# Perception of speech features by French-speaking children with cochlear implants

**Journal:** *Journal of Speech, Language, and Hearing Research*

**Manuscript ID:** JSLHR-H-10-0330.R2

**Manuscript Type:** Research Article

**Date Submitted by the Author:** 29-Apr-2011

**Complete List of Authors:**
- Bouton, Sophie; Laboratoire de Psychologie Cognitive, Université de Provence
- Sernaïlaes, Willy; Laboratoire de Psychologie de la Perception, Université Paris-Descartes
- Bertoncini, Josiane; Laboratoire de Psychologie de la Perception, Université Paris-Descartes
- Colé, Pascale; Université de Provence, LPC, Departement of Cognitive Psychology

**Keywords:**
- Categorical Perception
- Phonological Features
- Cochlear Implant
- Speech Development
Perception of speech features by French-speaking children with cochlear implants

Abstract

Purpose: The present study investigates the perception of phonological features in French-speaking CI children compared to normal-hearing (NH) children matched for listening age.

Method: Scores for discrimination and identification of minimal pairs for all features defining consonants (e.g. place, voicing, manner, nasality) and vowels (e.g. frontness, nasality, aperture) were measured in each listener. Results: The results indicate no differences in "categorical perception," specified as a similar difference between discrimination and identification between CI children and controls. However, CI children demonstrated a lower level of "categorical precision," i.e. lesser accuracy in both feature identification and discrimination, than NH children, with the magnitude of the deficit depending on the feature.

Conclusions: If sensitive periods of language development extend well beyond the moment of implantation, the consequences of hearing deprivation for the acquisition of categorical perception should be fairly important in comparison to categorical precision because categorical precision develops more slowly than categorical perception in NH children. Our results do not support the idea that the sensitive period for development of categorical perception is restricted to the first 1-2 years of life. The sensitive period may be significantly longer. Differences in precision may reflect the acoustic limitations of the cochlear implant, such as coding for temporal fine structure and frequency resolution.

Keywords: Categorical Perception, Phonological Features, Cochlear Implant, Speech Development
1. Introduction

Previous studies suggest that children with cochlear implants (CI) have less accurate perception of speech sounds than normal-hearing (NH) children (Geers, Brenner & Davidson, 2003; Medina & Serniclaes, 2009; Tye-Murray, Spencer & Gilbert-Bedia, 1995). There are two different possible reasons why CI users have more difficulty perceiving speech sounds than NH listeners. First, CI children have experienced a hearing deprivation period, which might correspond to a sensitive period for language development (Knudsen, 2004), especially for the acquisition of categorical perception (Medina, Hoonhorst, Bogliotti & Serniclaes, in press; Hoonhorst, Medina, Collin, Markessis, Radeau, et al., 2011), and this might create problems for the acquisition of "phonological," i.e. language-specific, features. However, if this period was flexible, and extended beyond the moment of implantation, it might only have reduced consequences for speech perception. Second, there are many differences between electrical and auditory stimulation (Shannon, Zeng, Kamath, Wygonski & Ekelid 1995; Dorman, Loizou, Spahr & Maloff, 2002; Xu & Zheng, 2007), and these differences might lead to less accurate feature perception. Thus, CIs and auditory deprivation give a unique opportunity to further elucidate speech and language development and give us a reason to revisit the classic topic of categorical perception.

1.1. Perceptual development of phonological features

The perceptual development of phonological features is a long-lasting process, which starts before one year of age and ends during adolescence. The pre-linguistic child is endowed with capacities to discriminate universal features irrespective of the language environment. These universal features are selected and combined after exposure to the sounds of a specific language between 6 and 12 months of age (Eilers, Gavin & Wilson, 1979; Rivera-Gaxiola, Silva-Peyrera & Kuhl, 2005; Burns, Yoshida, Hill & Werker, 2007; Hoonhorst, Colin, Markessis, Radeau, Deltenre & Serniclaes, 2009). At the age of one year, the child has
acquired the features that are necessary for perceiving the contrasts between sounds, which are phonologically relevant in his or her linguistic environment. However, as stated by Hoonhorst et al. (2009), theoretical perspectives have changed considerably over the years, emphasizing on learning in recent work rather than the strong innate assumptions in early studies (Burns et al., 2007; Maye, Weiss, & Aslin, 2008; Maye, Werker, & Gerken, 2002; for a review, see Kuhl, Conboy, Coffey-Corina, Padden, Rivera-Gaxiola, & Nelson, 2008).

Currently, there still remains a basic concern about the adaptation of the initial abilities to discriminate all of the phonetic contrasts present in the world languages (Werker & Tees, 1984) to the categories present in the native language. Interestingly, the perceptual development of phonological features is far from complete at the age of one year. Various studies show that the features, which play a role in the linguistic environment, are perceived with progressively greater precision up to the end of childhood and even during adolescence (Zlatin & Koenigsknecht, 1975; Simon & Fourcin, 1978; Krause, 1982; Elliott, Longinotti, Meyer, Raz, & Zucker, 1981; Elliott, Busse, Partridge, Rupert, & de Graaf, 1986; Burnham, Earnshaw & Clark, 1991; Hazan & Barrett, 2000; Medina et al., in press; Hoonhorst et al., 2011). This development of perceptual accuracy is not restricted to auditory speech features. Using different stimulus continua, Hoonhorst et al. (2011) found that categorical precision increased with age for both voicing and facial expressions, though not for colors. These results suggest that the development of categorical precision arises from a general cognitive maturation across different perceptual domains. It should be stressed that in all these studies it is specifically categorical precision, in both the discrimination and identification of features, which progresses during this period. For another property, namely "categorical perception," previous researches showed a different developmental pattern.

Whereas "categorical precision" is a matter of accuracy—it corresponds to the extent to which stimuli are correctly identified and discriminated—"categorical perception" reflects the coherence between discrimination and identification. According to the original definition,
perception is categorical when stimuli can be discriminated only insofar as they are classified into different categories (Liberman, Harris, Hoffman, & Griffith, 1957). Despite the huge number of studies which refer to categorical perception, the concept has proven to be highly resistant to various criticisms and misunderstandings. One frequent misconception is to reduce categorical perception to a sensory discontinuity or “boundary.” While there is little doubt that sensory discontinuities play an important role in the acquisition of categorical perception (for a review see Hoonhorst et al., 2009), categorical perception cannot be reduced to such discontinuities. Categorical perception is a much broader property, which can take place “…in the absence of a boundary effect since there need be no local maximum in discriminability” (Macmillan, Kaplan & Creelman, 1977, p. 453) and can then result from top-down effects through which upper-level categories constrain sensory processing (for brain evidence see e.g. Dufor, Serniclaes, Sprenger-Charolles, & Démonet, 2007). Categorical perception arises as a result either of sensory or of higher-level cognitive factors, but what remains essential for its definition is the coherence between sensation and cognition, i.e. the relationship between discrimination and identification responses (the “acid test” of categorical perception, in the words of Damper & Harnad, 2000). The link between sensation and cognition does not necessarily need to be acquired as evidenced by instances of categorical perception induced by psychophysical discontinuities (for data with a rise time continuum: Cutting & Rosner, 1974; for data with monkeys: Kuhl & Miller, 1975; 1978; for data with colors: Hoonhorst et al., 2011). However, when dealing with phonological features, such psychophysical discontinuities are not simply related to the language categories: categorical perception is acquired by combining different basic discontinuities through exposure to the sounds of the environmental language during the first year of age (e.g. Hoonhorst et al., 2009).
None of the previous studies on this issue, as far as we know, have evidenced a change in categorical perception after two years of age. However, many studies have shown that categorical perception changes before one year of age (starting with Werker & Tees, 1984; for a review see Kuhl et al., 2008). Human infants are born with predispositions for perceiving all the possible universal features, which are then activated or not as a function of the presence or absence of the corresponding features in the linguistic environment. For example, Werker & Tees (1984) observed that infants aged 6 to 8 months discriminated non-native (Thompson and Hindi) and native features (English), whereas infants aged 10 to 12 months discriminated only native features. Different universal features can thus be combined whenever they contribute to the perception of the same phonological feature in the language.

It is noteworthy that these studies have assessed categorical perception without collecting identification data instead using differences in discrimination between across- vs. within-category pairs. This is a simplification, which can be useful for assessing categorical perception when identification responses cannot be collected (typically in very young children, since the seminal study of Eimas, Siqueland, Jusczyk, & Vigorito, 1971, or in cross-linguistic studies, as in Strange & Dittmann, 1984). Although the optimal procedure is to collect identification responses, the location of the identification boundary can then be inferred from studies with other subjects (from adults in studies on children; from native-speakers in cross-linguistic studies). However, it should never be forgotten that using the boundary location for separating stimulus pairs into within- and across-category groups is not the optimal procedure; it is only a short-cut when no identification responses are available. No boundary is needed with the optimal procedure, only identification responses.

Categorical perception is perfect when speech sounds cannot be discriminated unless they are identified as members of different categories (Liberman et al., 1957). Although categorical perception is almost never perfect, the acquisition of adult-like categorical perception of language-specific features seems to start very early, before one year of age, but
all existing evidence on this point is indirect, as discussed above. There is only evidence to support that the relationship between discrimination and identification does not improve after some six years of age (Medina et al., in press; Hoonhorst, et al., 2011). This does not mean that categorical perception is perfect in six-year-old children, i.e. that discrimination between speech sounds at this age depends entirely on their classification into different categories, but that their degree of categorical perception is not less than that of adults with the same stimuli and more generally under the same conditions.

The abundant evidence on the long-lasting effect of age on categorical precision, up to the beginning of adolescence, is in sharp contrast with the presence of adult-like categorical perception performances at two years of age and perhaps even earlier. There is no doubt that categorical perception develops faster than categorical precision, although the precise age at which adult-like categorical perception is acquired remains unknown. The difference in the rate of development of these two properties of categorical behavior might have implications for language development in CI children. However, this depends on the flexibility of the sensitive periods during which these properties can be acquired. We share the views of several other authors that this is a concern for the development of speech in general (Werker & Tees, 2005; Kuhl et al. 2008, p.993; Gervain & Mehler, 2010) and with a CI in particular (Kral & Eggermont, 2007). For example, richer social environments extend the duration of the sensitive period for learning in owls and songbirds (Baptista & Petrinovich 1986; Brainard & Knudsen 1998). According to Werker and Tees (2005), biological and experiential factors contribute to variations in both the onset and offset of openness to speech perception experience. While this indicates that sensitive periods are fairly flexible, the sensory deprivation, which precedes cochlear implantation in infants who are congenitally deaf might nevertheless affect the development of speech perception, especially that of categorical perception whose development is relatively rapid. Hearing deprivation should have a lesser
effect on the development of categorical precision, which is fairly slow but the rate of development depends on the flexibility of the periods during which these categorical properties can be fully acquired. If these periods extend well beyond the moment of implantation, the consequences of hearing deprivation for the acquisition of categorical perception capacities should be fairly limited.

Although cochlear implantation in deaf infants has become a promising tool for auditory rehabilitation, the specific implications of the auditory deprivation period preceding cochlear implantation on speech perception remain to be elucidated. As explained above, categorical perception tells us about the coherence of speech perception. The aim of the present article is to study the implications of auditory deprivation for the development of two basically different properties of speech sound categorization, categorical perception and categorical precision.

1.2. Overview of consonant and vowel recognition in CI children

In general, perception of speech sounds is less precise in CI children than in NH children. Geers et al. (2003) administered a battery of speech perception tests to CI children of eight and nine years old who had received a cochlear implant by age five. They found that the children achieved an average level of about 50% speech perception skills through listening alone, although their performance depended on the difficulty of the test. Indeed, on an early speech perception test developed by the CID (Central Institute for the Deaf, St Louis, Missouri), in which the child hears a word and responds by pointing to one of 12 pictures on a single plate following each stimulus presentation, CI children obtained about 79% correct responses. This task, which uses words differing in all phonemes or with a similar first consonant, is simpler than the WIPI test (Word Intelligibility by Picture Identification test, Ross & Lerman, 1971), which is also a word identification task with picture pointing
responses. In this test, only one auditory stimulus word is presented for each plate, where there are 6 pictures corresponding to words with the same or similar vowel sound but different consonants (for example, hair, pear, stair). In the WIPI task, CI children correctly recognized about 41% of the words (Geers et al., 2003). Thus, tests that require fine auditory distinctions among consonant sounds appear to be more difficult than the early speech perception test where words are more broadly distinguished.

Other results show that the magnitude of the perceptual deficit in CI children depends on the phonological feature. Tye-Murray, Spencer and Gilbert-Bedia (1995) assessed the speech perception skills of CI children using the Children’s Audio-visual Feature Test (Tyler, Fryauf-Bertschy & Kelsay, 1991). The test is comprised of items that most 4- and 5-year-old children can recognize: seven letters of the alphabet and three common words. The stimuli are P, T, C, B, D, V, Z, key, knee, and me. When testing perception skills, an audiologist pronounced one word and the subject pointed to an item on the response form that bore an illustration of the test stimulus. They showed that children fitted at 7.3 years, with 34 months of CI experience at the moment of testing, had a score of 25% correct consonants in the auditory-only condition. The children scored highest for the nasality and voicing features (respectively 33% and 30% of correct responses on average) and lowest for the place and frication features (respectively 7% and 13% of correct responses on average).

Medina (2008) and Medina and Serniclaes (2009) showed that CI children aged from 5.9 to 11 years old and fitted at an average age of 3.3 years discriminated different vowels and consonants less successfully than NH children matched for chronological age. For vowels, the deficit was larger for nasality than for labiality, frontness and aperture. For the consonants, the deficit was larger for place than for nasality, manner and voicing, in agreement with the results of Tye-Murray et al. (1995).

The results of a previous study suggested that in some instances CI children do not have a deficit in either categorical perception or categorical precision when compared to NH
children of the same listening age (Medina & Serniclaes, 2009). However, this depended on
the phonological feature. Similar results between CI and NH children were found only for the
voicing feature (Medina & Serniclaes, 2009) and not for place of articulation, for which CI
children showed a deficit in categorical precision (Medina & Serniclaes, 2005). The absence
of a categorical perception deficit for the voicing feature suggests that early auditory
deprivation does not affect the development of phonological features. The fact that a
categorical precision deficit is present for the place feature but not for the voicing feature
suggests that cochlear implants’ poor transmission of certain acoustic cues such as, formant
transitions, has developmental implications.

The fact that deficits in phonemic discrimination are not equivalent for all the
investigated features suggests that the different acoustic cues are not transmitted with the
same fidelity by the cochlear implant device. These limitations may be a problem because
each phonemic feature depends on specific acoustic cues1. Studies on this issue started from
the dichotomy between two different dimensions of acoustic signals: temporal envelope and
temporal fine structure. For instance, envelope has decisive importance for the perception of
voicing and manner, whereas temporal fine structure is decisive for place and nasality (Rosen,
1992; Stevens, 2000). If envelope was better transmitted by the CI than temporal fine
structure, this might explain why CI children perceive manner and voicing better than place
and nasality. In general, the capacity to identify speech signals relies on the perception of both
temporal cues, the envelope (i.e. E, relatively slow variations in amplitude over time, between
2 and 50 Hz), and the temporal fine structure, (i.e. TFS, the rapid variations at a rate close to
the center frequency of the band, higher than 500 Hz). Thus, it is theoretically and practically
important to understand whether the perceptual resolution of temporal envelope and temporal
fine structure are key factors for speech perception in quiet for CI users. To our knowledge,
no data is available concerning the acoustic cue perception skills of children with cochlear
implants. All the research on this issue discussed below concerns adult listener skills.
Most current cochlear implant prostheses seem to deliver adequate temporal envelope information, but transmit temporal fine structure suboptimally. Existing types of cochlear implants cannot provide temporal fine structure information, as the modulated carrier in each channel is a single train of pulses with fixed delay, and the ability of the cochlear implant to use frequency-place coding cues is limited (Faulkner, Rosen & Moore, 1990). Indeed, large amounts of channel interaction have been observed for stimuli presented on adjacent electrodes, although the overall effect is individually variable and dependent on the stimulus parameters (Chatterjee & Shannon, 1998). The device’s discarding of temporal fine structure significantly decreases speech perception performance, especially in tonal languages (Xu & Pfingst, 2008; Stickney, Zeng, Litovsky, & Assman, 2004). Whereas some studies have shown that CI users’ temporal fine structure resolution is relatively poor compared to that of NH listeners, their processing of temporal envelope is less impaired (Fu & Shannon, 2000; Chen & Zeng, 2004; Stickney et al., 2004). Temporal envelope cues may, in fact, be particularly important for CI users precisely because their perception of temporal fine structure cues is significantly worse than that of listeners with normal hearing, and they may be forced to rely more heavily on those acoustic cues that are relatively well perceived with a cochlear implant (Zeng, Nie, Stickney, Kong, Vongphoe, Bhargave et al., 2005; Qin & Oxenham, 2003). Indeed, many studies have shown that temporal envelope information contributes significantly to consonant recognition (Xu, Thompson & Pfingst, 2005). Nie, Barco and Zeng (2006) performed a study with CI adults and showed that percent correct scores for consonant recognition vary with stimulation rate. Consonant recognition in quiet improved as the rate of stimulation increased.

The temporal envelope/temporal fine structure dichotomy does not seem to be as useful for studying the discrimination of vowel structure, which is mainly characterized by formants. Vowel perception relies on spectral resolution to resolve the formant pattern that distinguishes between vowels. Frontness is mainly defined by F2 and F3, aperture by F1 and
F2, and nasality by a special low F1 formant, i.e. a second low frequency resonance in the vicinity of F1, and antiformants (Calliope, 1989). The reliance of frontness on relatively high frequencies, in the F2 and F3 regions, and of nasality on lower frequencies, in the F1 region, could explain the differences in reception between these features. This possibility is supported by the suggestion of Harnsberger, Svirsky, Kaiser, Pisoni, Wright and Meyer (2001) that vowel perception by CI users may be limited by the listener’s formant frequency discrimination skills. Thus, differences in the reception of the different vowel features seem to be related to insufficient frequency resolution rather than to differences in the balance between temporal envelope and temporal fine structure cues.

In summary, cochlear implant devices’ superior transmission of temporal envelope cues as opposed to temporal fine structure cues might explain the unequal reception of different consonant features, namely the better reception of manner and voicing as opposed to nasality and place. However, differences between vowel features, namely the better reception of frontness than nasality, seem to be better explained by insufficient frequency resolution.

1.3. The present study

In summary, previous studies provide some evidence that child CI users have a deficit in feature perception when compared to NH children. However, there are no clear indications about the nature of this deficit. While previous studies suggest a deficit in the precision of the processes involved in feature discrimination and identification, little is known about a possible deficit in categorical perception, i.e. a possible deficit in the relationship between discrimination and identification. Although previous results suggest that categorical perception of the voicing feature is not impaired in CI children (Medina & Serniclaes, 2009), the question of whether this finding can be generalized to the other vowel and consonant features remains open. As previous results also suggest a fairly general deficit in categorical precision in CI children, one might wondered whether their categorical perception of the
different consonant and vowel features is preserved. The present study addresses these issues for the main phonological features in French (place, voicing, nasality and manner for consonants; nasality, aperture and frontness for vowels). Finally, the need to investigate and compare the ability to categorize speech sounds in young CI users is reinforced by the repeated demonstration of the important role of categorical perception in reading acquisition (Bogliotti, Serniclaes, Messaoud-Galusi, & Sprenger-Charolles, 2008; Serniclaes, Van Heghe Mousty, Carré & Sprenger-Charolles, 2004; Godfrey, Syrdal-Lasky, Millay & Knox, 1981).

The aim of the present study is to obtain a comprehensive picture of the perception of different consonant and vowel features by children with cochlear implants. We questioned whether the absence of categorical perception deficits evidenced in a previous study (Medina & Serniclaes, 2009) with a limited set of features might be generalized to the whole set of features present in a language. We also aimed to determine whether the severity of categorical precision deficits varies between features, in order to draw possible links with cochlear implants’ transmission of different speech cues. Two different hypotheses were tested:

Hypothesis 1. In accordance with previous results, CI children will not exhibit a lesser degree of categorical perception than NH children, in spite of a hearing deprivation period. We used the classical procedure for assessing categorical perception, which consists in converting the identification scores into “expected” discrimination scores—i.e. the amount of discrimination, which would be expected from their classification into different categories—before comparing them with the observed discrimination scores.

Hypothesis 2. CI children will exhibit a lesser degree of categorical precision than NH children, with the degree of categorical precision deficit depending on the phonological feature. The assessment of categorical precision was based on the mean values of the discrimination and identification scores (whereas categorical perception assessment was based on the difference between the discrimination and identification scores).
1.4. A simplified methodology for assessing categorical perception

Assessing the categorical perception of seven different features would be a formidable task with the usual methodology. Instead of using stimulus continua, we used minimal word pairs for assessing categorical perception. This simplified method is quite different from the usual one but, we believe, is entirely appropriate for measuring the degree of categorical perception.

As we have seen, what is essential for the definition of categorical perception is the relationship between identification and discrimination. With the classical definition of categorical perception (Liberman et al., 1957), there is no need to separate within- and across-category pairs of stimuli. In fact, within- and across-category pairs are not entirely separable unless there is a stepwise identification function (i.e. there is a jump from 0 to 100% between two successive stimuli) along some stimulus continuum. Without stepwise identification, all the stimuli belong at various degrees to both categories, and there is no clear-cut separation between within- and across-category pairs. Each pair can be considered as both across and within category pair. Liberman et al. (1957) argue categorical perception assessment does not need to segregate the within-category pairs from across-category ones: the classical method for assessing the degree of categorical perception is to compare the actual discrimination performance (observed discrimination) with the one afforded by differences in identification (expected discrimination) for each pair. The expected discrimination afforded by each single pair depends on the across- vs. within-category balance inside the pair (i.e. when both stimuli in a AX pair collect either 0 or 100 % identification scores there is 0% expected discrimination, and the percentage of expected discrimination increases with the difference between the identification scores of the stimuli inside the pair).

With this classical definition there is also no need for a stimulus continuum to assess categorical perception. A single pair is sufficient to compare discrimination and identification. Although categorical perception experiments have always (as far as we know) used pairs of
stimuli varying along some continuum, this was due to the reduction of categorical perception to sensory discontinuities and is by no means compulsory. Keeping in mind that each pair is a mixture of across-category and within-category pair, a single pair can already give sufficient information for comparing discrimination and identification performance. We do not need a continuum and a boundary. All that is needed is an ambiguous pair to keep performance from being asymptotic (ceiling effects). Therefore, to assess categorical perception with minimal pairs of natural speech stimuli as we did in the present study is entirely coherent with the definition of categorical perception, although as far as we know this is a completely new procedure. By comparison with the usual “continuum” paradigm, this new “minimal pairs” paradigm is much simpler because it makes it possible to considerably reduce the size of the experiment. This is an important advantage for research studies of broad scope like this one, and also for clinical studies.

2. Methods

The present study was designed to investigate CI children’s capacity to discriminate and to identify meaningful minimal pairs. Their performance was compared to that of NH children matched for listening age\(^3\). The stimuli were words and minimal pairs, comprising all the vowel and consonant features of French.

2.1. Participants

Twenty-five children with unilateral cochlear implants (11 boys and 14 girls) were recruited from 15 French institutes for the deaf located in different regions of France. All the children were congenitally deaf, had used a cochlear implant device for at least 5 years, and had been fitted with an implant before the age of 3;6 years. The implant was either the Clarion\(^4\) (Advanced Bionics), the Nucleus\(^5\) (Cochlear Corporation) or the Digisonic\(^6\) (Neurelec). Children recruited ranged from 7;11 to 11;6 years, and from grades 2 to 4. Age at implantation ranged in age from 1;10 to 3;6 years. Only one child had deaf parents. Table 1
describes the characteristics of each of the participants with cochlear implant. Before implantation, all children used conventional hearing aids and were still using them—even if only occasionally—in their non-implanted ear. During the speech perception tests, the hearing aid was taken off in order to prevent any sound from being perceived by the contralateral ear.

Before and after implantation, 9 out of 25 children used cued speech (early and intensive practice), and 11 children out of the remaining children used only spoken language, i.e. they exclusively used speech and audition to communicate. Five children had used spoken language and cued speech starting in grade 1. In our study, 19 out of 25 children were enrolled in mainstream classes with hearing children. Six children were in a spoken language classroom in a deaf school (special education with spoken language instruction).

Each child with CI was matched with one NH child with the same listening age to compare the performance of the two groups. All the NH children met the following criteria: (a) they were native speakers of French, and (b) they had no history of auditory and language disorders. All families, both those of CI children and those of NH children, were informed about the goals of the study and provided written consent before the participation of their child.

As indicated in Table 2, the chronological age of the CI group was significantly higher than that of the NH group ($t(25)=9.8; p<.001$). The listening age of the CI group and NH controls did not differ significantly ($t(25)<1$). Only children whose nonverbal cognitive development was considered to be within normal limits using a nonverbal reasoning test from Progressive Matrices (PM47, Raven, 1947) participated in the study. As indicated in Table 2, the scores of all children were within the normal range.
2.2. **Experimental tasks**

Two tasks were administered in order to assess subjects’ discrimination and identification skills. Possible differences between groups in categorical perception must be dissociated from differences in the precision of phonemic boundaries, the latter being assessed by the magnitude of identification scores (the better the d’ score, the higher the precision). In order to assess categorical perception while ruling out effects due to differences in categorical precision, we used the classical criterion of comparing the observed discrimination scores to those predicted from the identification responses (Liberman et al., 1957). Categorical perception is perfect when the observed discrimination scores coincide exactly with those expected from identification; the degree of categorical perception is assessed by the difference between the observed and expected discrimination data. The smaller the difference between the observed and predicted scores, the better the categorical perception would be.

According to the original definition (Liberman et al., 1957), observed discrimination scores correspond to actual discrimination responses (e.g. the correct discrimination scores of d’ transforms of stimulus pairs in an AX discrimination task), whereas expected discrimination scores are derived from identification responses (e.g. the identification scores of the same stimuli presented individually).

For all the tests, the children could potentially use lip-reading to understand instructions, but the items were pre-recorded to ensure that they listened to the stimuli without the help of lip-reading. This precaution is important because lip-reading has been shown to have an effect on speech perception in both deaf children (O’Donoghue, Nikolopoulos & Archobold, 2000; Bergeson, Pisoni & Davis, 2003; Bergeson, Pisoni & Davis, 2005; Lachs, Pisoni & Kirk, 2001; Leybaert & Colin, 2007; Spencer & Tomblin, 2009; Tyler, Fryauf-Bertschy, Kelsay, Gantz, Woodworth & Parkinson, 1997) and normal-hearing children (McGurk & MacDonald, 1976).

**Discrimination task.** Measures of speech feature discrimination for both consonants
and vowels were obtained for both groups of subjects. This test used a two-alternative forced choice procedure. The stimuli were presented in pairs (AX format), comprising either different stimuli (e.g. for the /b/-/m/ contrast, “mouche”-“bouche” or “bouche”-“mouche”) or the same stimuli presented twice (either “bouche” or “mouche” twice in a row). Pairs of stimuli were thus, for example, “mouche”-“bouche”, “mouche”-“mouche”, “bouche”-“bouche”, and “bouche”-“mouche”. The child heard the two words spoken successively, with a 100-ms interval between them. They had to indicate whether the stimuli within each pair were the same (either “bouche” or “mouche” twice) or different (e.g. “mouche”-“bouche” or “bouche”-“mouche”). There were 140 different pairs (70 stimulus conditions x 2 orders) and 140 same pairs.

We used one list of CVCV or CVC words pairs recorded by a female French speaker. This list assessed all features defining consonants (e.g. place, voicing, manner, nasality) and vowels (e.g. frontness, nasality, aperture). Pairs of stimuli were, for example:

- voicing: “poule”-“boule”;
- consonant nasality: “bouche”-“mouche”;
- manner: “vol”-“bol”;
- place of articulation: “bouche”-“douche”;
- frontness: “pull”-“poule”;
- aperture: “cil”-“salle”; and
- vowel nasality: “baton”-“bateau”.

The words chosen were drawn from Manulex (Lété, Sprenger-Charolles, Colé, 2004) and are very frequent, occurring one and a half times in every 100 words, which corresponds to a standard frequency effect of 85 (MANULEX, Lété et al., 2004). However, it should be kept in mind that the acoustic correlate of a phonological feature varies as a function of the context of the adjacent and simultaneous features. For instance, voicing cues change as a function of the vocalic aperture (an adjacent feature) and the mode of articulation (a
simultaneous feature). However, the major cues remain the same across contexts despite changes in their relative perceptual weightings (negative and positive VOT for voicing; noise spectrum and formant transitions for place of articulation). Our word list is, thus, not entirely balanced on acoustic grounds, although the imbalance affects only the relative weightings of these major cues.

Discrimination responses were given by pressing one of two differently colored keys on the computer keyboard. A green button was used to indicate that the two stimuli in a pair were the same, and a red button to indicate that they were different.

**Identification task.** The items used in this task were the same as in the discrimination task. One word was presented on each trial (example: “mouche” or “bouche”), and children had to indicate whether the picture presented on the computer screen was the same as or different from the spoken word. Word-picture pairs included, for example: the word “mouche” (fly) with a picture of a fly, the word “bouche” (mouth) with a picture of a mouth, the word “mouche” (fly) with a picture of a mouth, and the word “bouche” (mouth) with a picture of a fly. Two of them corresponded to the response “same,” and 2 to the response “different.” There were 140 different pairs (70 stimulus conditions x 2 orders) and 140 same pairs. The colored button responses were the same as in the discrimination task.

This picture-matching task required participants to compare stored representations of two stimuli, a picture and a word. This task cannot be performed by comparing the stimuli (the response would inevitably and always be “different”) but only by comparing the associated labels. Discrimination on this task is thus entirely constrained by labels, which is exactly what would be expected in the case of perfect categorical perception. Therefore, discrimination performance is functionally equivalent to identification performance and, when compared with performance on unconstrained discrimination of spoken words, is perfectly suited to assessing categorical perception.
2.3. Procedure

CI and NH listeners were tested individually in a quiet room (at home and at school, respectively). They received all features (place, voicing, manner, and nasality for consonants; frontness, nasality and aperture for vowels) during six sessions, which lasted around 20 minutes each. For all groups, both identification and discrimination tasks were presented in a random order and all AX pairs were presented only one time each in random order. In both tasks, we assessed the features, which define consonants (e.g. place, voicing, manner, nasality) and vowels (e.g. frontness, nasality, aperture).

Presentation was controlled by E-prime 2.0, running on a Dell PC. Stimuli were presented at an individual adapted comfortable level through headphones (Beyerdynamic DT290). The DT 290 is a double sided headphone headset with soft and large circumaural ear cups which cover ear and microphone. Before the test session, we first ensured that CI children were able to hear the stimuli. Children listened to several lists of 10 very frequent words (examples: “chat”-“cat”, “jambe”- “leg”, “petit”-“little”, “maison”-“house”, “enfant”-“child”, “rat”-“rat”, “ours”-“bear”, “rouge”- “red”, “vélo”-“bike” and “neige”-“snow”) at a comfortable level through headphones (70 dB SPL). They had to repeat one list and we considered that the hearing level of the CI children was sufficient if they repeated at least 80% of the words correctly. When they repeated less than 80% of the words correctly, another list of words was presented again with the level increased by 5dB. The stimuli were presented at 75 dB for 19 CI children and at 80 dB for 6 others CI children.

2.4. Data Analysis

The discrimination and identification responses collected for each of the seven feature contrasts, four for consonants and three for vowels, were converted into d’ scores by taking the difference between the normal deviate (z-value) corresponding to the proportion of correct change detection (i.e., the proportion of "different" responses to the "different" pairs or
"hits") and the proportion of false alarms (i.e., the proportion of "different" responses delivered in no change trials; as in Hary & Massaro, 1982; see equation 1). As 0% and 100% scores correspond to infinite z-values, response scores were adjusted following the classical procedure described by Macmillan & Creelman (2005) before conversion into z-values. Response scores above 50% were reduced by 1.25% (with 40 responses per couple of pairs, 1.25% corresponds to one half of the precision of the response scale, i.e. one half of 1/40), and those below 50% were increased by 1.25%. Finally, d’ on discrimination and identification tasks was computed for each participant and each feature.

According to the original definition, there is categorical perception when discrimination of stimuli varying along some continuum depends on identification (Liberman et al., 1957). Categorical perception is perfect when observed discrimination scores coincide with those expected from identification, and the degree of categorical perception is characterized by the difference between observed and expected discrimination scores. Observed discrimination scores correspond to actual discrimination responses (e.g. the correct discrimination scores, or d’ transforms of stimulus pairs in an AX discrimination task) whereas expected discrimination scores are derived from identification responses (e.g. the identification scores of the same stimuli presented individually).

The expected and observed discrimination scores were converted into d’ values using the same procedure and the d’ scores were entered into a repeated-measures analysis of variance (ANOVA). The design of the ANOVA run on the discrimination and identification scores comprised Task (observed vs. expected) and Phonetic Feature (voicing vs. consonant nasality vs. manner vs. place vs. aperture vs. frontness vs. vowel nasality) as within-subjects factors and Group (cochlear implant vs. normal hearing) as a between-subjects factor.

Differences in categorical perception between groups were tested with the Task x Group and Task x Group x Feature interactions. Differences in categorical precision between groups were tested with the Group effect and with the Group x Feature interaction.
Differences in categorical perception and categorical precision between features were tested with planned comparisons.

3. Results

Response Accuracy

- Figure 1 bout here -

Figure 1 presents mean discrimination scores, both observed and expected from the identification responses (expressed in d' values), for each group and all features taken together. Figure 2 presents the mean d' for each group and each task for each of the seven different phonetic features tested (consonants in the top panel, vowels in the bottom panel). These results were tested with a Task x Feature x Group repeated measures ANOVA. Among the main effects, significant effects were found for Group (F(1,48) = 47.2; p<.001), Task (F(1,48) = 35.3; p<.001), and Phonetic Feature (F(6,288) = 6.64; p<.001). Among the interactions, only the Group x Feature interaction was significant (F(6,288) = 3.36; p<.01). The group effect was due to the NH children’s greater observed and expected scores compared to the CI children. The Task effect indicates that categorical perception was not perfect: this was due to greater observed than expected discrimination scores (see Figure 1). However, the difference between the expected and observed discrimination scores was quite similar for both groups (as evidenced by the parallel lines in Figure 1, and either parallel or nearly parallel lines in Figure 2). Neither the Group x Task interaction nor the Group x Task x Feature interaction was significant (F<1; F(6,288) = 1.08; p =.20, respectively) indicating that there was no significant difference in categorical perception between the CI children and NH controls. The Group x Feature interaction was due to larger differences in both the expected and observed discrimination scores for some features, especially consonant and vowel nasality (see Figure 2).
In Figure 3 we present the differences between the two different groups (CI vs. NH) in mean d’ scores for both identification and discrimination tasks, and for the seven different features under study (nasality, place of articulation, manner, and voicing for consonants; nasality, frontness and aperture for vowels).

We observed that the mean difference between groups for each feature is different from chance (Place: t(24) = 8.4 ; p<.001; Consonant Nasality: t(24) = 6.4 ; p<.001; Voicing: t(24) = 1.2 ; p<.001; Manner: t(24) = 1.3 ; p<.001; Aperture: t(24) = 1.4 ; p<.001; Vowels Nasality: t(24) = 1.9 ; p<.001; Frontness: t(24) = 1.1 ; p<.001). For each feature, the accuracy scores of CI children are lower than those of NH children.

Figure 3 shows that, for consonant features, the differences between groups are larger for nasality and place than for voicing and manner. Newman-Keuls comparisons indicated significant differences between nasality and manner (p<.05), place and manner (p=.05), nasality and voicing (p<.01), and place and voicing (p<.05), but no significant difference between nasality and place (p>.20) or manner and voicing (p>.20). This is compatible with the results of previous study suggesting a larger deficit for nasality and place compared to manner and voicing (Tye-Murray et al., 1995). For vowel features, Newman-Keuls comparisons indicated significant differences between nasality and aperture (p<.05) and nasality and frontness (p<.05), but not between aperture and frontness (p>.20). Thus, there was a larger deficit for nasality compared to aperture and frontness. Previous study suggested that the deficit should also be larger for aperture vs. frontness, a difference, which was present as a trend in our results (see Figure 3), but was not significant (Medina & Serniclaes, 2009).

In summary, these results show that for both the CI children and NH controls the observed discrimination scores were larger than those expected from identification. This
means that categorical perception was not perfect. The magnitude of the difference between the two tasks did not depend on the group, which means that CI children and NH children matched for listening age show equivalent categorical perception skills. However, d’ for both identification and discrimination were lower for CI children than for NH children, indicating a lower level of categorical precision. Differences between features showed that for consonants, CI children have more difficulty perceiving nasality and place than manner and voicing; for vowel features, their difficulties are greater for nasality than aperture and frontness.

4. Discussion

The present study indicates that the categorical perception performance of children with cochlear implants is much the same as that of normal-hearing children for a fairly large set of different phonological features. The differences between observed and expected discrimination scores were fairly similar for CI and NH children, suggesting that discrimination was equally constrained by identification in both groups. However, CI children obtained lower discrimination and identification scores than NH children matched for listening age, indicating a lesser degree of categorical precision. Finally, this categorical precision deficit was larger for some features than for others. The differences between groups in the precision of consonant perception were the smallest for manner and voicing and the largest for nasality and place. For vowel perception, differences in precision between groups were smallest for frontness and aperture, and largest for nasality.

4.1. Categorical perception vs. precision

Differences in categorical precision between NH and CI children, as well as the fact that the magnitude of the deficit depends on the feature, have already been evidenced in previous studies (Tye-Murray et al., 1995; Medina & Serniclaes, 2009). There have also been some indications as to the absence of differences in categorical perception with synthetic
stimuli varying along a VOT continuum (Medina & Serniclaes, 2009). The present results make it possible to generalize these findings to a range of different consonant and vowel features in natural stimuli within a single study.

The present results indicate that the observed discrimination scores of both groups were larger than predicted from identification data. Categorical perception was thus not perfect, presumably due to the fact that discrimination can be based on all the acoustic differences between two sounds, whereas identification must be based on the particular acoustic differences that are relevant for making phonological distinctions. The important point for our purposes here is that the difference between discrimination and identification scores is similar in both groups, indicating that both groups exhibit the same degree of categorical perception. As explained in the Introduction, categorical perception depends on both sensory and cognitive factors, which jointly contribute to the coherence of identification and discrimination responses. In the course of perceptual development, categorical perception arises from the adaptation of universal boundaries (language-independent sensory discontinuities) to the phonological categories present in the environmental language (for review see: Hoonhorst et al., 2009). This is a very complex process through which fairly rigid sensory discontinuities are integrated with other sensory cues and become much more flexible after exposure to the sounds of a particular language. The end-product is often far from being perfect: the sensory discontinuities can still act independently in various conditions, giving rise to discrepancies between discrimination and identification performance, i.e. to different degrees of non-categorical perception (Medina et al., in press). However, what is most significant to us is that the degree of non-categorical perception was not larger in CI children compared to NH children, indicating that they possess unimpaired phonological integration capacities.

Our data thus suggest that the initial potential for acquiring categorical perception remains fairly intact at the end of the deprivation period. The fact that we did not find any
For Peer Review

evidence for differences in categorical perception between CI children and controls is even more striking in light of the fact that there were significant differences in precision. Indeed, CI children obtained lower scores than NH children for each feature, indicating that the development of categorical precision in CI children is slower than in NH children. At first sight, the absence of a categorical perception deficit in CI children despite the persistence of a categorical precision deficit is rather surprising in light of results on speech development in NH children. Categorical perception has already been acquired at six years of age, if not earlier, whereas adult-like categorical precision performance is only attained at the beginning of adolescence (see Introduction; Medina et al., in press; Hoonhorst et al., 2011). If the deficit in precision was due to the absence of auditory stimulation before the implantation, the effect of auditory deprivation should be even larger on categorical perception because the latter property develops faster than precision. Therefore, early auditory deprivation does not seem to explain the precision deficit of CI children, nor can it explain the absence of a categorical perception deficit, probably because the sensitive periods for acquiring these properties are fairly flexible. This is in accordance with other studies, which have found that an early deprivation period does not influence the development of the auditory pathway and the central auditory system if children receive a cochlear implant before 3;6 years of age (Ray, Gibson & Sanli, 2004; Sharma, Dorman & Kral, 2005).

As CI children’s lower precision in the perception of phonological features does not seem to be the consequence of auditory deprivation before cochlear implantation, the explanation has to be sought elsewhere. Limitations in the transmission of acoustic cues by the cochlear implant might offer some insight on this issue.

4.2. Differences between features

While the present results suggest that the categorical precision deficit of CI children compared to NH children is not likely explained by the effects of an auditory deprivation
period, the limitations of signal processing by the cochlear implant seem to be more relevant. Differences in categorical precision between NH and CI children might arise from differences between electrical and auditory stimulation. Indeed, different studies have shown that acoustic cues are transmitted with lesser accuracy by CI devices than by normal auditory stimulation (Dawson, McKay, Busby, Grayden & Clark 2000; Dorman & Loizou, 1997; Loizou, Dorman & Tu, 1999; Loizou, Poroy & Dorman, 2000). Furthermore, previous studies have demonstrated differences in the magnitude of perceptual deficits between phonological features (Tye-Murray et al., 1995; Medina & Serniclaes, 2009). The results of the present study confirm these findings and provide an overall picture of the reception of the different features in relation to the corresponding acoustic cues.

The present results reveal that, for consonant features, manner and voicing are better perceived by children using cochlear implant than nasality and place. These patterns of information reception are globally consistent with previous work on adult acoustic perception skills. The ordering of feature reception (manner = voicing > nasality = place) is reminiscent of the ordering of feature reception in Tye-Murray et al.’s (1995) study. The fact that we obtained much the same ordering some 15 years later, despite considerable changes in processing strategies and sound processor technology, is worth highlighting. The simplest explanation for this ordering is that voicing and manner information are mainly conveyed by E cues, whereas place and nasality are mainly conveyed by TFS cues, as would be anticipated from acoustic and phonetic considerations (Rosen, 1992; Verschuur & Rafaely, 2001). However, a recent study suggests that for normal-hearing listeners, the degradation between E and TFS speech is lesser for voicing and nasality than for manner and place (Bertoncini, Serniclaes & Lorenzi, 2009). Clearly, the relationship between each feature and the E vs. TFS cues needs further investigation and might be of relevance to the transmission of consonant features by cochlear implants. It would be worthwhile quantifying individual differences in the acoustic limitations of the CI devices through psychoacoustic measures of sensitivity to
temporal fine structure and envelope cues (as in: Won, Drennan, Kang & Rubinstein, 2010)
and to relate them to differences in categorical precision.

It is difficult to explain differences in the reception of vowel features based on the
dichotomy between E and TFS cues. Vowel perception relies on spectral resolution to resolve
the formant pattern that distinguishes between vowels, which seems to imply that all vowel
features are mainly defined by TFS cues. Rather, differences in perceptual precision between
vowel features are better explained by the consequences of insufficient frequency resolution
(see Introduction: Harnsberger et al., 2001). These consequences are more important for the
neural coding of low-frequency formants, and should therefore more severely affect nasality,
which depends on the transmission of low-frequency energy in the F1 region, than frontness
which depends on relatively high-frequency energy in the F2 and F3 regions (e.g. Calliope,
1989). We found that not only frontness but also aperture were better perceived than nasality,
which corresponds in part to our expectations. However, Tyler, Preece, Lansing and Gantz
(1992) and Fitzgerald et al. (2007) suggested that information about the F1 frequency, which
is the main acoustic correlate of aperture, was transmitted better than F2 and F3 frequencies,
which are the main acoustic correlates of frontness. The contradictions between these studies
suggest that performance in feature perception varies as a function of various factors, such as
differences in the electrophysiological profiles of the CI users and perhaps also the kind of CI
device. It is also consistent with the fact that there is no simple one to one correspondence
between acoustic and phonetic properties.

4.3. Conclusion

The present study suggests that CI children present the same degree of categorical
perception as normal-hearing children. However, categorical precision is lower for CI
children than for NH children matched for listening age. This finding is important because it
cannot be easily explained by auditory deprivation before implantation. If auditory
deprivation played an important role in the deficits of the CI children in this study, it should have stronger effects on categorical perception than on categorical precision because categorical precision develops more slowly than categorical perception in NH children (Hoonhorst et al. 2009, 2011). It would thus seem that the sensitive period for development of categorical perception is not restricted to the first 1-2 years of life. The sensitive period may be significantly longer. Among all possible factors, the precision deficit could be attributed both to limitations in cochlear implants’ transmission of acoustic cues and frequency resolution.

---
1 While it is evident that the perception of a single phonemic feature depends on multiple acoustic cues (for a review see Repp, 1982), it is also true that some of these cues have major perceptual importance. The other cues are secondary: they only act when the main cue is ambiguous (typically voice onset time for voicing: Abramson & Lisker, 1985). This has been shown even with naturally produced spontaneous speech (Saerens, Serniclaes & Beeckmans, 1989). Here we describe each feature which acts as a “major cue”, i.e. those with the largest perceptual weight.

2 The French vocalic repertoire includes three to four nasal vowels depending on the dialect (Malmberg, 1971).

3 The listening age was defined in terms of a child’s “aging” or maturity in listening once he has access to sound via a cochlear implant. A child who is 6 years old and has had an implant for 3.6 years would be expected to demonstrate auditory skills similar to a 3.6-year-old, not a six-year-old.

4 Clarion: The implant has 16 intracochlear electrodes. The implant uses CIS or SAS strategy. The CIS strategy covers the frequency band 350-6800 Hz. Each channel has an update rate of 833 pulses/s (pps). The frequency range of SAS strategy is 250-6800 Hz and a maximum of eight channels are available. The update rate is approximately 13000 pps.

5 Nucleus 24: The Nucleus 24 electrodes array contains 22 intracochlear electrodes and two extracochlear ground electrodes that permit up to 20 channels of information. The implant uses SPEAK, CIS or ACE strategy. The update rate is approximately 14400 pps. Just as the CIS, the ACE strategy has a higher rate of stimulation as compared to SPEAK, ranging between 500 and 2400Hz.

6 Digisonic: The implant has 20 electrodes on an array of 25 mm and all 20 electrodes can be activated. The default frequency distribution of the implant is 195-8003Hz. The stimulation rate may be set between 260 and 1000 pps per electrode in the SP. The implant uses main peak interleaved sampling (MPIS) strategy.

7 Equation 1 : \( d' = z_{hit} - z_{false} \) (false alarm)
References


Figure 1. Discrimination and identification scores for cochlear implant and normal-hearing children.

37x37mm (400 x 400 DPI)
Figure 2. Discrimination and identification scores for cochlear implant and normal-hearing children for each articulatory feature. The first row presents consonant features and the second row presents vowel features.

55x37mm (400 x 400 DPI)
Figure 3. Differences between normal-hearing and cochlear implant children in mean d' scores for identification and discrimination tasks, for articulatory features of vowels and consonants.

37x37mm (400 x 400 DPI)
<table>
<thead>
<tr>
<th>Chronological age (years; months)</th>
<th>Age at implantation (years; months)</th>
<th>Length of cochlear implant use (years; months)</th>
<th>Type of device</th>
<th>Communication mode</th>
<th>Education Placement</th>
</tr>
</thead>
<tbody>
<tr>
<td>9;1 (1;1)</td>
<td>2;7 (0;9)</td>
<td>6;6 (1;1)</td>
<td>12 Nucleus Freedom</td>
<td>9 early and intensive Cued Speech + oral</td>
<td>6 special education</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4 Nucleus SPrint</td>
<td>11 oral</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 Digisonic</td>
<td>5 late Cued Speech + oral</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 Clarion</td>
<td>19 mainstream</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Characteristics of children with cochlear implant (CI).

\(^a\) mean and standard deviation (in parentheses).

<table>
<thead>
<tr>
<th>Chronological Age</th>
<th>Listening Age</th>
<th>PM 47</th>
</tr>
</thead>
<tbody>
<tr>
<td>Years;months</td>
<td>Years;months</td>
<td>Number of correct responses / 36</td>
</tr>
<tr>
<td>CI</td>
<td>9;1 (1;1)***</td>
<td>6;6 (1;1)*** &gt;.20</td>
</tr>
<tr>
<td>NH</td>
<td>6;4 (1;2)</td>
<td>6;4 (1;2) &gt;.20</td>
</tr>
</tbody>
</table>

Table 2. Chronological age, listening age and non-verbal IQ level of cochlear implant (CI) and normal hearing (NH) participants.

\(^a\) mean and standard deviation (in parentheses).