemplar of the object (Lederberg, Spencer & Prezbindowski, 2000). All children passed this test. The object-matching task was scored on-site, but also recorded on video for off-line scoring validation. The dependent variable was the number of testing trials with novel objects answered correctly by the child.

Procedure

Testing took place individually in a quiet room at the children's school and in a quiet testing room at a university for the adults. Children completed the two word learning tasks on two separate days, and always completed the picture-matching task first. Presentation of nonword pairs in the on-line and off-line word learning tasks was counterbalanced across children in each group such that subsets of children were presented with different nonword pairs in the same task, but were never presented with the same nonword pair twice across tasks. Task instructions for the children with a CI were provided in sign-supported speech, but no sign support was available during the familiarization and testing phases in the word learning tasks.

The picture-matching task was administered on a DELL© Latitude D630 laptop using two external speakers (Trust© SP-2310). Participants completed this task at a sound level within their own range of comfort. The four word learning blocks (stimulus sets) were presented in counterbalanced order across participants with alternated presentation of consonant and vowel contrasts. The object-matching task was administered

321

live by the experimenter, with alternated presentation of consonant and vowel contrasts.

Statistical analysis

Linear and logit mixed models were computed using the statistical computing environment R (R development core team, 2011, v3·0·2) and the lme4 package (Bates, Maechler, Bolker & Walker, 2013). Mixed-effect modeling was preferred to ANOVAs because of its greater power for repeated observations, flexibility in dealing with missing values, and the absence of assumptions of homogenous variances (for discussion see, e.g. Baayen, Davidson & Bates, 2008; Barr, Levy, Scheepers & Tily, 2013; Jaeger, 2008).

In all analyses, models with a maximal random-effects structure were first attempted (Barr *et al.*, 2013). In our design, item and subject effects are correlated because subsets of subjects completed different items per task (see 'Procedure'). Therefore, models with a maximal random-effects structure included random intercepts for subjects and items, as well as random slopes for subjects for within-subject fixed effects (i.e. Contrast and/or Trial type) and related interaction terms (e.g. random slope for the interaction term of Trial type and Contrast). If the initial maximal model did not converge, random slopes were removed one-by-one until the model converged.

For the on-line picture-matching task, fixed effects of Group (NH children, CI children, adults), Trial type (target, control) and Contrast (/f/–/s/, /b/–/p/, /ɑ/–/a/, /ɪ/–/i/) and the respective interaction terms were included in the model. The factor Trial type was not included in the accuracy analysis for this task, because of a ceiling effect in scores on control trials for NH children and adults (98–100% correct; see Table 2). This choice of removing Trial type from the analysis was further motivated by the fact that even a model with minimal random-effects structure failed to converge with Trial type as fixed effect. Thus, only target trials were included in the model for picture-naming accuracy. In the analysis of mispronunciation effects in the picture-matching task, reaction times for mispronunciation (MP) trials were directly compared to the corresponding correct pronunciation (CP) trials and entered into the model as a fixed effect of Trial type. For the off-line object-matching task, fixed effects of Group (NH children, CI children) and Contrast (/f/–/s/, /b/–/p/, /ɑ/–/a/, /ɪ/–/i/) and the respective interaction term were included in the model.

Only planned comparisons for fixed effects were performed: for Group, adults were compared to children (if adult data were available) and the NH children were compared to the CI children, while for Contrast,

322

TABLE 2. *Observed means and standard deviations (between parentheses) of accuracy scores (proportion correct) for each tested contrast for NH children, CI children, and adults in the picture-matching task and object-matching task (NH and CI children only)*

Task[a]	Trial type	Group	/b/–/p/	/f/–/s/	/ɑ/–/a/	/ɪ/–/i/
Picture-matching[b]	Target trials	NH	·64 (·48)	·58 (·50)	·66 (·48)	·64 (·48)
		CI	·48 (·50)	·48 (·50)	·62 (·49)	·61 (·49)
		Adults	·99 (·11)	·92 (·28)	·92 (·28)	·92 (·28)
	Control trials	NH	·98 (·16)	1·00 (·00)	1·00 (·00)	1·00 (·00)
		CI	·83 (·38)	·88 (·32)	·88 (·32)	·89 (·32)
		Adults	1·00 (·00)	1·00 (·00)	1·00 (·00)	1·00 (·00)
Object-matching[c]		NH	·74 (·45)	·74 (·45)	·89 (·31)	·82 (·39)
		CI	·46 (·48)	·32 (·51)	·73 (·45)	·59 (·50)

NOTES: [a]Four CI children and three NH children did not complete all four blocks of the picture-matching task and/or object-matching task, but did complete each task for at least one vowel and one consonant contrast.
[b]NH: N = 20, CI: N = 13, Adults: N = 21.
[c]NH: N = 20, CI: N = 13.

consonant contrasts were compared to vowel contrasts, and each of these were compared to the other consonant or vowel contrast (i.e. /f/–/s/ versus /b/–/p/, and /ɑ/–/a/ versus /ɪ/–/i/). *P*-values for the fixed effects in linear mixed models were calculated based on maximum likelihood *t*-tests with Satterthwaite's approximation to degrees of freedom, as implemented in the lmer Test package (Kuznetsova, Brockhoff & Bojesen, 2012), and with LaPlace approximation for logit mixed models.

RESULTS

Table 2 and 3 present the descriptive statistics of the observed proportion correct scores and reaction times, respectively, for all available groups, trial types, and contrasts for the picture-matching task and the object-matching task (observed proportion correct scores only).

On-line picture-matching: accuracy

The maximally converging model for accuracy on target trials in the on-line picture-matching task only included random intercepts for subjects and items (see Table 4). This model yielded a significant effect of Group for the comparison between the adults and the two child groups, with the adults scoring higher than the children (Estimate = 2·54, $SE = 0·32$, $p < ·001$). Furthermore, the NH children scored marginally higher than the CI children (Estimate = 0·33, $SE = 0·14$, $p = ·08$). The model also yielded a marginal effect of Contrast (Estimate = –0·76, $SE = 0·40$, $p = ·06$),

323

TABLE 3. *Observed means and standard deviations (between parentheses) of reaction times (ms) for each tested contrast in the picture-matching task for NH children, CI children, and adults*

Trial type[a]	Group[b]	/b/–/p/	/f/–/s/	/ɑ/–/a/	/ɪ/–/i/
Target trials	NH	2590 (1053)	2723 (1126)	2823 (1012)	2359 (945)
	CI	2179 (1136)	2470 (980)	2188 (814)	2496 (1176)
	Adults	1218 (380)	1142 (363)	1192 (347)	1371 (345)
Control trials	NH	2147 (918)	1997 (620)	2034 (746)	1926 (667)
	CI	1761 (721)	2058 (864)	1840 (928)	1868 (893)
	Adults	1058 (327)	1012 (374)	965 (264)	1128 (361)
CP trials	NH	2652 (951)	2147 (688)	2320 (672)	2136 (787)
	CI	1747 (842)	2337 (1000)	2105 (917)	1997 (821)
	Adults	1160 (346)	1131 (450)	1047 (263)	1231 (317)
MP trials	NH	2625 (878)	2462 (911)	2657 (983)	2467 (871)
	CI	2116 (711)	2087 (879)	2195 (845)	2571 (888)
	Adults	1739 (524)	1640 (627)	1557 (455)	1672 (352)

NOTES: CP = correct pronunciation trials; MP = mispronunciation trials.
[a]Four CI children and three NH children did not complete all four blocks of the picture-matching task and/or object-matching task, but did complete each task for at least one vowel and one consonant contrast.
[b]NH: N = 20, CI: N = 13, Adults: N = 21.

which reflected marginally better performance on the voicing /b/–/p/ contrast than the place of articulation /f/–/s/ contrast. Finally, a significant interaction between Group and Contrast was found (Estimate = −1·37, $SE = 0·64$, $p = ·03$). This interaction was further investigated by fitting models on the adult and child data separately. Whereas the adults performed similarly on vowel and consonant contrasts (Estimate = −1·30, $SE = 1·16$, $p = ·26$), the children performed marginally better on vowel contrasts than consonant contrasts (Estimate = 0·32, $SE = 0·18$, $p = ·08$).

On-line picture-matching: reaction times

The maximally converging model for reaction times in the on-line picture-matching task included random intercepts for subjects and items, random slopes for subjects on Trial type, and a random slope for subjects on Contrast (see Table 5). This model yielded a significant effect of Group for the comparison between the adults and the two child groups, reflecting faster reaction times for the adults (Estimate = 1096·92, $SE = 103·75$, $p < ·001$). NH and CI children did not differ significantly. Furthermore, a significant effect of Trial type was found, indicating faster responses on control trials than target trials (Estimate = 414·82, $SE = 49·05$, $p < ·001$). These two main effects were qualified by a significant interaction between Group and Trial type (Estimate = 339·83, $SE = 95·18$, $p < ·001$). This

324

TABLE 4. *Summary of the fixed and random effects in the maximally converging model of accuracy scores in the picture-matching task*

Fixed effects	coefficient	SE	z	p	Random effects	SD
(Intercept)	1·21	0·16	10·17	<·001	Subjects	
Group1	2·54	0·32	8·00	<·001	(Intercept)	0·00
Group2	0·33	0·19	1·78	·08	Items	
Contrast2	−0·76	0·40	−1·90	·06	(Intercept)	0·00
Group1:Contrast1	−1·37	0·64	−2·15	·03		

NOTE: Estimated coefficients and standard error (SE) of fixed effects and standard deviation (SD) of random effects are reported in logits. Coefficients reflect twice the difference between the (average of the) relevant level(s) of the factor and the predicted grand mean. Only fixed effects with $p < ·08$ are reported. Group1 = adults vs. children, Group2 = NH children vs. CI children, Contrast1 = consonants vs. vowels, Contrast2 = /b/–/p/ vs. /f/–/s/.

TABLE 5. *Summary of the fixed and random effects in the maximally converging model of reaction times in the picture-matching task*

Fixed effects	coefficient	SE	t	p	Random effects	SD
(Intercept)	1866·35	51·78	36·04	<·001	Subjects	
Group1	1096·92	103·75	10·57	<·001	(Intercept)	186·80
Trial type	414·82	49·05	8·46	<·001	Trial type	244·80
Group1:Trial type	339·83	95·18	3·57	<·001	Contrast	
					/b/–/p/	394·80
					/f/–/s/	304·30
					/ɑ/–/a/	258·40
					/ɪ/–/i/	364·50
					Items	
					(Intercept)	0·00

NOTE: Estimated coefficients and standard error (SE) of fixed effects and standard deviation (SD) of random effects are reported in ms. Coefficients reflect twice the difference between the (average of the) relevant level(s) of the factor and the predicted grand mean. Only fixed effects with $p < ·08$ are reported. Group1 = adults vs. children.

interaction was further investigated by fitting models on the adult and child data separately. Although clearly significant for both the adults (Estimate = 186·17, $SE = 25·83$, $p < ·001$) and the children (Estimate = 553·76, $SE = 85·40$, $p < ·001$), this interaction may reflect a larger effect on reaction times for the adults.

On-line picture-matching: mispronunciation effect

The maximally converging model for reaction times on mispronunciation (MP) trials and correct pronunciation (CP) trials in the on-line picture-matching task included random intercepts for subjects and items,

325

random slopes for subjects on Trial type, and a random slope for subjects on Contrast. This model yielded a significant fixed effect of Group for the comparison between the adults and the two child groups, reflecting faster reaction times for the adults (Estimate = 902·58, SE = 100·47, p < ·001). Furthermore, the CI children responded marginally faster than the NH children (Estimate = −239·93, SE = 128·22, p = ·07). Additionally, the model yielded a significant fixed effect of Trial type, with slower responses on MP trials than CP trials (Estimate = 316·16, SE = 49·20, p < ·001).

The two main effects were qualified by two significant interactions. First, a significant interaction between Group and Trial type was found (Estimate = −297·80, SE = 98·13, p = ·003), which was further investigated by fitting models to the adult and child data separately and likely reflected enhanced sensitivity to the mispronunciations for the adults (Estimate = 528·95, SE = 64·34, p < ·001) compared to the children (Estimate = 210·80, SE = 71·31, p = ·003). Second, a significant interaction between the two child groups and the two consonant contrasts was observed (Estimate = −670·90, SE = 199·63, p = ·001), but only for CP trials, not MP trials (three-way interaction, Estimate = 889·27, SE = 343·86, p = ·01). This interaction was further investigated by fitting models on the CP trials for the NH and CI children separately. The NH children responded faster on CP trials with the place of articulation /f/–/s/ contrast than the voicing /b/–/p/ contrast (Estimate = 508·33, SE = 176·29, p = ·008), whereas the CI children showed an opposite trend (Estimate = −562·94, SE = 287·80, p = ·06).

Interestingly, in the MP trials the children chose the phonologically similar novel object the majority of the time, which shows that they perform similarly to the adults, who chose this novel object in all MP trials. However, this pattern occurred less frequently in the CI than in the NH group (9 out of 156 versus 18 out of 100 trials, χ^2 (1, N = 256) = 9·66, p = ·002). In infant studies using the MP paradigm, young children typically look longer at a phonologically dissimilar familiar object (e.g. *car*) when they are presented with a mispronunciation in familiar or novel words (e.g. *vaby* instead of *baby*) than when the words are correctly pronounced (e.g. Swingley & Aslin, 2002). Our finding that the CI children chose the phonologically novel object less often than the NH children and the adults might further point to delayed word learning abilities due to limited hearing experience.

Off-line object-matching task: accuracy
The maximally converging model for accuracy in the object-matching task only included random intercepts for subjects and items (see Table 7).

326

TABLE 6. *Summary of the fixed and random effects in the maximally converging model of mispronunciation effects in the picture-matching task*

Fixed effects	coefficient	SE	t	p	Random effects	SD
(Intercept)	1851·66	50·24	36·86	<·001	Subjects	
Group1	902·58	100·47	8·98	<·001	(Intercept)	216·40
Group2	−239·93	128·22	−1·87	·07	Trial type	108·30
Trial type	365·79	49·06	7·46	<·001	Contrast	
Group1:Trial type	−297·80	98·13	−3·04	·003	/b/–/p/	377·50
Group2:Contrast2	−670·90	199·63	−3·36	<·001	/f/–/s/	303·60
Group2:Trial type: Contrast2	889·30	343·86	2·59	·01	/ɑ/–/a/	133·00
					/ɪ/–/i/	202·00
					Items	
					(Intercept)	0·00

NOTE: Estimated coefficients and standard error (SE) of fixed effects and standard deviation (SD) of random effects are reported in ms. Coefficients reflect twice the difference between the (average of the) relevant level(s) of the factor and the predicted grand mean. Only fixed effects with $p < ·08$ are reported. Group1 = adults vs. children, Group2 = NH children vs. CI children, Contrast1 = consonants vs. vowels, Contrast2 = /b/–/p vs. /f/–/s/.

TABLE 7. *Summary of the fixed and random effects in the maximally converging model of accuracy scores in the object-matching task*

Fixed effects	coefficient	SE	z	p	Random effects	SD
(Intercept)	0·82	0·19	4·36	<·001	Subjects	
Group	1·37	0·37	3·73	<·001	(Intercept)	0·26
Contrast1	1·01	0·32	3·22	·01	Items	
					(Intercept)	<0·01

NOTE: Estimated coefficients and standard error (SE) of fixed effects and standard deviation (SD) of random effects are reported in logits. Coefficients reflect twice the difference between the (average of the) relevant level(s) of the factor and the predicted grand mean. Only fixed effects with $p < ·08$ are reported. Contrast1 = consonants vs. vowels.

This model yielded a significant fixed effect of Group (Estimate = 1·37, SE = 0·37, $p < ·001$), with the CI children scoring lower than the NH children. Additionally, the model yielded a significant effect of Contrast for consonants versus vowels (Estimate = 1·01, SE = 0·31, $p = ·001$), reflecting better performance on the vowel contrasts than consonant contrasts. Accuracy scores for the two vowel contrasts did not differ significantly and neither did accuracy scores between the two consonant contrasts (ps > ·10).

If we descriptively compare the performance of both groups of children across tasks (see Table 2), it seems that performance substantially improved for the children with normal hearing in the object-matching task

(0·14 increase in average proportion correct score across the four contrasts), but clearly less so for the children with a CI (0·05 decrease in average proportion correct score across the four contrasts). The only exception appeared to be the low vowel contrast /ɑ/–/a/ (0·11 increase in average proportion correct score for the CI children). Furthermore, the children with a CI performed at chance for consonant minimal pairs in both the picture-matching task and the object-matching task. Thus, the relatively low demands of the object-matching task had a positive impact on the performance of the children with normal hearing, but not on the children with a CI, who still experienced substantial difficulty, especially with consonant minimal pairs.

Relationship with age at implantation and hearing age

To examine the role of age factors in explaining inter-individual variation in performance by the CI children, we examined correlations between their word learning scores and their age at implantation and hearing age (years of hearing experience). As is often the case in studies that test CI children with different ages at implantation, but similar chronological ages, age at implantation and hearing age correlated significantly ($r = -0.71$, $p = .007$). Neither age at implantation nor hearing age correlated significantly with scores on the object-matching task or scores on target trials in the picture-matching task ($ps > .10$). Furthermore, for the off-line object-matching task, entering hearing age as continuous predictor (grand mean centered) to the maximally converging model did not remove the effects of Group and Contrast. However, for the on-line picture-matching task, the effects of Group were no longer significant when hearing age was added as predictor, suggesting that differences in hearing experience for NH and CI children may have contributed to the marginally higher accuracy scores for NH children in this task.

DISCUSSION

The present study examined the role of acoustic salience in the rapid word learning of five- to six-year-old children with a CI and children with normal hearing (NH) of the same age. Two vowel (/ɑ/–/a/ and /ɪ/–/i/) and two consonant contrasts (/f/–/s/ and /b/–/p/) were tested in two different word learning tasks, an on-line picture-matching task and an off-line object-matching task. The CI children's accuracy was only marginally lower than that of the NH children on target trials in the picture-matching task, and both groups of children exhibited sensitivity to mispronunciations (switching of novel word–object mappings) in their reaction times. The NH children clearly did not yet perform at the level obtained by adults on the picture-matching task. Importantly, the scores

328

of the NH children markedly improved in the less demanding off-line object-matching task, on which they performed significantly higher than the CI children. In the on-line picture-matching task, NH and CI children performed only marginally better with vowel than consonant minimal pairs, whereas enhanced performance for vowel minimal pairs was statistically robust in the off-line object-matching task.

Although children can successfully learn phonologically minimal pairs and show sensitivity to small mispronunciations in recently learned words before the age of two (Werker & Curtin, 2005), our results show that novel phonologically minimal pairs still provide a challenge to five- to six-year-old children in rapid word learning tasks, especially in the context of high task demands. Our findings are in line with other studies that have reported difficulties for older children in tasks that require the processing of phonetic detail in novel or familiar words (e.g. Bowey & Hirakis, 2006; Gerken *et al.*, 1995; Havy *et al.*, 2011). It remains unclear whether this apparent discrepancy between the results obtained with infants and (older) children might be due to, for example, methodological differences between studies, developmental changes in using overall acoustic or feature information structures versus phonological or segmental information, or increased lexical competition resulting from crowding in phonological space (for discussion see, e.g. Henderson, Weighall, Brown & Gaskell, 2013).

The CI children did not appear to benefit from a reduction in task demands to the same extent that the NH children did, demonstrating clear difficulties with learning phonologically minimal pairs. Several methodological differences between the picture-matching task and the object-matching task used in the present study made the latter less cognitively taxing than the former: (i) tangible objects were used rather than drawings; (ii) an experimenter presented the novel words rather than using audio-recordings, which increases the social-interactive nature of the task and allows for the use of visual speech cues; and (iii) no mispronunciation trials were presented. Interestingly, Havy *et al.* (2011) suggested that children aged four and older might have pragmatic difficulties with conflict situations in word learning tasks. Although different in nature, the on-line picture-matching task in the present study also included a conflict dimension (mispronunciation trials), which might have influenced children's processing of the remaining, non-conflict trials. This conflict dimension was not present in the object-matching task, which might explain the better performance by NH children on that task. It might also explain why performance differences for consonant and vowel contrasts were more pronounced in the object-matching task.

Furthermore, the availability of visual speech cues may have contributed to the enhanced performance of the NH children in the object-matching task. The benefit of visual speech information for typically developing

children, and particularly children with hearing loss, is well established (e.g. Woodhouse, Hickson & Dodd, 2009), although recent research has shown reduced sensitivity to visual speech for children in the (chronological) age range of the present study (e.g. Jerger, Damian, Spence, Tye-Murray & Abdi, 2009). Furthermore, none of the tested contrasts had very strong visual correlates. Indeed, whereas the children with a CI would be expected to particularly benefit from the audiovisual context (e.g. Bergeson *et al.*, 2005), their performance in the object-matching task clearly did not improve to the same extent as that of the NH children. Nevertheless, it remains possible that more robust improvements would be observed when contrasts with clear visual correlates are tested (see, e.g. Giezen, Baker & Escudero, 2014).

Interestingly, although the difference between consonant and vowel minimal pairs for NH and CI children did not reach statistical significance in the on-line picture-matching task, in the off-line object-matching task the NH and CI children performed better with vowel minimal pairs than consonant minimal pairs. One possible explanation for the observed better performance with vowel contrasts is that these contain more acoustically salient information than consonant contrasts, e.g. in terms of spectral and durational properties, and that this helps children to perceive and/or encode vowel sounds in novel lexical representations (cf. Curtin *et al.*, 2009). The relative acoustic salience of some vowel contrasts might especially benefit children with a CI who have difficulty with processing the rapid temporal fluctuations that characterize formant transitions in consonant sounds. This explanation is in line with studies reporting more accurate vowel than consonant perception in this population (Kishon-Rabin *et al.*, 2002; Pisoni *et al.*, 1999), as well as an earlier study with Dutch-speaking children with a CI that examined the perception of the same sound contrasts as tested in the present study and reported better performance with vowel than consonant contrasts (Giezen *et al.*, 2010).

It is important to note that, in contrast to the present findings, other word learning studies have found better performance with consonant contrasts than vowel contrasts (e.g. Havy *et al.*, 2011; Nazzi, 2005; Nazzi *et al.*, 2009). One possible explanation for the divergent findings are methodological differences between the studies. Indeed, the substantial discrepancy between performance in the picture-matching task and the object-matching task for the NH children in the present study suggests that task demands affect successful learning of phonologically minimal pairs. Their overall performance was lower in the on-line picture-matching task and no statistically robust differences between consonant and vowel minimal pairs were observed in that task.

Another possible explanation for the divergent findings may follow from the relative phonological distance between the consonant and vowel

330

contrasts that were tested in the present study. Havy *et al*. (2011) compared less discriminable (one-feature change, e.g. /p/–/t/ or /u/–/y/) versus more discriminable (three-feature change, e.g. /t/–/j/ or /y/–/o/) consonant and vowel contrasts and found better performance for the more discriminable contrasts. In the present study, the consonant contrasts involved one-feature contrasts (voicing for /b/–/p/ and center of gravity for /f/–/s/) and the vowel contrasts three-feature contrasts (height, backness, and duration for /ɑ/–/a/ and /ɪ/–/i/). It is therefore possible that the better performance with vowel contrasts in the object-matching task was at least partly driven by relative phonological distance in the consonant and vowel contrasts. It is important to note, however, that although both /ɑ/–/a/ and /ɪ/–/i/ are three-feature contrasts, /ɑ/–/a/ acoustically is a more salient contrast than /ɪ/–/i/ because of a larger difference in acoustic energy of the first and second formants as well as a larger difference in duration. This could explain why performance with the /ɑ/–/a/ minimal pairs was relatively high in the off-line object-matching task for both NH and CI children, and the only vowel contrast that yielded improved accuracy scores in the off-line object-matching task compared to the on-line picture-matching task for the CI children (see Table 2). Both number of feature changes and acoustic salience may therefore contribute to differential performance for vowel versus consonant minimal pairs in novel word learning tasks. Of course, it remains to be seen whether the results obtained in the present study with these particular vowel and consonant contrasts generalize to other vowel and consonant contrasts with similar feature changes, but differences in acoustic salience. In addition, it should be noted that the mixed models for accuracy scores only converged with a minimal random-effects structure (i.e. random intercepts). Although it is not uncommon that logistic mixed models with random slopes in datasets with small samples do not converge, it is nevertheless important to mention that random intercepts only models can be anti-conservative and therefore have to be interpreted with caution (Barr *et al*., 2013).

The fact that the consonant contrasts were only distinguished by one feature may have caused particular difficulty for the children with a CI, as also shown in a recent study by Havy *et al*. (2013). They tested one-feature and three-feature consonant and vowel contrasts in their study with three- to- six-year-old children with a CI. The children with a CI performed at chance for the one-feature contrasts, but above chance for the three-feature contrasts. The results from that study and the study presented here thus suggest that minimal, one-feature contrasts present severe challenges to children with a CI, even when visual speech cues are available and with reduced task demands (the object-matching task in the present study). According to Havy *et al*. (2013), the lower performance with the less discriminable contrasts might result from their perceptual

331

problems or from more general problems with the use of phonetic detail in a word learning context, similar to those experienced by typically developing children. Results from (non-)word discrimination studies with children with a CI showing similar difficulties with several one-feature contrasts speak in favor of a perceptual explanation (e.g. Bouton *et al.*, 2012). Furthermore, Davidson *et al.* (2014) recently reported significant positive correlations between aided pure-tone auditory thresholds, open-set word recognition scores, and novel word learning scores in a large sample of 101 children with a CI between six and twelve years old. However, it should be noted that several other studies with smaller samples have failed to find significant correlations between rapid word learning scores and speech perception measures for children with a CI (e.g. Houston *et al.*, 2012; Willstedt-Svensson *et al.*, 2004).

An alternative or additional explanation is that CI children are more likely to try to relate novel words to words they know than children with normal hearing, making any task that requires the processing of nonwords relatively difficult for them (see, e.g. Schwartz, Steinman, Ying, Mystal & Houston, 2013). Because consonant contrasts tend to be associated with larger phonological neighborhoods than vowel contrasts (Cutler, Sebastián-Gallés, Soler-Vilageliu & Van Ooijen, 2000), this could also explain why both the NH and CI children in the present study appeared to have more difficulty with consonant minimal pairs. Although all minimal pairs in the present study had relatively large phonological lexical neighborhoods, the consonant minimal pairs were associated with larger neighborhoods than the vowel minimal pairs, especially the /b/–/p/ minimal pairs.

Further support for an increased tendency to relate novel words to familiar words among children with a CI is found in the fact that CI children scored relatively low even on control trials (one novel object, one familiar object) in the on-line picture-matching task compared to the NH children (see Table 2). However, this could also be the result of difficulty in maintaining (auditory) attention to tasks that involve high cognitive demands. Sustained auditory attention was not explicitly assessed in the present study, but other work has shown, for example, that infants with a CI exhibit reduced attention to speech compared to infants with normal hearing (Houston & Bergeson, 2014). Clearly, more research is needed to determine the role of attentional demands in explaining attested performance difference between children with a CI and children with normal hearing.

Havy *et al.* (2013) reported a significant correlation between hearing age (but not age at implantation or chronological age) and word learning scores in their study on novel minimal pair learning in CI children. They interpreted this relationship as evidence for a strong role for auditory

experience in learning to encode fine phonetic detail in novel words. Neither age at implantation nor hearing age correlated with word learning scores for the CI children in the present study. Differences in the sampled range of age at implantation and hearing age between the two studies or the small sample size and resulting limited statistical power to detect significant effects of age predictors in the present study might explain the discrepant findings between the two studies. Ideally, a future study should compare samples of children with a CI implanted at different ages with normally hearing children matched on hearing age to more closely investigate the role of hearing experience and early access to sound in learning phonologically minimal pairs.

Future studies should also more directly look at the role of CI children's vocabulary size and growth in learning phonologically minimal pairs. It is well known that accumulated vocabulary knowledge is an important predictor of rapid word learning performance in typically developing children (for discussion see, e.g. Gathercole, 2006). Houston *et al*. (2012) found strong correlations between rapid word learning performance and receptive vocabulary in their study of younger children with a CI. Furthermore, Davidson *et al*. (2014) found that novel word learning scores predicted independent variance in receptive vocabulary after accounting for other known predictors. Of course, it is both possible that a larger vocabulary enhances word learning skills and that enhanced word learning skills in turn facilitate vocabulary growth (e.g. Kan & Kohnert, 2012). Vocabulary measures were not available for the children, which is an important limitation of the present study. However, the substantial inter-individual variation in the time the children with a CI had been using the implant predicts variability in spoken language experience and vocabulary size (e.g. Fagan & Pisoni, 2010). Also, it seems safe to assume their vocabulary size was smaller than that of the children with normal hearing. If for typically developing children a growing mental lexicon increases the need for more detailed phonetic distinctions in lexical representations and thus enhances their sensitivity to phonetic detail (e.g. Beckman & Edwards, 2000; Metsala & Walley, 1998), then delays in vocabulary development for the children with a CI in the present study may have contributed to their observed difficulties with learning phonologically minimal pairs. Indeed, a recent study by Walker and McGregor (2013) showed that rapid word learning scores in children with a CI did not differ significantly from those of vocabulary-matched children with normal hearing.

In sum, the present results show that, at age six, the ability to encode phonetic detail in newly established lexical representations has not yet fully developed and appears to be mediated by acoustic salience (or discriminability) of the contrast as well as task demands. Although deaf

333

children with a CI performed better with acoustically more salient vowel minimal pairs than acoustically less salient consonant minimal pairs, they showed substantial difficulty with learning phonologically minimal pairs and did not benefit as clearly from a reduction in task demands as their peers with normal hearing did. In addition to limited perceptual abilities, lower novel word learning performance in children with a CI might be associated with vocabulary delays or inherent difficulties with tasks that involve processing nonwords.

REFERENCES

Baayen, R. H., Davidson, D. J. & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* **59**, 390–412.

Barr, D. J., Levy, R., Scheepers, C. & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *Journal of Memory and Language* **68**, 255–278.

Bates, D. M., Maechler, M., Bolker, B. & Walker, S. (2013). lme4: linear mixed-effects models using Eigen and S4. R package version 1.1–7.

Beckman, M. E. & Edwards, J. (2000). The ontogeny of phonological categories and the primacy of lexical learning in linguistic development. *Child Development* **71**, 240–249.

Bergeson, T. R., Pisoni, D. B. & Davis, R. O. A. (2005). Development of audiovisual comprehension skills in prelingually deaf children with cochlear implants. *Ear and Hearing* **26**, 149–164.

Bonatti, L. L., Peña, M., Nespor, M. & Mehler, J. (2005). Linguistic constraints on statistical computations: the role of consonants and vowels in continuous speech processing. *Psychological Science* **16**, 451–459.

Bouton, S., Serniclaes, W., Bertoncini, J. & Colé, P. (2012). Perception of speech features by French-speaking children with cochlear implants. *Journal of Speech, Language and Hearing Research* **55**, 139–153.

Bowey, J. A. & Hirakis, E. (2006). Testing the protracted lexical restructuring hypothesis: the effects of position and acoustic-phonetic clarity on sensitivity to mispronunciations in children and adults. *Journal of Experimental Child Psychology* **95**, 1–17.

Carreiras, M., Duñabeitia, J. A. & Molinaro, N. (2009). Consonants and vowels contribute differently to visual word recognition: ERPs of relative position priming. *Cerebral Cortex* **19**, 2659–2670.

Curtin, S., Fennell, C. T. & Escudero, P. (2009). Weighting of vowel cues explains patterns of word–object associative learning. *Developmental Science* **12**, 725–731.

Cutler, A., Sebastián-Gallés, N., Soler-Vilageliu, O. & Van Ooijen, B. (2000). Constraints of vowels and consonants on lexical selection: cross-linguistic comparisons. *Memory and Cognition* **28**, 746–755.

Davidson, L. S., Geers, A. E., Blamey, P. J., Tobey, E. A. & Brenner, C. (2011). Factors contributing to speech perception scores in long-term pediatric cochlear implant users. *Ear and Hearing* **32**, 19–26.

Davidson, L. S., Geers, A. E. & Nicholas, J. G. (2014). The effects of audibility and novel word learning ability on vocabulary level in children with cochlear implants. *Cochlear Implants International* **15**, 211–221.

Dietrich, C., Swingley, D. & Werker, J. F. (2007). Native language governs interpretation of salient speech sound differences at 18 months. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 16027–16031.

Duyck, W., Desmet, T., Verbeke, L. P. C. & Brysbaert, M. (2004). WordGen: a tool for word selection and nonword generation in Dutch, English, German, and French. *Behavior Research Methods, Instruments & Computers* **36**, 488–499.

Escudero, P., Mulak, K. & Vlach, H. (2015). Cross-situational learning of minimal word pairs. *Cognitive Science*, online: <doi:10.1111/cogs.12243>.

Fagan, M. K. & Pisoni, D. B. (2010). Hearing experience and receptive vocabulary development in deaf children with cochlear implants. *Journal of Deaf Studies and Deaf Education* **15**, 149–161.

Garlock, V. M., Walley, A. C. & Metsala, J. L. (2001). Age-of-acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults: implications for the development of phoneme awareness and early reading ability. *Journal of Memory and Language* **45**, 468–492.

Gathercole, S. E. (2006). Nonword repetition and word learning: the nature of the relationship. *Applied Psycholinguistics* **27**, 513–543.

Gerken, L. A., Murphy, W. D. & Aslin, R. N. (1995). Three- and four-year-olds' perceptual confusions for spoken words. *Perception and Psychophysics* **57**, 475–486.

Gerrits, E. (2001). *The categorisation of speech sounds by adults and children: a study of the categorical perception hypothesis and the developmental weighting of acoustic speech cues.* Utrecht: LOT Dissertation Series 42.

Giezen, M. R., Baker, A. E. & Escudero, P. (2014). Relationships between spoken word and sign processing in children with cochlear implants. *Journal of Deaf Studies and Deaf Education* **19**, 107–125.

Giezen, M. R., Escudero, P. & Baker, A. E. (2010). Use of acoustic cues by children with cochlear implants. *Journal of Speech, Language and Hearing Research* **53**, 1440–1457.

Havy, M., Bertoncini, J. & Nazzi, T. (2011). Word learning and phonetic processing in preschool-age children. *Journal of Experimental Child Psychology* **108**, 25–43.

Havy, M., Nazzi, T. & Bertoncini, J. (2013). Phonetic processing during the acquisition of new words in 3-to-6-year-old French-speaking deaf children with cochlear implants. *Journal of Communication Disorders* **46**, 181–192.

Hazan, V. & Barrett, S. (2000). The development of phonemic categorization in children aged 6–12. *Journal of Phonetics* **28**, 377–396.

Henderson, L., Weighall, A., Brown, H. & Gaskell, G. (2013). Online lexical competition during spoken word recognition and word learning in children and adults. *Child Development* **84**, 1668–1685.

Hicks, C. B. & Ohde, R. N. (2005). Developmental role of static, dynamic, and contextual cues in speech perception. *Journal of Speech, Language and Hearing Research* **48**, 960–974.

Houston, D. M. & Bergeson, T. R. (2014). Hearing versus listening: attention to speech and its role in language acquisition in deaf infants with cochlear implants. *Lingua* **139**, 10–25.

Houston, D. M., Stewart, J., Moberly, A., Hollich, G. & Miyamoto, R. T. (2012). Word learning in deaf children with cochlear implants: effects of early auditory experience. *Developmental Science* **15**, 448–461.

Jaeger, T. F. (2008). Categorical data analysis: away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* **59**, 434–446.

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

Kan, P. F. & Kohnert, K. (2012). A growth curve analysis of novel word learning by sequential bilingual preschool children. *Bilingualism: Language and Cognition* **15**, 452–469.

Kishon-Rabin, L., Taitelbaum, R., Muchnik, C., Gehtler, I., Kronenberg, J. & Hildesheimer, M. (2002). Development of speech perception and production in children with cochlear implants. *Annals of Otology, Rhinology and Laryngology* **111**, 85–90.

Kuznetsova, A., Brockhoff, P. B. & Bojesen, R. H. (2012). Tests in Linear Mixed Effects Models. R package version 2·0–20.

Lederberg, A. R., Spencer, P. E. & Prezbindowski, A. K. (2000). Word learning skills of deaf preschoolers: the development of novel mapping and rapid word learning strategies. *Child Development* **71**, 1571–1585.

Mani, N. & Plunkett, K. (2008). Fourteen-month-olds pay attention to vowels in novel words. *Developmental Science* **11**, 53–59.

335

Marian, V., Bartolotti, J., Chabal, S. & Shook, A. (2012). CLEARPOND: Cross-Linguistic Easy-Access Resource for Phonological and Orthographic Neighborhood Densities. *PLoS ONE* **7**, e43230.

Mayo, C. & Turk, A. (2005). The influence of spectral distinctiveness on acoustic cue weighting in children's and adults' speech perception. *Journal of the Acoustical Society of America* **118**, 1730–1741.

Metsala, J. L. & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: precursors to phonemic awareness and early reading ability. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89–120). Mahwah, NJ: Lawrence Erlbaum Associates.

Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words: differences between consonants and vowels. *Cognition* **98**, 13–30.

Nazzi, T., Floccia, C., Moquet, B. & Butler, J. (2009). Bias for consonantal information over vocalic information in 30-month-olds: cross-linguistic evidence from French and English. *Journal of Experimental Child Psychology* **102**, 522–537.

Nespor, M., Peña, M. & Mehler, J. (2003). On the different roles of vowels and consonants in speech processing and language acquisition. *Lingue e Linguaggio* **ii**, 221–247.

Nittrouer, S. & Miller, M. E. (1997). Predicting developmental shifts in perceptual weighting schemes. *Journal of the Acoustical Society of America* **101**, 2253–2266.

Peterson, N. R., Pisoni, D. B. & Miyamoto, R. T. (2010). Cochlear implants and spoken language processing abilities: review and assessment of the literature. *Restorative Neurology and Neuroscience* **28**, 237–250.

Pisoni, D. B., Cleary, M., Geers, A. E. & Tobey, E. A. (1999). Individual differences in effectiveness of cochlear implants in children who are prelingually deaf: new process measures of performance. *Volta Review* **101**, 111–164.

Pisoni, D. B., Conway, C. M., Kronenberger, W. G., Horn, D. L., Karpicke, J. & Henning, S. C. (2008). Efficacy and effectiveness of cochlear implants in deaf children. In M. Marschark & P. C. Hauser (Eds.), *Deaf cognition: foundations and outcomes* (pp. 52–101). Oxford & New York: Oxford University Press.

Polka, L. & Werker, J. F. (1994). Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance* **24**, 745–749.

R development core team (2011). R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Online: < http://www.r-project.org >.

Schwartz, R. G., Steinman, S., Ying, E., Mystal, E. Y. & Houston, D. M. (2013). Language processing in children with cochlear implants: a preliminary report on lexical access for production and comprehension. *Clinical Linguistics and Phonetics* **27**, 264–277.

Shatzman, K. B. & McQueen, J. M. (2006). Prosodic knowledge affects the recognition of newly acquired words. *Psychological Science* **17**, 372–377.

Stager, C. L. & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word learning tasks. *Nature* **388**, 381–382.

Svirsky, M. A., Teoh, S.-W. & Neuburger, H. (2004). Development of language and speech perception in congenitally, profoundly deaf children as a function of age at cochlear implantation. *Audiology and Neuro-Otology* **9**, 224–233.

Swingley, D. & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science* **13**, 480–484.

Walker, E. A. & McGregor, K. K. (2013). Word learning processes in children with cochlear implants. *Journal of Speech and Hearing Research* **56**, 375–387.

Werker, J. F. & Curtin, S. (2005). PRIMIR: a developmental framework of infant speech processing. *Language Learning and Development* **1**, 197–234.

Willstedt-Svensson, U., Löfqvist, A., Almqvist, B. & Sahlén, B. (2004). Is age at implant the only factor that counts? The influence of working memory on lexical and grammatical development in children with cochlear implants. *International Journal of Audiology* **43**, 506–515.

336

Woodhouse, L., Hickson, L. & Dodd, B. (2009). Review of visual speech perception by hearing and hearing-impaired people: clinical implications. *International Journal of Language and Communication Disorders* **44**, 253–270.

Xu, L. & Pfingst, B. E. (2008). Spectral and temporal cues for speech recognition: implications for auditory prostheses. *Hearing Research* **242**, 132–140.

Yoshida, K. A., Fennell, C. T., Swingley, D. & Werker, J. F. (2009). Fourteen-month-old infants learn similar-sounding words. *Developmental Science* **12**, 412–418.

APPENDIX

Nonword and familiar word stimuli in the rapid word learning tasks

Nonword pairs	Familiar words
/kɪχ/ – /kiχ/	/døʀ/ 'door'
/χɪk/ – /χik/	/hut/ 'hat'
/tɪχ/ – /tiχ/	/myʀ/ 'wall'
/kɑχ/ – /kaχ/	/wɛχ/ 'road'
/tɑχ/ – /taχ/	/vɔrk/ 'fork'
/tɑt/ – /tat/	/ʝʏrk/ 'dress'
/byk/ – /pyk/	/tɑuw/ 'rope'
/bet/ – /pet/	/trœy/ 'sweater'
/beχ/ – /peχ/	
/fyk/ – /syk/	
/fot/ – /sot/	
/fet/ – /set/	

337