Check for updates

Consonant and vowel production in children with cochlear implants: acoustic measures and multiple factor analysis

Sophie Fagniart[1,2]*, Brigitte Charlier[3,4], Véronique Delvaux[1,2,5], Anne Huberlant[4], Bernard Georges Harmegnies[1,2,3], Myriam Piccaluga[1,2] and Kathy Huet[1,2]

[1]Language Sciences and Metrology Unit, University of Mons (UMONS), Mons, Belgium, [2]Research Institute for Language Science and Technology, University of Mons (UMONS), Mons, Belgium, [3]Université Libre de Bruxelles, Free University of Brussels, Brussels, Belgium, [4]Functional Rehabilitation Center "Comprendre et Parler", Brussels, Belgium, [5]National Fund for Scientific Research, Brussels, Belgium

Introduction: The acoustic limitations of cochlear implants (CIs) can lead to perceptual limitations and consequently to imprecise phonological representations and production difficulties. The aim of the study is to document the phonological and phonetic skills of children with CIs and their typically hearing peers. Phonetically, three types of segments were targeted, each characterized by contrasting acoustic information: nasal/oral vowels, fricative segments, and voiced/voiceless stops.

Methods: Forty-seven typically hearing children (TH) and 23 children with CIs performed a picture-naming task. Productions were analyzed to obtain phonological measures (percentages of correct phonemes, types of errors), and various acoustic measures were collected to characterize the productions on the three types of segments investigated. Multiple factor analyses were conducted to study productive profiles on the various acoustic measures, and the dimensions were correlated with phonological measures.

Results: The results showed lower performance in lexical (target word retrieval) and phonological (percentages of correct phonemes) skills among children with CIs (CI group), although with better performances among children exposed to CS. Acoustically, children in the CI group exhibited productions significantly different from those of the TH group in terms of the distinction of fricative consonants, marking nasalization through nasal resonance cues, and in the production of voiceless stops. However, the CI group demonstrated compensatory strategies (lengthening of VOT for voiced stops, marking of nasalization through oropharyngeal configuration cues).

Conclusions: The results indicate that children with CIs are at risk of experiencing difficulties in both phonetic and phonological domains. However, there are opportunities for compensation through the exploitation of acoustic cues better coded by the CI and/or through perceptual means (utilization of visual cues).

1 Introduction

Cochlear implantation is now commonly provided to people with severe to profound deafness, and has been shown to effectively restore hearing function and promote oral language development in children (Sharma et al., 2020; Tamati et al., 2022). However, numerous studies on speech sound production by children with cochlear implants have shown specificities compared to age-matched peers with typical hearing, as well as significant variability in performance. Difficulties in production can be explained primarily by delayed access to oral language associated with a lack of oral language stimulation during sensitive periods in the development of the auditory areas associated with language. Another explanatory factor is related to perceptual difficulties that may arise from processing speech through a cochlear implant, as productive skills require precise support from acoustically and phonologically specified representations (Stackhouse and Wells, 1993). The cochlear implant degrades the spectral structure of sound before transmitting it to the auditory nerve. This degradation is related to the limited number of electrodes capable of independently coding the frequency information of the original sound without activation diffusion or interactions between adjacent electrodes (channel-to-channel interactions). Furthermore, frequency ranges perceived via the implant may be limited in both high and low frequencies. The coding of low frequencies depends on the shallowness of the array insertion and potential mismatches in frequency mapping (Başkent and Shannon, 2005; Başkent et al., 2016). Frequencies above ∼8,000 Hz reach the limits of the processor in current implants (Loizou, 2006; Reidy et al., 2017), meaning that speech sounds with acoustic cues relying on high frequencies are more likely to be perceived and encoded imprecisely by individuals with cochlear implants. The present study aims to investigate how French-speaking children with cochlear implants produce three types of speech segments: nasal and oral vowels, where the distinction is primarily carried by low-frequency information; fricative consonants, where acoustic cues are mainly carried by high-frequency information; and voiced/unvoiced plosive consonants, where the voicing contrast is supported by temporal acoustic cues, presumed to be better encoded by the cochlear implant than spectral cues.

In French, the production of contrastive nasal vowels involves nasal resonance and a specific vowel quality associated with a characteristic oropharyngeal configuration (lip, tongue, and larynx positioning). The acoustic coupling of nasopharyngeal and oropharyngeal cavities results in various acoustic changes compared to oral vowels, including shifts in frequency, intensity, and bandwidth of the first formant (Delattre, 1954; House and Stevens, 1956; Delattre and Monnot, 1968; Maeda, 1993), as well as changes in intensity ratios between the first harmonics and among different formants (Chen, 1995, 1997; Delvaux, 2002; Delvaux et al., 2002). These acoustic differences between vowels contrasting for nasalization necessitate the precise processing of acoustic information with a sufficient degree of frequency selectivity and sensitivity to amplitude variations, particularly among low-frequency harmonics, which may pose challenges for cochlear implant recipients. The study of nasal and oral vowels in CI users has been the subject of a limited number of studies, possibly due to the non-contrastive nature of vowel nasalization

in many languages worldwide. However, Bouton et al. (2012) highlighted difficulties in discriminating minimal pairs based on nasal and oral vowels among French-speaking children with cochlear implants, attributing the challenges to insufficient spectral resolution and difficulty in coding low-frequency information. Borel (2015) and Borel et al. (2019) noticed challenges in identifying nasal vowels among adult French speakers with cochlear implants, particularly when these vowels were phonetically similar in oropharyngeal configuration to other oral vowels in the French system. This observation led to the development of a discrimination task involving phonologically contrasting nasal and oral vowels (according to the nasal-oral distinction in the French phonological system: /ã/-/a/, /ɔ̃/-/ɔ/, /ɛ̃/-/ɛ/) as well as phonetically divergent pairs in which the oral and nasal vowels were close in terms of oropharyngeal configuration (/ã/-/ɔ/, /ɔ̃/-/o/, /ɛ̃/-/a/). A recent study (Fagniart et al., 2024) confirmed these findings in children CI recipients, who have greater difficulty discriminating phonetically matched nasal-oral pairs. Intensive exposure to Cued Speech led to a better utilization of temporal acoustic cues, resulting in improved performance in these children. Subsequent analyses of nasal and oral vowel productions from the same children revealed reduced differentiation based on acoustic cues related to nasal resonance, but increased differentiation based on formant frequencies (i.e., oropharyngeal configuration) and segmental length.[1] These results support the hypothesis of increased difficulty in detecting nasal anti-resonances and other acoustic cues related with phonetic nasality, although this can be compensated for by exploiting more accessible cues conveying the oral-nasal contrast such as formant values or temporal differences.

The production of fricative consonants involves a constriction in the vocal tract generating turbulent airflow. The resulting aperiodic signal (noise source) covers a wide frequency range with significant energy in the high frequencies. It is then filtered by the vocal tract, resulting in a concentration of energy in the mid to high frequencies depending on the location of the constriction (place of articulation). Due to limitations in processing high frequencies by the implant processor, these segments are prone to causing perceptual and productive difficulties in CI recipients. Identifying and discriminating the places of articulation is more challenging for children with CIs (Lane et al., 2001; Mildner and Liker, 2008; Bouton et al., 2012), especially for the phonemes /s/ and /ʃ/ (Giezen et al., 2010; Hedrick et al., 2011). On the production side, late and imprecise emergence of fricative consonants has been observed in the phonemic repertoires of children with implants, although performance improves with age and duration of CI use (Warner-Czyz and Davis, 2008). Concerning phonological accuracy, some authors (Kim and Chin, 2008) identified typical error patterns in CI children, which are associated with fortition errors (e.g., cessation of fricatives, devoicing). These errors match those observed in the early stages of phonological development in typically hearing children (Jakobson, 1968), suggesting delayed

---

1   Fagniart, S., Delvaux, V., Harmegnies, B., Huberlant, A., Huet, K., Piccaluga, M., et al. (under review). Producing nasal vowels without nasalization? Perceptual judgments and acoustic measurements of nasal/oral vowels produced by children with cochlear implants and typically hearing peers. *J. Speech Lang. Hear. Res.*

acquisition patterns that are not unique to CI children. In the same vein, Faes and Gillis (2016) have shown that phonological accuracy in fricative consonants is delayed when comparing CI and typically hearing children based on age, but not when matched in terms of vocabulary size. Several acoustic studies have also documented difficulties related to the production of fricatives segments in children with CI compared to their age-matched typically hearing peers, such as: diminished differentiation in the /s/-/ʃ/ contrast (Mildner and Liker, 2008; Todd et al., 2011; Reidy et al., 2017), specific patterns in implementing the /f/-/s/ contrasts in French (Grandon and Vilain, 2020), and overall lower spectral values (Yang and Xu, 2023).

The production of stop consonants involves the active and complete closure of the vocal tract by movements of the articulators toward each other, followed by a quick opening that releases a burst of acoustic energy. In voiced stops, vocal cord vibration accompanies the closing phase, contributing to the addition of a periodic sound source voiced. Voice Onset Time (VOT) serves as the acoustic marker for the voicing contrast in stop consonants. VOT represents the duration of the period of time between the release of the oral closure and the onset of vocal cord vibration (Lisker and Abramson, 1964). Since the voicing contrast in stop consonants is carried by temporal cues, one could presume that it is appropriately encoded by CI. This was suggested by Bouton et al. (2012), who noted better performance in children with CIs in discriminating minimal pairs opposing stop consonants on the basis of the voicing feature, compared to other distinctive features. However, this finding has not been consistently verified. For instance, Peng et al. (2019) reported lower performance in discriminating minimal pairs involving voiced vs. voiceless stops among young cochlear implant recipients compared to their hearing peers. Studies using categorical perception paradigms have also yielded contradictory results regarding the performance of children with CIs, with some studies showing lower categorical perception (Giezen et al., 2010), while others did not find any difference when compared to typically hearing children (Medina et al., 2004) for the voicing contrast. Few studies have examined VOT measurements to objectively assess how voiced and voiceless stops are distinguished in the speech productions of children with CI. Uchanski and Geers (2003) and Horga and Liker (2006) observed shorter VOT values for voiceless stops, leading to a reduced voiced-voiceless distinction compared to typical-hearing peers. Grandon et al. (2017) observed shorter VOT values for voiceless stops in French-speaking CI children, but only for the velar consonant /k/. Despite the voicing feature of stop consonants being indicated by temporal cues, studies on the perception and production of this distinctive feature show contrasting results, warranting including them in our study of the speech productions of French-speaking children with CI.

As most studies have focused on a single distinguishing feature in isolation, the main purpose of the present study is to document the productive skills of different types of distinction with the same children, to jointly observe their productive profiles based on phonological and phonetic analysis. To this purpose, we focused on three types of segments: nasal/oral vowels, fricative consonants, and stop consonants, to examine whether there are common production profiles across different types of targeted phonetic features. Productions will be collected through a picture-naming task, to study the phonological representations stored in the

children's memory. Taking the literature into account, it can be expected that, children with a cochlear implant (CI):

A) May have difficulty finding the precise phonological form of target words considering their perceptual limitations. These difficulties may manifest in lower naming performance (less retrieval of the target word in the first instance) and/or in more phonemic substitution when producing the target word;
B) May distinct nasal and oral vowels relying more on better-encoded cues, like formant frequencies related to oropharyngeal configuration rather than nasal resonance cues (see text footnote 1);
C) May produce fricative consonants with less distinction of place of articulation (Mildner and Liker, 2008; Todd et al., 2011; Reidy et al., 2017; Grandon and Vilain, 2020);
D) May produce voiceless stops with shorter values (Uchanski and Geers, 2003; Horga and Liker, 2006; Grandon et al., 2017).

The originality of the study lies in jointly examining these different segments, aiming to identify distinct profiles of common difficulties and/or compensatory strategies that may be observed among the children. In addition to studying these different hypotheses through comparisons between CI children and typical hearing peers, different variables likely to have an impact on performance will also be studied, namely chronological age as well as hearing age, age of implantation and exposure to Cued Speech among CI children.

2 Materials and method

2.1 Participants

Two groups of children were recruited: a group of children with typical hearing (TH group) and a group of children with cochlear implants (CI group). The TH group comprises 47 French-speaking children with typical hearing, with an average age of 56 ± 13 months, who do not exhibit any learning delays or auditory disorders. The CI group consisted of 23 French-speaking children (mean age: 67 ± 15 m.) with congenital bilateral profound hearing loss, 22 of whom had bilateral implants, and one child with a unilateral implant. All CI participants received "oralist" auditory rehabilitation, both at their rehabilitation center and in their family environment. This group was divided based on their exposure to Cued Speech: eight of the children were not exposed to CS (CS0), while 15 were exposed to CS during their speech therapy sessions (two at three sessions per week) and/or in their family context (CS1). Implantation age groups were also created, with children who received their first implant before 16 months considered as early implantations (CI/EI, $n = 12$), and those implanted after 16 months considered as late implantations (CI/LI, $n = 11$). The age of 16 months was chosen to be in line with various studies showing a significant benefit from implantation before 18 months (Sharma et al., 2020). Given the distribution of implantation ages, we lowered the threshold to 16 months, enabling us to create equivalent groups. The list of participants and their characteristics are presented in Table 1.

Both groups were divided into three/four chronological age groups: 2;6-3;6 years (only for TH group), 3;7-4;6 years, 4;7-5;6 years, and 5;6-7 years (see Table 2). For children in the CI group,

TABLE 1 Characteristics of the CI children.

Subject	Sex	Chronological age (years; months)	Chronological age group	Age at first implantation (months)	Implantation age group: early (E)—late (L)	Auditory age group	Implantation type	CS exposure group
CI1	M	4;6	3;7-4;6 y.	9	E	2;6-3;6 y.	Bilateral	CS0
CI2	M	6;5	5;7-7 y.	39	L	2;6-3;6 y.	Bilateral	CS0
CI3	F	5;10	4;7-5;6 y.	15	E	3;7-4;6 y.	Bilateral	CS0
CI4	F	6;7	5;7-7 y.	7	E	3;7-4;6 y.	Bilateral	CS0
CI5	F	6;6	4;7-5;6 y.	31	L	3;7-4;6 y.	Bilateral	CS1
CI6	F	4;7	4;7-5;6 y.	7	E	3;7-4;6 y.	Unilateral	CS1
CI7	F	7;3	5;7-7 y.	13	E	4;7-5;6 y.	Bilateral	CS1
CI8	M	4;7	4;7-5;6 y.	13	E	2;6-3;6 y.	Bilateral	CS1
CI9	M	4;9	4;7-5;6 y.	13	E	2;6-3;6 y.	Bilateral	CS1
CI10	M	4;6	3;7-4;6 y.	12	E	2;6-3;6 y.	Bilateral	CS1
CI11	F	4;6	3;7-4;6 y.	18	L	2;6-3;6 y.	Bilateral	CS1
CI12	M	5;6	4;7-5;6 y.	9	E	2;6-3;6 y.	Bilateral	CS0
CI13	F	7;3	5;7-7 y.	24	L	4;7-5;6 y.	Bilateral	CS1
CI14	G	7;10	5;7-7 y.	26	L	4;7-5;6 y.	Bilateral	CS1
C15	G	6;11	5;7-7 y.	23	L	4;7-5;6 y.	Bilateral	CS1
C16	F	6;9	5;7-7 y.	20	L	4;7-5;6 y.	Bilateral	CS1
C17	M	6;0	5;7-7 y.	20	L	3;7-4;6 y.	Bilateral	CS0
CI18	F	3;9	3;7-4;6 y.	23	L	3;7-4;6 y.	Bilateral	CS0
CI19	F	5;0	4;7-5;6 y.	12	E	2;6-3;6 y.	Bilateral	CS1
CI20	F	3;8	3;7-4;6 y.	32	L	2;6-3;6 y.	Bilateral	CS1
CI21	M	4;11	4;7-5;6 y.	11	E	2;6-3;6 y.	Bilateral	CS0
CI22	F	6;7	5;7-7 y.	17	L	2;6-3;6 y.	Bilateral	CS1
CI23	F	5;0	4;7-5;6 y.	13	E	3;7-4;6 y.	Bilateral	CS1

TABLE 2 Groups and age subgroups distribution.

Group	Chronological age subgroups years; months (N)	Auditory age subgroups years; months (N)
CI	3;7-4;6 y. (5)	2;6-3;6 y. (11)
	4;7-5;6 y. (9)	3;7-4;6 y. (7)
	5;7-7 y. (9)	4;7-5;6 y. (5)
TH	2;6-3;6 y. (9)	N/A (typical hearing)
	3;7-4;6 y. (10)	
	4;7-5;6 y. (17)	
	5;7-7y. (11)	

auditory age groups were also formed, considering their age from the time of their first implantation.

2.2 Data collection and treatment

2.2.1 Procedure

Children's speech samples were collected using a picture naming activity (Philippart De Foy et al., 2018). Target words were carefully chosen by the authors to include all French phonemes in initial, medial, or final syllabic position. In addition, these words were selected for their high lexical frequency and early age of acquisition, to facilitate their retrieval by young children. In terms of target segments, the target words contained 25 fricative consonants, 13 nasal vowels and 69 oral vowels, as well as 42 stop consonants.

The target word pictures were presented to the child one at a time via a booklet, and he or she was asked to orally name each picture. Different prompts were provided if the child did not respond or if the produced word did not match the target (semantic paraphasia or random response). First, semantic cues related to the target word were provided (e.g., for example: "you can use it when it rains" for "umbrella"). If the target word was still not produced, a phonological cue was offered by presenting its initial phoneme (e.g., "it starts with/s/" for "/suri/"—mouse). If these two cues were not sufficient for the child to retrieve the target word, the experimenter would produce the target word and ask the child to repeat it. Thus, each target word could be elicited through four types of elicitation: spontaneous naming, naming after semantic prompt, naming after semantic and phonological prompts, or simple repetition. Production based on naming and on repetition can imply different mechanisms: while naming requires retrieval of a phonological form stored in memory, repetition relies on auditory skills while allowing direct imitation of the stimulus. Based on this principle, the effect of the type of elicitation (direct naming or prompt vs. repetition) will also be studied within productions. The children's productions were recorded using a H5 Zoom portable recorder.

2.2.2 Phonological analysis

All the audio files were annotated by an initial examiner and subsequently verified by the first author using the Phon 3.1 software (Hedlung and Rose, 2020). By comparing them with the canonical

phonological content of the target words, these annotations made it possible to the extraction of the Percentage of Correct Phonemes (PCP), Correct Fricatives (PCF), Correct Nasal vowels (PCN), Correct Stops (PCS), and to identify the various types of production errors made by the children such as substitution based on place or manner of articulation or voicing.

2.2.3 Acoustic analysis

The annotations performed in Phon were subsequently exported to Praat (Boersma and Weenink, 2023). Phoneme alignments were manually corrected to enable the use of semi-automated scripts for extracting acoustic measures on the segments of interest.

2.2.3.1 Nasal vowels

The acoustic description of vowels aimed to study the two main aspects of nasal/oral vowel production: the adoption of an articulatory configuration specific to the vowel quality, on one hand, and the resonance with the nasal cavities (only for nasal vowels) on the other hand. To investigate the acoustic characteristics associated with oropharyngeal configuration, formant values were examined. For the study of nasal resonance, Nasalization from Acoustic Features (NAF) values (Carignan et al., 2023) were generated. A total of 6,605 vowels were analyzed.

Formant measurements were obtained using a semi-automated procedure. For F1, F2, and F3, the formant value used is the median value of the series of values obtained every 5 ms in the interval between 25 and 75% of the total vowel duration. Given the sensitivity of formant value detection to spectrogram parameters, particularly in children, several precautions and verifications were implemented to avoid errors in formant detection. Initially, formant detection parameters were adjusted individually for each vowel and child. After extracting the formant values based on these parameters, a visualization of the productions in the F1/F2 space was utilized to identify any aberrant values. Aberrant values were identified by checking if F1, F2, or F3 values fell beyond plus or minus three standard deviations from the mean formant values of the subject. All outliers were manually verified, with spectrograms examined to correct formant values or exclude vowels with unreadable or unclear signals (with a negligible number of occurrences, around 1%).

To assess the degree of nasality in the vowel productions, a procedure largely inspired by the NAF (Nasalization from Acoustic Features) method (Carignan, 2021; Carignan et al., 2023) was employed. First, a large array of measures was collected through semi-automated procedures to extract acoustic indices at 11 time points within the vowels. These measures included overall amplitude, formant bandwidths for F1, F2, and F3, as well as relative amplitude deltas between formants and poles: A1-P0, A1-P1, A3-P0 (measured using the "Nasality Automeasure Praat" script by Styler, 2017) and various indices proposed by Carignan (spectral moments and nasal murmur). Note that some acoustic indices used in Carignan's initial method, such as formant frequency values and Mel-frequency spectral coefficients (MFCC), were not included here since effects pertaining to oropharyngeal configuration alterations were measured separately with formant values. Secondly, a model was built to reduce the various acoustic cues linked to vowel nasality to a value that would characterize the

oral-nasal dimension. Indeed, it is currently complicated to isolate a single acoustic metric to reflect the degree of nasal resonance (Carignan, 2021). Based on this principle, we drew inspiration from the NAF method to build a machine learning model that predicts a metric value quantifying the oral/nasal character of children's productions based on the series of acoustic cues collected. A supervised machine learning technique was employed: the gradient-boosted decision tree model. This technique necessitates training the model on a portion of the data, requiring a training and test sample. For this purpose, part of the time points over which acoustic measurements were collected within each vowel were used for training, the other for testing. To avoid capturing the effects of pre- and post-vocalic phonetic context, we excluded the time points corresponding to the 0, 10, 90, and 100% portions of the vowel, leaving 7 time points. Next, we partitioned the dataset by extracting the time points at 20, 40, 60, and 80% of the duration of each vowel from the children in the TH group to form the training sample. We chose to include these time points because they represent a relatively stable portion of the vowel that is most likely to carry information related to vowel nasality. The training sample was made up of children from the TH group only, so that the model could be trained on supposedly typical productions. Within the training sample, productions were labeled as oral (0) or nasal (1) based on the target vowel to be produced. Subsequently, a gradient-boosting decision tree model (XGBoost R Package, Chen and Guestrin, 2016) was trained on the scaled selected acoustic features with multiple iterations to optimize hyperparameters and minimize cross-validation errors. Finally, the trained model was used to predict nasality responses on the testing sample. The model was defined with minimized linear regression error, to permit the obtention of values on a scale from 0 to 1 on an oral-nasal mapping dimension. The resulting NAF values ranged numerically from 0 to 1, with higher values indicating a higher predicted degree of nasality, and intermediate values corresponding to those that are halfway to the acoustic characteristics of nasal and oral vowels.

To examine strategies employed in the phonetic implementation of the phonological contrast between nasal and oral vowels, paired comparison analyses were conducted, considering the phonetic (/ã/-/ɔ/, /ɔ̃/-/o/, /ɛ̃/-/a/) and phonological (/ã/-/a/, /ɔ̃/-/ɔ/, /ɛ̃/-/ɛ/) proximity (Borel, 2015) of oral-nasal pairs in French. We also included the pairs /ã/-/o/, as the distinction between /o/ and /ɔ/ is sometimes subtle in children's productions, and /ɔ̃/-/u/, as these segments are also very close phonetically (Fagniart et al., 2024). For each child, each produced nasal vowel was paired with all orally produced vowels that were phonetically or phonologically similar, resulting in a listing of all oral/nasal pairs produced. A total of 30,402 pairs were formed, allowing for comparisons of acoustic cues within each nasal-oral pair. Euclidean distances in the F1-F2-F3 (Bark) planes (as described in Calabrino, 2006) and differences between NAF values were examined for each pair.

2.2.3.2 Fricative consonants

The acoustic characterization of fricative consonants was conducted using recently developed measures (Shadle et al., 2023), allowing for the examination of both the place of articulation, i.e., the location of airflow obstruction, and the quality of the frication noise generated by analyzing intensity ratios across low, mid, and high-frequency bands. These measurements were conducted within spectra generated by the Multitaper Method (MTPS; Blacklock, 2004), which averages a series of periodograms obtained through the collection of mutually orthogonal windows (tapers). The MTPS method is renowned for its minimized errors and enhanced temporal precision (Sfakianaki et al., 2024).

A total of 1917 fricatives were analyzed. A R script adapted from the script developed and provided by Reidy et al. (2017)[2] generated a MTPS using eight tapers at the temporal midpoint of the phoneme. Three acoustic measures were then collected from the generated spectra: spectral peak, levelD, and ampDiff for each target sibilant /s,z,ʃ,ʒ/ or ampRange for each target non-sibilant /f-v/. The spectral peak was obtained by extracting the frequency of the amplitude peak in the mid frequencies, levelD was obtained by calculating the difference in acoustic power between mid and high frequencies, and ampDiff represented the amplitude difference between low and mid frequencies. It is worth noting that the indices levelD and ampDiff quantify the energy ratios in low, mid, and high frequencies. A good frication noise source should have a significant portion of acoustic energy in mid and, particularly, high frequencies. Therefore, a good noise source should result in high ampDiff values (as mid frequencies are reinforced compared to lows) and low levelD values (indicating a large proportion of energy in high frequencies). These three measures required the definition of ranges for low, mid, and high frequencies within the spectrum. Since there were no references available for young children, these ranges were established through a meticulous analysis of the spectra, employing trial-and-error to identify parameters that most accurately represented our data. Finally, the values proposed by Shadle for adult females (Shadle et al., 2023) with slight modifications were adopted. Notably, the maximum threshold for the mid-frequency range in the detection of spectral peaks for /s, z/ was adjusted to 10,000 Hz instead of 8,000 Hz, and to 8,000 Hz instead of 4,000 Hz for /ʃ, ʒ/.

2.2.3.3 Stop consonants

A total of 3,012 stops were analyzed. To calculate VOT, stop consonants were manually annotated on Praat by identifying the consonant burst, which represents the moment of stop release, and the onset of voicing, which could precede the burst in the case of voiced consonants or follow it in the case of voiceless consonants. Subsequently, a Praat script was used to extract the VOT of all the annotated stops.

2.3 Statistical analysis

Linear generalized mixed models, employing the lme4 package (version 1.1-34; Bates et al., 2015) within the R software (R Core Team, 2020), were used to compare groups among the various acoustic measures on the children's speech productions. These models were constructed by including subject and stimulus characteristics (the variables and their levels are specified in Table 3) and the interaction among these variables. It is worth noting that it was the expected segments relative to the target

---

2 Freely at: https://github.com/MontrealCorpusTools/SPADE/blob/main/jss_sibilant_revised_jm_0618.R.

TABLE 3 Variables related to subject and stimulus characteristics and their levels.

LGM variables	Variables and their levels
Subject characteristics	Auditory status: cochlear implant children (CI) – typical hearing children (TH)
	CS exposure: CI/CS0 = no CS exposure - CI/CS1 = CS exposure – TH
	Chronological age group: 2;6-3;6/3;7-4;6/4;7-5;6/5;7-7
	Auditory age group: 2;6-3;6/3;7-4;6/4;7-5;6/5;7-7
	Implantation age group: CI/EI = early (<16 m.) – CI/LI = late (>16 m.) – TH
Stimulus characteristics	Segment identity: - Nasal/oral pair:/ã/-/a/, /ã/-/ɔ/, /ã/-/o/, /ɔ̃/-/ɔ/, /ɔ̃/-/o/, /ɔ̃/-/u/, /ɛ̃/-/ɛ/, /ɛ̃/-/a/ - Fricative consonants: /f/-/s/-/ʃ/-/v/-/z/-/ʒ/ - Stop consonants: /p/-/t/-/k/-/b/-/d/-/g/
	Voicing type: voiced – voiceless (for fricative and stop consonants)
	Elicitation type: naming – repetition

word that allowed for labeling the identity of the productions. For example, the /ã/ in "pantalon" (/pãtalɔ̃/- "pants") was labeled as/ã/regardless of the actual production of the segment, i.e., even if it was denasalized. To address inter-subject variability, a random intercept effect for the subject was integrated into the model. Significance assessment of fixed effects were examined using Chi-squared tests and corresponding *p*-values, conducted via the ANOVA function of the Car package (Fox and Weisberg, 2018) applied to the model. Additionally, *post-hoc* analysis were conducted using the emmeans package (Lenth et al., 2024).

Multiple factor analyses were conducted using the FactoMineR package (Le et al., 2008), and graphical representations were created using Factoextra (Kassambara and Mundt, 2020). They were performed on a dataset consisting of subject-wise averages of various acoustic measures aggregated as means, namely:

- Euclidean distance values of F1-F2-F3 and NAF for all nasal-oral pairs, where higher average values indicate a greater distinction between nasal and oral configurations in terms of oropharyngeal configuration and nasal resonance.
- Spectral peak values by place of articulation, averaged levelD and ampDiff for fricatives, where one would expect to observe better articulation places marked by higher and well-distinguished spectral peak values, lower values for levelD, and higher values for ampDiff, representing a reinforcement of high-frequency acoustic energy associated with a good frication source (Shadle et al., 2023);
- Mean differences between VOT values of voiceless and voiced stops, where higher values indicate a greater distinction in voicing between voiced and voiceless stops.

The subjects' characteristics (hearing status, age groups, and CS exposure) were added as supplementary variables not actively involved in constructing the dimensions. This addition allows for observing the distribution of different subgroups based on the constructed dimensions. The description of the

generated dimensions along with their constituent variables and the additional variables was performed using the dimdesc function (package FactoMineR). Finally, to determine whether a relationship exists between children's phonological performance and their acoustic profiles, we conducted Pearson correlations between the dimensions of the multiple factorial analysis and the various phonological accuracy scores obtained.

3 Results

3.1 Naming task performance

As explained in Section 2.2.1, children produced all target words of the naming task but may have done so using different types of elicitation: spontaneous naming or after semantic prompt, after semantic and phonological prompts, or through simple repetition. The percentages of the first type of elicitation, spontaneous naming, are significantly higher in the TH group (84.4%) than in the CI group [77.4%; $\chi^2_{(1)} = 4.96$; $p = 0.02$]. No group effect is observed for the second type of elicitation, i.e., naming on semantic cue [TH: 2.79%; CI: 1.96%; $\chi^2_{(1)} = 1.26$; $p = 0.26$], while the third, based on phonological priming, is found significantly more frequently in the TH group [6.22%; CI: 2.03%; $\chi^2_{(1)} = 10.05$; $p = 0.001$]. Production based on repetition of the target word, the fourth type of elicitation, is significantly more common among children in the CI group [18.38%; TH: 5.95%; $\chi^2_{(1)} = 17.06$; $p < 0.001$]. An effect of CS exposure is observed on the percentage of spontaneous naming (elicitation 1): only children in the CI/CS0 group differ significantly from the TH group [70.8%; $t_{(67)} = -13.65$; $p = 0.02$], with the CI/CS1 group showing similar performance [80.3%; $t_{(67)} = -4.15$; $p = 0.47$]. No effect of chronological or auditory age or age of implantation group was observed.

3.2 Phonological analysis

The percentages of correct phonemes are analyzed to document phonological accuracy. Children in the CI group have significantly lower percentages of correct total phonemes [CI: 77.5%—TH: 91.1%; $\chi^2_{(1)} = 31.87$; $p < 0.001$], correct nasal vowels [CI: 74%—TH: 91.5%; $\chi^2_{(1)} = 35.43$; $p < 0.001$], correct fricative consonants [CI: 74.3%–TH: 90.4%; $\chi^2_{(1)} = 36.67$; $p < 0.001$] and correct stop consonants [CI: 76.9%–TH: 90.7%; $\chi^2_{(1)} = 29.07$; $p < 0.001$]. Table 4 presents the percentages of different error types on the target segments. The most frequently observed error type is denasalization of nasal vowels with significantly higher rate than TH children [$\chi^2_{(1)} = 27.07$; $p < 0.001$]. Fricativization [$\chi^2_{(1)} = 10.19$; $p = 0.001$] and stopping [$\chi^2_{(1)} = 10.8$; $p = 0.001$] errors are also retrieved at a significantly higher rate in the CI group as well as voicing of voiceless stops [$\chi^2_{(1)} = 25.96$; $p < 0.001$], these errors being negligible in the TH group (<1%). Devoicing errors are retrieved in the two groups, with a marginally higher rate in the CI group [$\chi^2_{(1)} = 3.04$; $p = 0.08$] while nasalization of oral vowels is negligible in the two groups.

CS exposure displays a significant effect on the correct percentages of nasal vowels [$\chi^2_{(2)} = 43.14$; $p < 0.001$], with the

TABLE 4 Percentage correct phonemes among the cochlear implant (CI) and typically hearing (TH) groups.

Measure	Group performances (%)				Significance of group comparison tests			
	CI	CI/CS0	CI/CS1	TH	CI/TH	CI/CS0-CI/CS1	CI/CS0-TH	CI/CS1-TH
% correct phonemes (PCP)	77.5	72.5	79.6	91.1	***		***	**
% correct nasal vowel (PCN)	74.0	65.7	77.6	91.5	***	*	***	***
% correct fricatives (PCF)	74.3	69.4	76.5	90.4	***		***	***
% correct stops (PCS)	76.9	71.7	79.2	90.7	***		***	***
% vowel nasalization	0.53	1.28	0.2	0.16		**	***	
% vowel denasalization	15.25	23.23	11.76	2.65	***	*	***	**
% voicing errors	2.94	3.33	2.77	0.41	***		**	***
% devoicing errors	5.14	3.33	5.93	3.72				*
% stopping errors	3.37	3.33	3.38	0.56	**			*
% fricativization errors	1.29	2.26	0.87	0.31	**	*	***	

ns, *p < 0.05, **p < 0.01, ***p < 0.001.

CI/CS1 group showing significantly higher score than the CI/CS0 group [77.6 vs. 65.7; $t_{(67)} = -11.8$; $p = 0.05$] but lower than the TH group [$t_{(67)} = -13.9$; $p < 0.001$]. As for the different error types, the CS/CS0 group also shows higher percentages of nasalization of oral vowels and of fricativization of stops than the two other groups. For nasal vowels denasalization, the CI/CS1 group show lower error percentages than the CI/CS0 group but the percentage remain higher than the TH group. Devoicing of voiced stops was observed at a higher percentage in the CI/CS1 group compared to the two others. No effect of chronological or auditory age was observed, nor were there any effects of the age group at implantation.

3.3 Acoustic analysis

Table 5 presents the means, as well as the significance of the associated group comparison tests, for the various acoustic measurements carried out on the studied segments, grouped according to auditory status (TH vs. CI) and exposure to CS (CS0 vs. CS1).

3.3.1 Nasal/oral vowels

This section will focus on the analysis of acoustic differences within pairs of nasal-oral vowels. Formant and NAF values averaged per target phoneme and per child group, as well as the p-values associated with group difference tests, are available in the appendices. Considering nasal-oral pairwise comparisons in terms of Euclidean distances in the F1/F2/F3 plane, an auditory group*pair interaction is observed [$\chi^2_{(7)} = 201.6$; $p < 0.001$]. The CI group exhibits higher values for 5 out of 8 pairs, indicating a greater differentiation in terms of oropharyngeal configuration for these pairs, namely/ã/-/o/, /ɔ̃/-/o/, /ɔ̃/-/ɔ/, /ɔ̃/-/u/, and /ɛ̃/-/a/pairs

(see Figure 1). An interaction between elicitation type (naming vs. repetition), pair, and group is observed [$\chi^2_{(21)} = 330.6$; $p < 0.001$]. Indeed, the CI group show higher Euclidean distances between oral and nasal vowels in the repetition condition for all pairs except /ã/-/a/ and /ɛ̃/-/a/, while the TH group shows higher values in the repetition condition for /ã/-/a/, /ɔ̃/-/ɔ/, /ɛ̃/-/a/, and /ɛ̃/-/ɛ/. A significant CS exposure group*pair interaction is also found [$\chi^2_{(14)} = 309.55$; $p < 0.001$], with the TH group showing lower values than the CI/CS0 and CI/CS1 groups for /ɔ̃/-/o/ and /ɔ̃/-/u/, while the CI/CS0 group shows the highest values compared to other groups for /ã/-/o/ and the lowest for /ɛ̃/-/ɛ/. An interaction between elicitation type (naming vs. repetition), pair, and CS exposure group is also observed [$\chi^2_{(35)} = 466.1$; $p < 0.001$]. While the CI/CS1 group showed higher values in the repetition condition for all the pairs except/ɛ/-/a/, the CI/CS0 group is characterized by higher values only for /ɔ̃/-/u/ and /ɛ̃/-/a/, with, conversely, lower values in the repetition condition for /ɛ̃/-/ɛ/ and /ã/-/a/. An interaction between chronological/auditory age group, auditory status group, and pair is observed. Indeed, an auditory age group effect is only observed in the /ɔ̃/-/u/pair, with decreasing values for age groups following 3;7-4;6. In the TH group, a chronological age group effect was observed in /ã/-/a/, /ã/-/ɔ/, and /ã/-/o/, with decreasing values in older age groups. When comparing the groups based on age of implantation, there's an observed interaction effect between the age of implantation groups and pair [$\chi^2_{(14)} = 299.1$; $p < 0.001$]. Specifically, the group of children with later implantation (CI/LI) shows significantly higher values than the group of children with early implantation (CI/EI) for the pair /ɔ̃/-/ɔ/.

The statistical analysis of nasal/oral differences in terms of NAF values revealed an interaction between auditory status group and pair [$\chi^2_{(7)} = 201.5$; $p < 0.001$], with significantly higher values in the TH group for the /ã/-/o/, /ɔ̃/-/o/, /ɔ̃/-/ɔ/, /ɔ̃/-/u/ and /ɛ̃/-/a/ pairs. An interaction between elicitation type, group, and pair

TABLE 5 Acoustic analysis according to auditory status and exposure to Cued Speech (CS).

Segment type	Measure	Target vowel	Group means				Significance of group comparison tests			
			CI	CI/CS0	CI/CS1	TH	CI—TH	CI/CS0—CI/CS1	CI/CS0—TH	CI/CS1—TH
Vowel	E.D. F1-F2-F3 (brk)	/ã/-/a/	4.41	4.39	4.42	4.6				
		/ã/-/ɔ/	3.62	3.34	3.72	3.49				
		/ã/-/o/	3.45	3.97	3.28	3.06	*	*	**	
		/ɔ̃/-/ɔ/	4.01	3.5	4.19	3.68	*	*		*
		/ɔ̃/-/o/	3.42	3.47	3.42	2.74	***		*	**
		/ɔ̃/-/u/	3.32	3.52	3.27	2.84	**		*	*
		/ɛ̃/-/ɛ/	3.97	3.37	4.17	4.01		*	*	
		/ɛ̃/-/a/	3.14	3.11	3.16	2.86	0.06			
	delta NAF	/ã/-/a/	0.114	0.08	0.13	0.17	***		**	*
		/ã/-/ɔ/	0.019	0.002	0.027	0.06	***		*	
		/ã/-/o/	0.071	0.034	0.087	0.127	**		**	*
		/ɔ̃/-/ɔ/	0.052	0.05	0.05	0.11	***		**	*
		/ɔ̃/-/o/	0.109	0.097	0.114	0.178	***			**
		/ɔ̃/-/u/	0.051	0.062	0.045	0.08	*			*
		/ɛ̃/-/ɛ/	0.076	0.05	0.08	0.113	*			
		/ɛ̃/-/a/	0.08	0.06	0.09	0.124	**		*	*
Fricatives	Spectral peak (Hz)	/f/	6,689	6,009	6,990	7,611	***	*	***	*
		/s/	6,085	5,749	6,232	6,702	**		*	
		/ʃ/	4,720	4,207	4,943	5,014		0.08	*	
		/v/	6,421	5,948	6,620	6,618				
		/z/	5,775	5,284	5,981	6,925	***		**	*
		/ʒ/	4,921	4,302	5,198	4,587				
	levelD (dB)	/f/	6	7.81	5.21	3.93	*		**	
		/s/	6.98	8.89	6.17	2.72	***		***	**
		/ʃ/	8.27	10.4	7.34	7.29		0.07	*	
		/v/	5.19	5.13	5.17	5.26				
		/z/	6.99	7.96	6.57	1.89	***		***	***
		/ʒ/	8.46	10.23	7.67	9.23				

(Continued)

TABLE 5 (Continued)

Segment type	Measure	Target vowel	Group means				Significance of group comparison tests			
			CI	CI/CS0	CI/CS1	TH	CI/CS0–CI/CS1	CI–TH	CI/CS0–TH	CI/CS1–TH
	ampDiff (dB)	/f/	13.99	13.69	14.13	13.13				
		/s/	19.09	22.23	17.77	16.91	*	*	**	
		/ʃ/	19.11	20.43	18.55	19.38				
		/v/	12.93	14.6	12.26	11.42				
		/z/	16.16	13.15	17.34	13.21	0.08	**		**
		/ʒ/	16.06	16.54	15.84	16.63				
Stops	VOT (ms)	/p/	23	29.6	20.1	32.1		*		*
		/t/	34	33	34.4	44.6		*		
		/k/	38.2	34.8	39.7	45.3				
		/b/	−57.3	−67.5	−53.4	−48.9		0.07	*	
		/d/	−78.8	−87	−75.7	−57.3		*	0.09	
		/g/	−42.6	−60.2	−36.7	−36.6				

ns, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

$[\chi^2_{(21)} = 155.9; p < 0.001]$ was also observed. Indeed, while TH children benefited from repetition which results in an increase in nasal-oral differences in terms of NAF for the pairs /ã/-/a/, /ã/-/ɔ/, /ɔ̃/-/ɔ/ and /ɔ̃/-/u/, for children in the CI group this is only the case for the pairs /ɛ̃/-/ɛ/ and /ɛ̃/-/a/. On the contrary, children in the CI group showed a decrease in NAF values in the repetition condition for the pairs /ã/-/a/, /ã/-/ɔ/, /ɔ̃/-/o/, /ɔ̃/-/ɔ/ and /ɔ/-/u/. Considering CS exposure, an interaction between CS exposure and pair is observed $[\chi^2_{(14)} = 131.6; p < 0.001]$. Indeed, the CI/CS0 group had lower values compared to the two other groups for the /ɔ̃/-/o/ pair and lower compared to the TH group for the /ã/-/ɔ/ pair. The TH group shows the highest values compared to the other groups for the /ɛ̃/-/a/ pair. An interaction between elicitation type and CS exposure group $[\chi^2_{(2)} = 62; p < 0.001]$, as well as between elicitation type, CS group, and pair $[\chi^2_{(35)} = 280.9; p < 0.001]$, was observed. Indeed, while children in the TH and CI/CS0 groups benefited from the repetition condition by seeing their nasal/oral difference values in terms of NAF increase, children in the CI/CS1 group see their overall values decrease. The increase in values in the repetition condition was found significant in the TH group for the /ã/-/a/ pairs, /ã/-/ɔ/, /ɔ̃/-/ɔ/, /ɔ̃/-/u/ and in the CS0 group for the pairs /ã/-/ɔ/, /ɔ̃/-/u/, /ɛ̃/-/ɛ/ and /ɛ̃/-/a/. In the CS1 group, values were significantly lower in the repetition condition for the pairs /ã/-/ɔ/, /ɔ̃/-/ɔ/, /ɔ̃/-/u/. Again, an interaction between chronological/auditory age group, auditory status group, and pair is observed $[\chi^2_{(52)} = 323.2; p < 0.001]$. Indeed, a chronological age effect was observed for /ã/-/o/, /ɔ̃/-/o/, /ɛ̃/-/ɛ/ and /ɛ̃/-/a/ in the TH group with no increasing chronological/auditory age effect on values in the CI group. In comparing the groups formed on the basis of age of implantation, an interaction effect between age of implantation and the pair is observed $[\chi^2_{(14)} = 108.1; p < 0.001]$. Specifically, the group of children with early implantation (CI/EI) exhibited significantly higher values than the group of children with later implantation (CI/LI) for the pair /ɔ̃/-/ɔ/.

3.3.2 Fricative consonants

Concerning spectral peak values, an auditory status group effect is observed, indicating lower values in the CI group $[\chi^2_{(1)} = 9.4; p = 0.002]$. A significant interaction effect is observed between group and phoneme type $[\chi^2_{(5)} = 23.9; p < 0.001]$, with significant group differences noted for the phonemes /f/, /s/, and /z/, suggesting a more posterior place of articulation for these segments in the CI group (see Figure 2). This spectral peak decreased values have an impact on the distinction of the different places of articulation: the CI group shows no significant differences between places of articulation among the voiceless /f/-/s/, /s/-/ʃ/ and the voiced fricatives /v/-/z/ and /z/-/ʒ/, while this phonemes are significantly distinguished in the TH group (/f-s/: $z = 5.8$; $p < 0.001$ -/v-z/: $z = 11.2$; $p < 0.001$ -/z/-/ʒ/: $z = 10.8$; $p < 0.001$). An elicitation type*auditory status group interaction effect is observed $[\chi^2_{(5)} = 506.6; p = 0.05]$. Indeed, while the repetition condition led to increasing spectral peak values in the TH group, it led to decreased values in the CI group. This effect is significant in the TH group for /f/ [naming: 7,566 Hz—repetition: 9,342 Hz; $t_{(1,890)} = -2.5$; $p = 0.01$] and marginal in the CI group for /v/ [naming: 6,538 Hz–repetition: 4,641 Hz; $t_{(1,890)} = 1.8$; $p = 0.07$]. An interaction
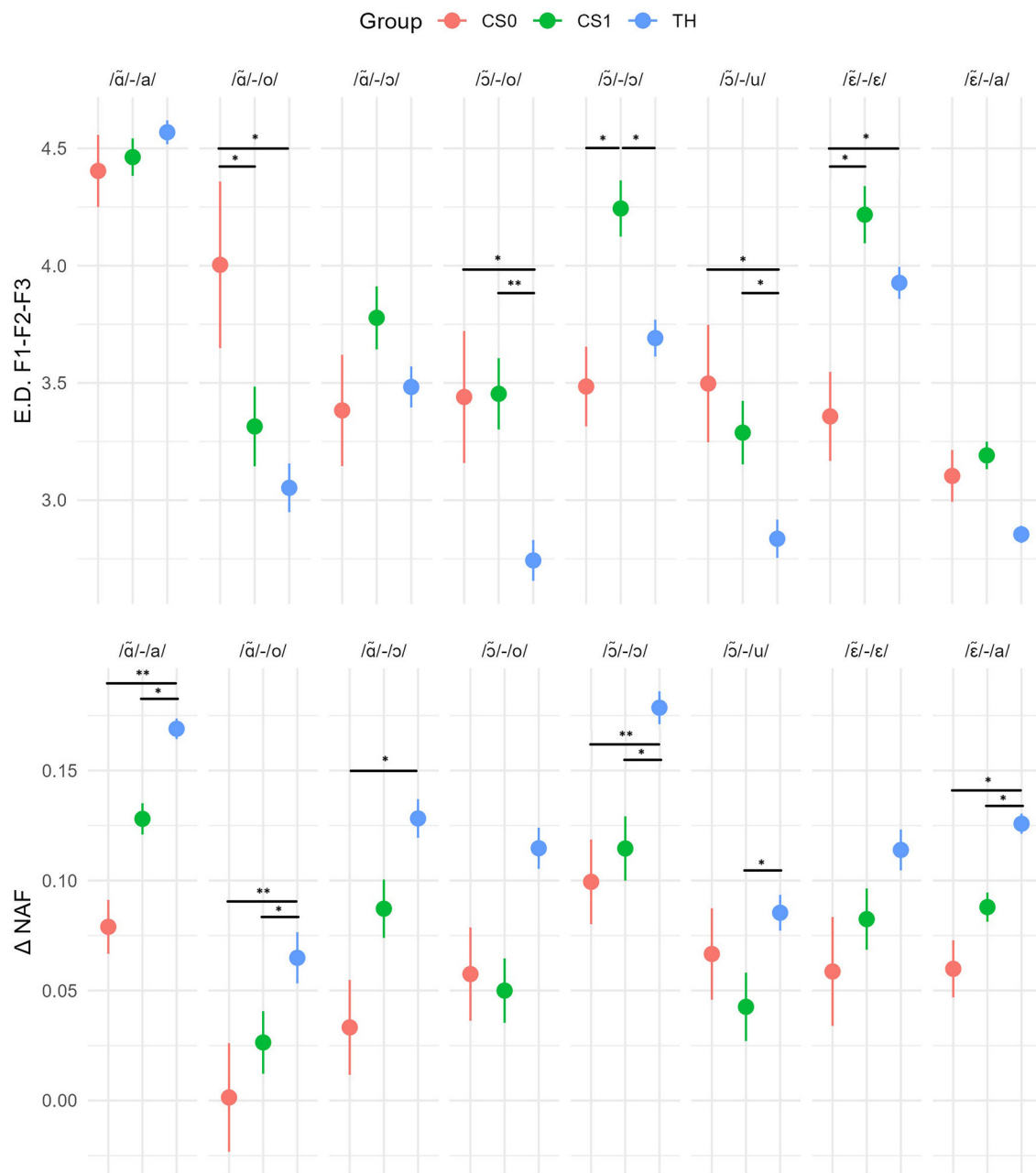
FIGURE 1
Means and 95% confidence intervals of the Euclidean distances and delta NAF values for the different nasal/oral pairs among the CS exposure groups (TH, CI/CS0, and CI/CS1). Significance levels for pairwise comparison tests are shown when the difference is significant at 0.05 (*), 0.001 (**), or < 0.001 (***).

between chronological/auditory age group, auditory status group and phoneme effect is observed [$\chi^2_{(35)} = 66.0$; $p = 0.001$]: while no chronological/auditory group effect appears in the CI group, an effect of chronological age is observed in the TH group, with spectral peak values decreasing with age for /f/ and /s/, resulting in improved distinction of articulation places among voiceless fricatives /f/, /s/, /ʃ/. A CS exposure grouping effect [$\chi^2_{(2)} = 14.7$; $p < 0.001$] as well as an interaction between CS grouping and phoneme [$\chi^2_{(10)} = 28.4$; $p = 0.001$] are obtained: spectral peak values are significantly lower in the CI/CS0 group compared to the TH group ($z = -3.5$; $p = 0.001$) and marginally to the CI/CS1

groups ($z = -2.1$; $p = 0.09$), while the TH and the CI/CS1 group had similar mean values. Regarding phoneme type, CI/CS0 had significantly lower values for /f/ compared to TH ($z = -4.4$; $p < 0.001$) and CI/CS1 group ($z = -2.7$; $p = 0.02$), as well as marginally lower values than TH group for /ʃ/ ($z = -2.1$; $p = 0.08$) and /v/ ($z = -2.2$; $p = 0.08$). For /s/ and /z/, TH group has significantly higher spectral peak values than the other two groups. An effect of age of implantation group [$\chi^2_{(2)} = 10.1$; $p = 0.006$] as well as an interaction between age of implantation group and phoneme type is observed [$\chi^2_{(10)} = 30.5$; $p < 0.001$]. Specifically, values are generally lower in the late implantation group compared to the TH group (z
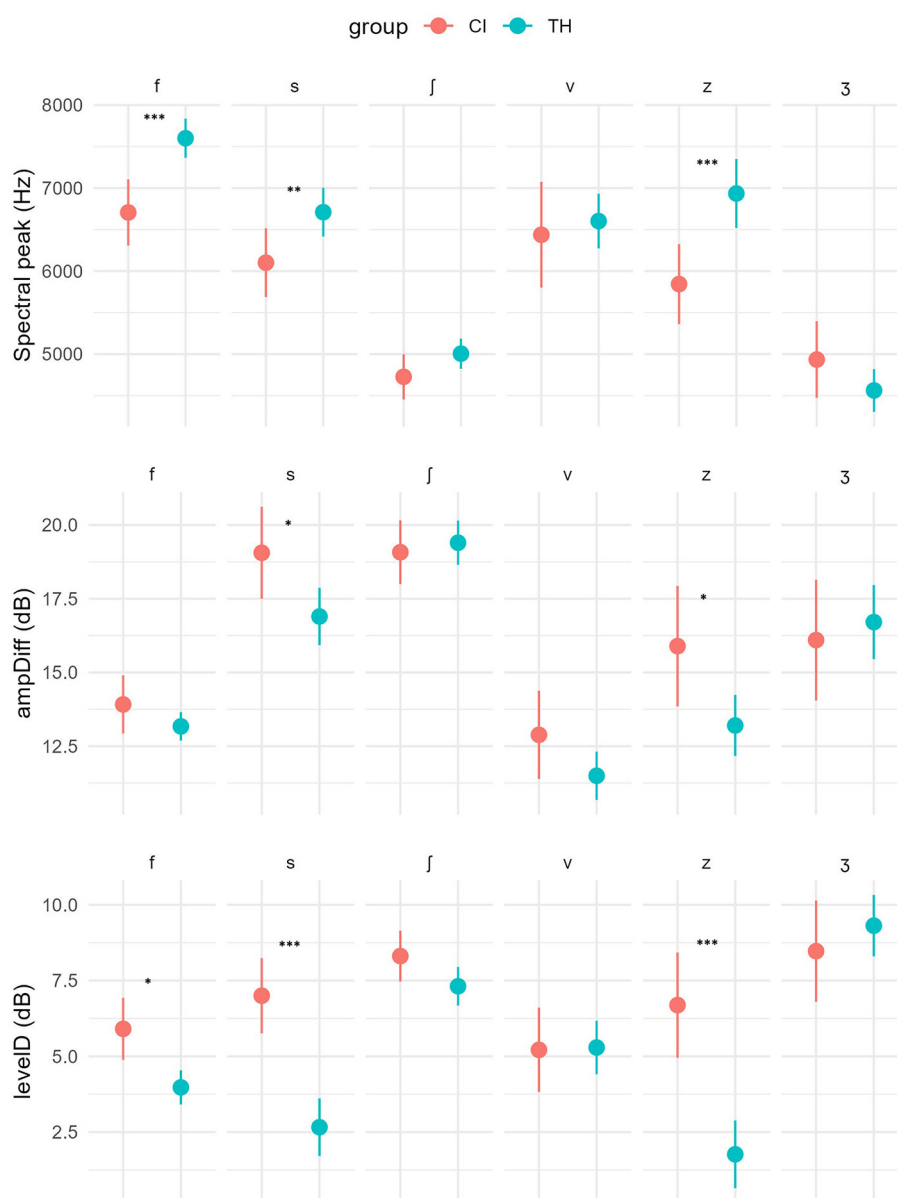
FIGURE 2
Means and 95% confidence intervals of the spectral peak, ampDiff, and levelD values of the different fricative segments of the CI and TH groups. Significance levels for pairwise comparison tests are shown when the difference is significant at 0.05 (*), 0.001 (**), or < 0.001 (***).

$= -2.9$; $p = 0.008$), with this difference being significant for the phoneme /z/ ($z = -3.6$; $p < 0.001$).

AmpDiff values, which reflects amplitude differences between mid- and low-frequency ranges within the fricative spectrum, exhibited an auditory status group effect [$\chi^2_{(1)} = 3.5$; $p = 0.05$], with lower values in the TH group, as well as a group*phoneme interaction effect [$\chi^2_{(5)} = 13.5$; $p = 0.02$] with significantly lower values in the TH group for /s/ ($z = 2.8$; $p = 0.004$) and /z/ ($z = 2.9$; $p = 0.003$). The higher values observed in the CI group may indicate greater reinforcement of mid-frequency areas compared to TH children. No elicitation type effect was observed. A voicing type effect is observed [$\chi^2_{(1)} = 71.6$; $p < 0.001$], with a significant decrease of the voiced fricatives ampDiff values in the TH ($z = 7.4$; $p < 0.001$) and the CI group ($z = 4.3$; $p < 0.001$), as expected.

An interaction between chronological/auditory age group, auditory status group and phoneme are obtained [$\chi^2_{(15)} = 56.7$; $p < 0.001$]: ampDiff values increase with chronological age in the TH group for all phonemes except /f/, while CI group displays a decrease of the values in the older auditory age group for /s/. An interaction between CS exposure and phoneme is observed [$\chi^2_{(10)} = 34.4$; $p < 0.001$], with significantly higher values in the CI/CS1 group compared to the CI/CS0 ($z = -2.3$; $p = 0.06$) and TH groups ($z = 3.5$; $p = 0.001$) whereas the CI/CS0 group had significantly higher AmpDiff values for /s/ compared with the TH group ($z = 2.9$; $p = 0.01$). No effect of implantation group is observed on the ampDiff values.

Regarding the levelD values, which reflects sound level differences between the mid- and high frequency ranges, a

significant group effect was observed, with significantly higher values in the CI group [$\chi^2_{(1)} = 5.6$; $p = 0.01$] as well as a group*phoneme interaction effect [$\chi^2_{(5)} = 98.6$; $p < 0.001$], with values significantly higher for /f/, /s/ and /z/ in the CI group. The higher values of the levelD values in the CI group indicate less reinforcement of high frequencies compared to children in the TH group. An interaction between elicitation type and group was observed [$\chi^2_{(5)} = 47.4$; $p < 0.001$], with repetition condition leading to higher levelID values in the CI group only [naming: 6.7 dB—repetition: 8.9 dB; $t_{(1,891)} = -2.4$; $p = 0.01$]. This trend was significant for /v/ in the CI group [naming: 4.8 dB—repetition: 10.3 dB; $t_{(1,869)} = -1.9$; $p = 0.04$], while in the TH group the repetition condition led to significantly decreased values for /f/ [naming: 3.9 dB—repetition: 0.2 dB; $t_{(1,832)} = 2.1$; $p = 0.03$] and marginally so for /ʃ/ [naming: 7.3 dB—repetition: 3.4 dB; $t_{(1,835)} = 1.9$; $p = 0.06$]. Considering voicing, a marginal group*voicing interaction effect was observed [$\chi^2_{(1)} = 2.7$; $p = 0.09$], with a significant decrease of levelD values for the voiced fricatives only in the TH group ($z = -2.6$; $p = 0.007$). An interaction between chronological/auditory age group, auditory status and phoneme was observed [$\chi^2_{(15)} = 25.4$; $p = 0.04$], with no chronological/auditory age group effect in the CI group, compared to decreased values in older children of the TH group for /f/, /s/, /ʃ/ and /z/. A CS exposure grouping effect [$\chi^2_{(2)} = 7.3$; $p = 0.02$] as well as an interaction between CS grouping and phoneme were observed [$\chi^2_{(10)} = 106.9$; $p < 0.001$]. Indeed, levelD values were significantly lower in general in the CI/CS0 group compared to the TH group, but significantly lower values in the TH group for the phoneme /s/ and /z/, compared to the other two groups. An elicitation type*CS exposure group interaction effect was also retrieved [$\chi^2_{(2)} = 10.1$; $p = 0.006$], with higher values for CI/CS0 group for /f/ compared to CI/CS1 [$\chi^2_{(157)} = 2.3$; $p = 0.06$] and TH group [$\chi^2_{(210)} = 1.9$; $p < 0.001$], with significantly lower values for /z. CI/CS1 had lower values than the CI/CS0 group for /ʃ/ in the repetition condition [$\chi^2_{(231)} = 2.6$; $p = 0.02$]. An effect of age of implantation [$\chi^2_{(2)} = 6.1$; $p = 0.04$] in interaction between group and phoneme type [$\chi^2_{(10)} = 55.7$; $p < 0.001$] was observed. Specifically, the later implanted children showed significantly higher values than the TH children ($z = 2.45$; $p = 0.04$) for the phonemes /f/ and /s/.

3.3.3 Stop consonants

An interaction between auditory status group and voicing type (voiced vs. voiceless) was observed on the VOT of the stop consonants [$\chi^2_{(1)} = 30.58$; $p < 0.001$], with higher VOT for voiceless stops and lower for voiced stops in the TH group when compared to the CI group. Phoneme*group pairwise comparisons shown that this group effect was significant for the voiceless stop /t/ and the voiced /b/ and /d/ (see Figure 3).

An elicitation type effect [$\chi^2_{(1)} = 6.4$; $p = 0.01$] as well as an interaction between elicitation type, auditory status group and voicing type [$\chi^2_{(3)} = 39.6$; $p < 0.001$] is observed. Indeed, VOT values are overall higher in the repetition condition, particularly for the voiceless stops, in the TH group (naming: 39.9 ms—repetition: 46.9 ms; $z = 7-6.95$; $p = 0.02$), and to a greater extent in the CI group (naming: 33.2 ms—repetition: 43.8 ms; $z = -10.6$; $p = 0.001$), allowing them to reach similar mean values than in the TH

group. An interaction between chronological/auditory age group and auditory status group is observed [$\chi^2_{(10)} = 48.4$; $p < 0.001$]. Indeed, in the TH group, an increase of the mean values from younger to older age groups is observed for voiced and voiceless stops, while no (chronological or auditory) age effect is observed in the CI group. A CS exposure grouping*voicing type interaction is observed [$\chi^2_{(2)} = 10.92$; $p = 0.004$], with the CI/CS0 group showing higher VOT values for voiced stops compared to the CI/CS1 [$t_{(68)} = -0.015$; $p = 0.08$] and TH groups [$t_{(68)} = -0.021$; $p = 0.002$], whereas the CI/CS1 group shows the lowest VOT values for the voiceless stops [$t_{(68)} = -0.011$; $p = 0.003$]. Phoneme*CS grouping pairwise analysis reveals that the higher values in the CI/CS0 group is significant for voiced /b/ and /g/ compared to the other groups, while CI/CS1 children shows higher values for /d/ compared to TH children. The CI/CS1 group shows lower values than the TH group for voiceless /t/ and /k/. An interaction between CS exposure grouping, voicing type and elicitation type is also observed [$\chi^2_{(7)} = 40.9$; $p < 0.001$], with a significant increase of the voiceless stops VOT in the repetition condition in the CI/CS1 group (naming: 32 ms—repetition: 42.3 ms; $z = -2.6$; $p = 0.007$), this increase being only marginal in the CI/CS0 group (naming: 36 ms—repetition: 46 ms; $z = -1.7$; $p = 0.08$). An interaction effect is observed between age of implantation and voicing type in plosives [$\chi^2_{(2)} = 34.5$; $p < 0.001$]. Specifically, children in the late implantation group showed significantly longer negative VOT values than children in the TH group ($z = 2.7$; $p = 0.02$). An interaction effect appears between age of implantation and phoneme type [$\chi^2_{(10)} = 47.5$; $p < 0.001$], with the lengthening of negative VOTs in the late implantation group being significant for the phonemes /b/ and /d/.

3.3.4 Multiple factor analysis of acoustic measures

A multiple factor analysis was conducted, integrating subject-averaged values of NAF and Euclidean distances in F1-F2-F3 plane between each nasal vowel and the averaged values of the associated oral vowels, the differences between positive and negative VOT values, as well as spectral peak values by location (/f/-/v/-/s/-/z/-/ʃ/-/ʒ/) and ampDiff and levelD mean values. These variables were grouped according to the type of segment characterized (fricative, stop, nasal/oral vowels) but also the production mechanism associated (place vs. frication noise for fricatives, formant vs. NAF values for vowels). Among the 8 dimensions generated, the first three will be analyzed, capturing 61.84% of the explained variance. The first dimension, contributing to explaining 28.5% of the total variance, is more correlated with groups of variables associated with fricative consonants (place = 0.79; frication = 0.62) and with spectral peak variables (SP/s/-/z/= 0.79; SP/f/-/v/= 0.78), while variables associated with frication quality are negatively correlated (levelD = −0.8, ampDiff = −0.44). In other words, positive values on the first dimension indicate high values of spectral peaks as well as lower values of levelD and ampDiff (indicating more reinforcement of high frequencies in the frication), while negative values indicate lower spectral peaks and higher values of levelD and ampDiff (enhancement of mid-range frequencies in the frication). The correlations between additional categorical variables and dimension 1 show
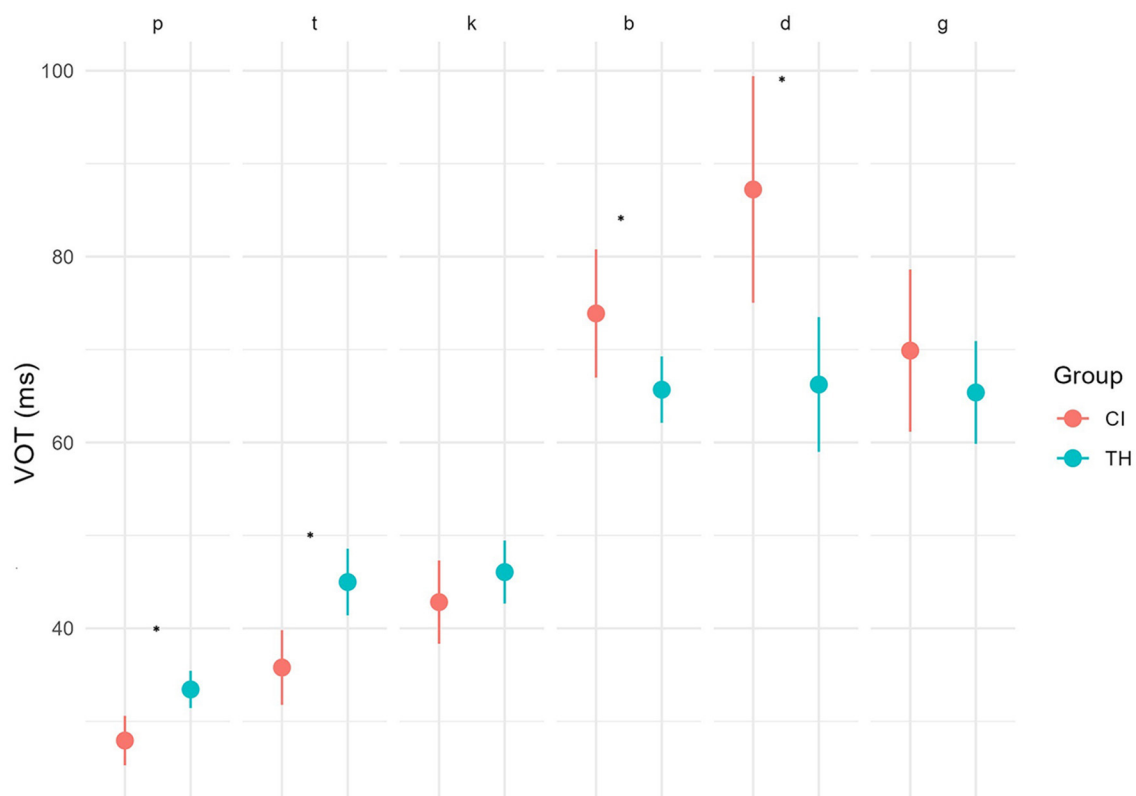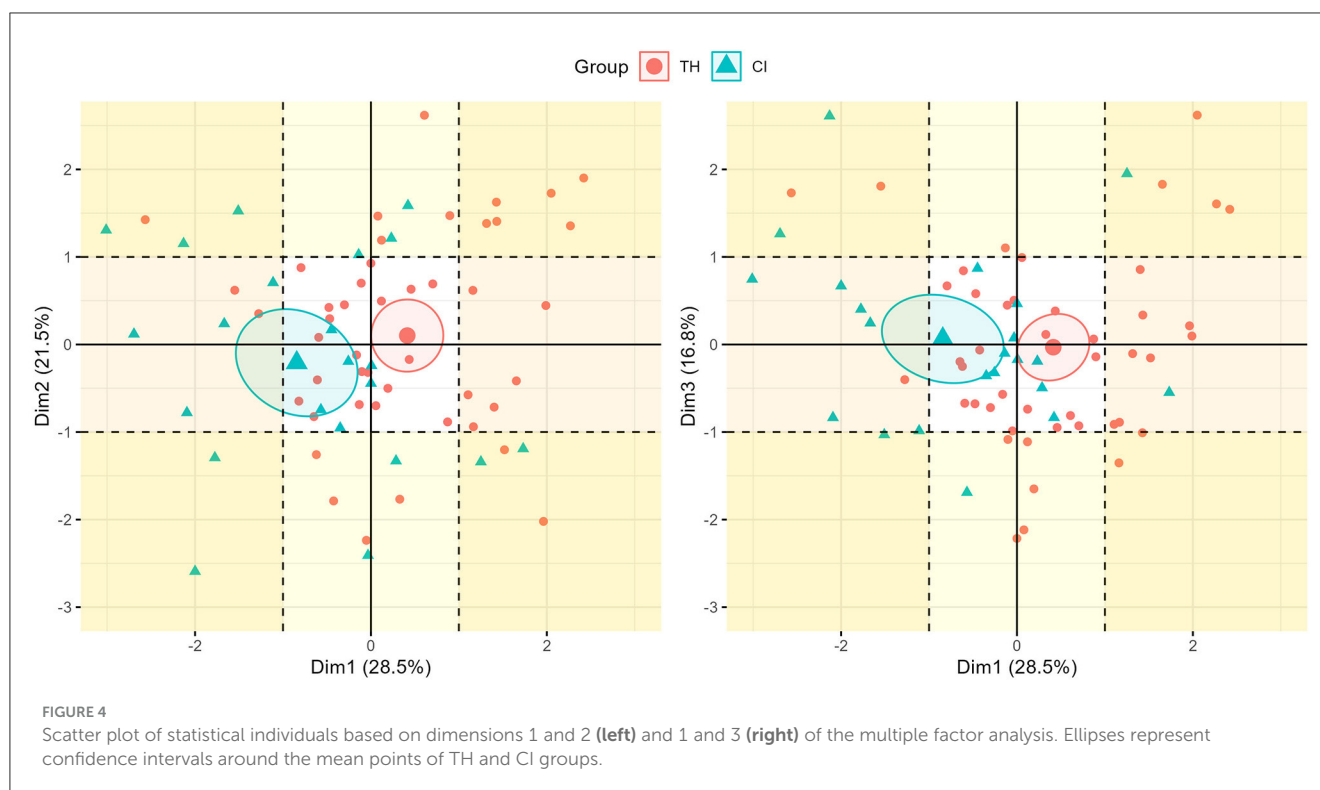
FIGURE 3
Means and 95% confidence intervals of the VOT values of the different voiced/voiceless stops among the CI and TH groups. Significance levels for pairwise comparison tests are shown when the difference is significant at 0.05 (*), 0.001 (**), or < 0.001 (***).

that children in the CI group are negatively associated with the dimension (−0.63), whereas children in the TH group are positively associated with the dimension (0.93). The second dimension, contributing 21.5% of the total variance, is associated with the variable related to nasal/oral differences in terms of NAF values (0.47), with positive correlations associated with the NAF values mean differences (0.69), as well as with the variable associated with VOT values (0.54) and negatively with spectral peak of the posterior fricatives/ʃ/-/ʒ/(−0.63). Positive values are then associated with greater nasal/oral distinction based on NAF values and greater voiced/voiceless VOT values distinction. An association is observed with categorical supplementary variables of chronological/auditory age group, with the older age group positively correlated with the dimension (0.66) and younger negatively correlated (−0.69). The third dimension, contributing 16.8% of the total variance, is associated with the variable related to nasal/oral differences in terms of F1-F2-F3 E.D (0.63) with positive correlations associated with the of the F1-F2-F3 E.D. mean differences (0.79), but negatively with the variable associated with VOT values (−0.54). A link with chronological/auditory age group is observed, with the dimension being negatively correlated with the older chronological/auditory age group (−0.47).

Figure 4 illustrates the distribution of children from the CI and TH groups along dimensions 1 and 2 (left) and 1 and 3 (right), along with ellipses representing 95% confidence intervals

around the group means. The ellipses of the two groups are primarily distinguished on dimension 1, with children from the TH group located on the positive side and CI on the negative side. This is consistent with the analyses on fricatives, showing a clear effect of auditory status on productions, with children in the CI group exhibiting lower spectral values, as well as higher levelD values indicating less utilization of high frequencies in their noise frication. On dimension 2, the group mean tends more toward positive values for the TH group and negative values for the CI group, while on dimension 3, both groups are close to 0. It is important to note the large variability around the ellipses. Note the contrasting situation between the two groups in the dimension 1/dimension2 plan: only children from the TH group are situated in the extreme right-hand quadrant (values >1 in dimensions 1 and 2) and only children from the CI group in the extreme left-hand quadrant (values <-1 in dimensions 1 and 2), testifying to contrasting profiles.

Considering the other group variables, different trends between the CI/CS0 and CI/CS1 groups for dimensions 2 and 3 can be observed in Figure 5 (top graphs). Indeed, on dimension 2, the ellipse of CI/CS1 children tends more toward negative values, while the CS0 group leans toward around zero, with more variability. The CI/CS1 group tends to distinguish less nasal/oral vowels based on the NAF values. For dimension 3, the CI/CS1 group has average values toward positive values, showing less nasal/oral distinction based on F1-F2-F3 E.D. values,

FIGURE 4
Scatter plot of statistical individuals based on dimensions 1 and 2 **(left)** and 1 and 3 **(right)** of the multiple factor analysis. Ellipses represent confidence intervals around the mean points of TH and CI groups.

whereas the CI/CS0 group is situated in negative values with again a large variability. Considering the chronological age groups (middle graphs), we can see a trend for younger children to be positioned more to the right on dimension 1, in negative values on dimension 1, and positive on dimension 2. It seems that younger chronological age group 2;6-3;6 (age group represented only by children from the TH group) produce their fricatives with high spectral peaks with frication noise rich in high frequencies, while they mark the nasal/oral distinction more based on the oropharyngeal configuration (F1-F2-F3 E.D.) and less on nasal resonance. This effect is attenuated when considering auditory age, thus including children from the CI group within the 2;6-3;6 age group. When considering implantation age groups on dimension 1, early implanted children (CI/EI) have their average values intermediate between those of the late implantation group and (CI/LI) the TH group. It can also be seen that the CI/EI group is situated toward negative values on dimension 2, while the group with later implantation seems to be more situated toward positive values for dimension 3.

3.4 Link between phonological performance and acoustic dimensions

The study of correlations between various phonological scores and error types with the three dimensions of multiple factor analysis has revealed moderate and significant correlations between dimension 2, related to the marking of nasal/oral distinctions by NAF values, and various phonological scores among the CI and TH groups (see Table 6). In this regard, high values on

the dimension, indicating a better nasal/oral distinction in terms of NAF as well as a better marking of the distinction between voiced and voiceless stops, are associated with better phonological performance. Among children in the TH group, it is also observed that dimension 2 is negatively correlated with the occurrence of errors in oral vowel nasalization and nasal vowel denasalization. A negative correlation between dimension 1 values and the number of voiced errors is observed in the CI group, while a positive correlation is observed in the TH group with the number of voiceless errors.

4 Discussion

The present study investigates the phonological and phonetic skills of a group of 23 children with cochlear implants (CI) and 47 children with typical hearing (TH) through the analysis of productions obtained with a naming task. Phonological skills are examined by assessing correct phoneme scores, while phonetic skills are studied through acoustic analysis of three types of segments: nasal and oral vowels, fricative consonants, and stop consonants. These segments have been chosen because each is primarily supported by rather contrasting acoustic cues, namely low-frequency cues, high-frequency cues, and temporal information, respectively. The effect of auditory status, as well as the effects of chronological/auditory age, exposure to Cued Speech, and age at implantation, are studied. Factor analyses were conducted on all acoustic variables, and the resulting dimensions were correlated with phonological scores.
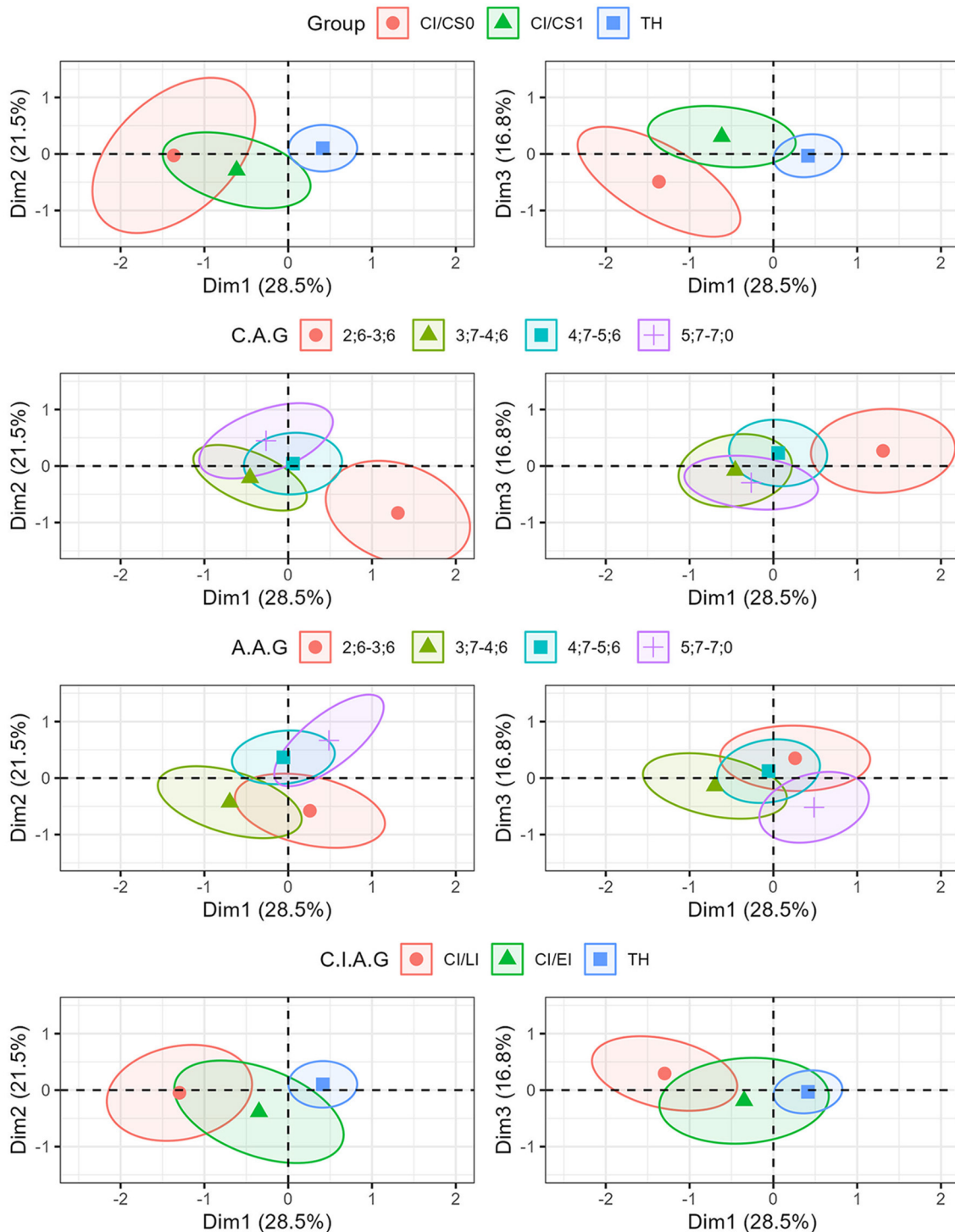
FIGURE 5
Individuals plot representing ellipses around the mean values of the groups based on exposure to CS **(top graphs)**, chronological age **(middle-top graph)**, auditory age **(middle-bottom graphs)**, age of implantation **(bottom graphs)**, according to dimensions 1 and 2 **(left side)**, and dimensions 1 and 3 **(right side)**.

4.1 Phonological form retrieval of the target words

It was hypothesized that, given the perceptual limitations of children with CI, their ability to retrieve the phonological form of their target words could be impacted, with repercussions both lexically (target word retrieval) and phonologically (accuracy of the retrieved phonological form). At the lexical level, children in the TH group demonstrated greater ease in retrieving target words, as evidenced by their significantly higher percentage of spontaneous naming (84%), as well as their higher percentage of retrieval based on phonological cueing. Children in the CI

TABLE 6 Correlations between phonology and acoustic analysis.

Factor analysis dimension	Group	% correct phonemes (PCP)	% correct nasal vowel (PCN)	% correct fricatives (PCF)	% correct stops (PCS)	% vowel nasalization	% vowel denasalization	% voicing errors	% devoicing errors
1	CI	0.19	0.11	0.23	0.19	−0.006	−0.03	−0.44**	0.01
	TH	−0.07	−0.23	−0.17	−0.05	−0.02	0.21	0	0.35*
2	CI	0.45*	0.42*	0.46*	0.47*	−0.07	−0.18	−0.26	−0.33
	TH	0.46**	0.54***	0.54***	0.45**	−0.31**	−0.38**	0	−0.25
3	CI	0.11	0.2	0.06	0.07	−0.22	−0.18	0.15	0.11
	TH	0.05	0.18	0.03	0.08	−0.03	−0.11	0.19	0.047

ns, * p < 0.05, ** p < 0.01, *** p < 0.001.

group showed lower percentages in spontaneous naming (77%) and relied more on repetition (18%). Semantic and phonological cueing provided little assistance in target retrieval, suggesting differences in lexical storage rather than access difficulties compared to their typically hearing peers, who benefited to a greater extent from phonological prompts. A considerable number of studies investigating lexical production in children with cochlear implants have shown comparable performances to typically hearing peers of the same chronological age (Caselli et al., 2012; Luckhurst et al., 2013) or when matched for auditory age (Duchesne et al., 2010) or in early implanted children (Manrique et al., 2004; Connor et al., 2006; Maner-Idrissi et al., 2009; Rinaldi et al., 2013). Other studies show more moderate lexical performances (Young and Killen, 2002; Nittrouer et al., 2018) or with clear difficulties identified (Cambra et al., 2021). Our results seem to align more with these studies, with significantly lower performance than those of children with typical hearing, without a positive effect of chronological, auditory age, or age of implantation. However, a beneficial effect of exposure to Cued Speech is observed, with performances among children exposed to Cued Speech reaching those of the TH group. These findings support literature that has highlighted a positive impact of Cued Speech on children with CIs, both for perceptual skills (Leybaert and LaSasso, 2010; Van Bogaert et al., 2023) and productive abilities (Machart et al., 2021). Studies have also shown a positive impact on early lexical development (Moreno-Torres and Torres, 2008; Rees and Bladel, 2013). Cued Speech, providing complete visual access to all distinctive features of speech sounds, may enable the child to develop more precise phonological representations and thus be more efficient in the storage and retrieval of lexical targets.

On the phonological level, lower performances are also observed in children in the CI group for all types of targeted phonemes: fricatives, nasals, and stops. Certain types of errors were predominantly found in children in the CI group, such as voicing errors, denasalization of nasal vowels, stopping, or fricativization. While stopping errors have been previously reported in children with moderate (Teveny and Yamaguchi, 2023) and profound deafness (Baudonck et al., 2010), and can be classified, along with denasalization errors, as typical errors in development according to Jakobson's markedness theory (Jakobson, 1968), voicing and fricativization errors suggest a more atypical developmental profile. Furthermore, we did not find any effects of chronological/auditory age or age of implantation on phonological scores and error patterns, suggesting more an effect of auditory status than developmental delay. These results support the notion of phonological development constrained by the limitations of the CI described previously, which may lead to underspecified phonological representations and consequently result in production errors. Within this study, this proposition is supported by the observation of a positive effect of exposure to Cued Speech on performances, although scores of CS1 group do not reach the levels of typically hearing children. The group exposed to CS also made fewer errors of oral vowels nasalization, which is consistent with previous studies on vowel nasality perception (Fagniart et al., 2024) and production (see text footnote 1), as well as fewer errors of fricativization, indicating greater stability of phonological representations regarding manner of articulation.

4.2 Nasal-oral vowels distinction

The acoustic analyses characterizing the distinction between nasal and oral vowels reveal an increased marking of the nasal/oral contrast based on indices related to oropharyngeal configuration (formant values) in the CI group compared to TH group. This result is consistent with previous findings obtained in a pseudo-word repetition task and supports the hypothesis that CI children may be more inclined to employ perceptually salient acoustic cues both in perception and production (Fagniart et al., 2024; see text footnote 1). However, the results showed lower values of NAF, representing the degree of nasalization predicted based on a series of acoustic indices related to nasal resonance, suggesting a lesser exploitation of nasal resonance through velopharyngeal opening to distinguish nasal and oral vowels. As described in Section 1, indices related to nasal resonance, primarily carried by low-frequency information associated with fine spectral resolution, are more likely to be poorly transmitted by the CI. This could explain the difficulties observed in phonological production [percentage of correct nasals and (de)nasalization errors], as already noted in the literature on nasal/oral vowel perception (Bouton et al., 2012; Borel, 2015; Borel et al., 2019; Fagniart et al., 2024). These perceptual difficulties may therefore lead to atypically specified phonological representations (marking related more to visually accessible cues such as information related to oropharyngeal configuration), thus resulting in these atypical productions compared to hearing peers. Children exposed to CS exhibit the lowest values in terms of NAF, suggesting a productive profile even more reliant on a phonological system constructed around the most salient cues to access the distinctive features of their oral language. Specifically, in the case of nasal vowels, this relies more on oropharyngeal configurations at the expense of cues related to nasal resonance. The comparison of productions according to the type of elicitation (spontaneous naming or repetition of the target word) supports these findings. Indeed, while children in the TH group improve the marking of nasal-oral distinction in repetition condition for both types of cues as well as NAF values, children in the CI group see their values increase only for the F1-F2-F3 E.D. cue, and on the contrary, their NAF values decrease. In perception, they thus seem to be able to correctly exploit visually accessible information (lip rounding, mouth opening) but not the information related to velopharyngeal opening.

4.3 Fricatives production

Regarding the acoustic study of fricatives, the results confirmed various findings already reported in the literature. Indeed, lower acoustic values had been observed for the center of gravity of fricatives in children with CI (Yang and Xu, 2023), as well as in French-speaking children (Grandon and Vilain, 2020). However, these studies had been limited to the investigation of fricatives /s/-/z/ or all voiceless fricatives in French (/f/, /s/, /ʃ/), and this observation is here extended to voiced segments. The differences between groups were significantly observed for the phonemes /f/, /s/, and /z/, whose spectral peaks are on average higher than those of segments /ʃ/ and /ʒ/, characterized by lower values. This is

entirely consistent with the acoustic limitations in high frequencies mentioned previously. These lowered thresholds also result in a lack of distinction among the three places of articulation among children in the CI group, as the peaks of the segments /f, v/, /s, z/, and /ʃ, ʒ/ are not significantly different. An effect of CS is observed to produce /f/ and marginally for /ʃ/, with values for the CI/CS1 group approaching those of the TH group. However, it is noteworthy that the distinction between the three places of articulation is still not significant in this group. Unlike the productive skills of nasal vowels, the contribution of CS is only moderate for the distinct production of the places of articulation of fricatives. The use of manual cues to provide visual support during the perception of fricative segments may not be enough to develop sufficiently specified representations. It is possible that the acoustic limitations for this distinction are too significant to be compensated for using CS, or that these segments, being among the last to be acquired in the development of children with typical hearing, may be even more challenging for children with CIs. To our knowledge, there is no study documenting fricative productions in terms of frication noise among CI users. The results of the present study show a clear tendency in the CI group to express frication by exploiting energy in mid-range frequencies and less in high frequencies, unlike children with typical hearing. This trend could also directly result from the perceptual limitations of the implant, restricting the perception of frequency ranges above ~8,000 Hz. Indeed, the quality of fricative noise can only be perceived auditorily, with no visual/temporal cues supporting this type of production. This is supported by the study of values in the repetition condition: while TH children see improvements in their productions during repetition (increased spectral peaks, increased energy in high frequency resulting in decreased levelD values), children with CIs, on the contrary, experience slight deterioration in their productions (lower spectral peaks, increased levelD values). The perceptual limitations of CIs do not allow them to access the acoustic information related to the characteristics of fricative segments, thus preventing them from benefiting from repetition for these segments. Possible difficulties in adequately perceiving characteristics related to fricative sound could explain the higher occurrence of errors in articulation mode for stopping or fricativization errors observed in the study, and more broadly in the literature among individuals with moderate (Teveny and Yamaguchi, 2023) or severe deafness (Baudonck et al., 2010). However, an effect of the age of implantation on levelD values was observed, with higher values (and thus less reinforcement in high frequencies) in children with late implantation. It is therefore possible that early implantation allows, to some extent, better exploitation of high-frequency information, despite the technical limitations of the implant, due to the stimulation provided during the sensitive periods of the development of auditory cortical areas.

4.4 Voiced/voiceless stops production

Regarding the production of the voicing feature of stop consonants, a differentiated group effect is observed depending on the type of segments. Indeed, for voiceless consonants, there is a shortening of VOT values in the CI group compared to the TH

group, which is congruent with the literature (Uchanski and Geers, 2003; Horga and Liker, 2006; Grandon et al., 2017). However, in Grandon et al.'s (2017) study on French-speaking children, only the phoneme/k/showed a significant shortening in the CI group, whereas in the present study, it is precisely the phonemes /p/ and /t/ that are significantly shorter in terms of VOT. Grandon had suggested that obtaining a difference only on the phoneme/k/could be attributed to a difficulty in coordinating the articulatory gesture, as/k/has the longest positive VOT in canonical production. However, it is noteworthy that the children in Grandon's study were in a higher age range (6;6-10;6). This difference may explain why, within the TH group, the productions were not sufficiently differentiated between the places of articulation of the voiceless segments, as the average positive VOT of/k/(45 ms) differed little from/t/(44 ms). As a result, the results did not show differences between the CI and TH groups for this phoneme, but rather for the more anterior phonemes /p/ and /t/, whose values were significantly lower in the CI group. Children in the CI group seem to have difficulty coordinating the articulatory gestures associated with the production of voiced stops in a picture naming task. However, when these segments are to be produced in repetition, children in the CI group produce the segments with elongation, allowing them to reach values similar to those of children in the TH group: thus, they are capable of effectively exploiting the acoustic information related to VOT to adjust their productions. On the other hand, for voiced segments, it is the TH group that exhibits a shortening of VOT values, for the phonemes /b/ and /d/. The study of the effects of exposure to CS showed that it mainly involves an elongation of VOT values found among children in the CI/CS1 group. It is possible that relying on temporal cues is a more prevalent strategy in the CI/CS0 group.

4.5 Acoustic profiles

The factorial analyses revealed two distinct trends in the productive profiles of the three investigated segments. Firstly, Dimension 1, which discriminates children well according to their auditory status, consisted of variables related to the quality of fricative production, both in terms of spectral peak and in terms of the utilization of high-frequency energy in frication. It was observed that children in the TH group were predominantly situated on the positive values on the dimension 1, indicating fricatives with high average spectral peak values, and frication noise containing a higher concentration of high frequencies. Dimension 2, on the other hand, was mainly associated with marking the nasal/oral distinction in terms of NAF values, but also, to a lesser extent, with the distinction between voiced and voiceless stop consonants. It is quite interesting to note that among children in both groups, positive correlations are observed between the values on this dimension and various phonological scores. Better marking of the nasal/oral distinction in terms of nasal resonance thus seems to be associated with better phonological performances, both among TH and CI children. Therefore, despite significantly lower NAF values among CI children, there seems to be some variability in the exploitation of nasal resonance cues, which may contribute to part of the variability in linguistic outcomes.

In this regard, it is more surprising to see that Dimension 3, more associated with marking vowel nasality through cues related to oropharyngeal configuration (E.D. F1-F2-F3), is not positively correlated with phonological scores in the CI group. One might have expected that this marking strategy, reflecting a greater reliance on information assumed to be better coded by the CI, would be beneficial phonologically overall. In the study by Fagniart et al. (see text footnote 1) the use of this strategy was associated with better intelligibility of nasal and oral segments. This study seems to indicate that this improvement in segment production is not necessarily associated with better phonological performances overall. These findings support the notion that while the perception-production of fricatives remains critical among the CI population, despite aids such as CS, the perception/production of nasal/oral vowels and stops entails significant variability, which may indicate possible compensations of the perceptual system in children with CIs. These findings are important to consider in the management and evaluation of language skills in children with CIs, to refine auditory stimulation techniques more based on perceptual skills accessible through the CI for critical segments, such as nasal vowels, and to quickly diagnose difficulties that may manifest subclinically.

The various findings of this study must be viewed considering certain limitations. Indeed, it is challenging to assemble a sizable sample with homogeneous characteristics among children with CIs, which complicates the generalization of results. Nevertheless, the results presented here are largely supported by existing literature and can therefore be taken seriously. Regarding the acoustic analyses, it should be noted that the target words were selected to create a list with frequent words, easily imaginable, and with low age of acquisition. These constraints did not allow for controlling various elements, such as phonemic neighborhood or overall syllabic context. Protocols targeting specific segments, with better control over parameters influencing the acoustic characteristics of productions, could be developed to address this bias in future investigations.

5 Conclusion

This study aimed to investigate and correlate phonological and phonetic skills through the analysis of picture naming tasks among children with CIs and their hearing peers. The following observations can be highlighted:

1) Children in the CI group exhibit more difficulties in lexical and phonological domains, which may be compensated for by exposure to Cued Speech.
2) CI users can exploit visually accessible information (such as oropharyngeal configuration) or information better coded by the CI to compensate for their perceptual difficulties, as noted in the production of nasal/oral vowels or voiced/voiceless stops, particularly among children using CS.
3) Distinctive features relying on information not accessible through the implant and less compensable visually and/or temporally, such as the distinction of fricative consonants, are critical among CI children.

4) Adequate exploitation of nasal resonance in distinguishing nasal/oral vowels is associated with better phonological performances.

These findings emphasize the perceptual system's ability to adapt and compensate for the limitations of CIs, a phenomenon that should be prioritized in children's management. Segments most at risk, such as fricative consonants, warrant particular attention to avoid significant phonological underspecification and associated linguistic delays.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical approval was not required for the study involving human samples in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin. Written informed consent was obtained from the minor(s)' legal guardian/next of kin for the publication of any potentially identifiable images or data included in this article.

Author contributions

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Başkent, D., Gaudrain, E., Tamati, T. N., and Wagner, A. (2016). "Perception and psychoacoustics of speech in cochlear implant users," in *Scientific Foundations of Audiology Perspectives from Physics, Biology, Modeling, and Medicine* (San Diego: Plural Publishing), 285–319.

Başkent, D., and Shannon, R. V. (2005). Interactions between cochlear implant electrode insertion depth and frequency-place mapping. *J. Acoust. Soc. Am.* 117, 1405–1416. doi: 10.1121/1.1856273

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Soft.* 67:i01. doi: 10.18637/jss.v067.i01

Baudonck, N., Dhooge, I., D'haeseleer, E., and Van Lierde, K. (2010). A comparison of the consonant production between Dutch children using cochlear implants and children using hearing aids. *Int. J. Pediatr. Otorhinolaryngol.* 74, 416–421. doi: 10.1016/j.ijporl.2010.01.017

Blacklock, O. (2004). *Characteristics of Variation in Production of Normal and Disordered Fricatives, Using Reduced-Variance Spectral Methods.* Southampton: University of Southampton.

Boersma, P., and Weenink, D. (2023). *Praat: Doing Phonetics by Computer.*

Borel, S. (2015). *Perception auditive, visuelle et audiovisuelle des voyelles nasales par les adultes devenus sourds. Lecture labiale, implant cochléaire, implant du tronc cérébral.* Paris: Université Sorbonne Paris Cité.

Borel, S., Serniclaes, W., Sterkers, O., and Vaissière, J. (2019). L'identification des consonnes et voyelles nasales par les adultes implantés cochléaires francophones. *Audiol. Direct* 1:1. doi: 10.1051/audiodir/2019030001

Bouton, S., Serniclaes, W., Bertoncini, J., and Colé, P. (2012). Perception of speech features by french-speaking children with cochlear implants. *J. Speech Lang. Hear. Res.* 55, 139–153. doi: 10.1044/1092-4388(2011/10-0330)

Calabrino, A. (2006). "Effets acoustiques du débit sur la production de la parole chez des locuteurs enfants et adultes," in *Actes du Xe Colloque des étudiants en sciences du langage* (Montréal, QC: Université du Québec à Montréal), 61–83.

Cambra, C., Losilla, J. M., Mena, N., and Pérez, E. (2021). Differences in picture naming between children with cochlear implants and children with typical hearing. *Heliyon* 7:e08507. doi: 10.1016/j.heliyon.2021.e08507

Carignan, C. (2021). A practical method of estimating the time-varying degree of vowel nasalization from acoustic features. *J. Acoust. Soc. Am.* 149, 911–922. doi: 10.1121/10.0002925

Carignan, C., Chen, J., Harvey, M., Stockigt, C., Simpson, J., and Strangways, S. (2023). An investigation of the dynamics of vowel nasalization in Arabana using machine learning of acoustic features. *Lab. Phonol.* 14:9152. doi: 10.16995/labphon.9152

Caselli, M. C., Rinaldi, P., Varuzza, C., Giuliani, A., and Burdo, S. (2012). Cochlear implant in the second year of life: lexical and grammatical outcomes. *J. Speech. Lang. Hear. Res.* 55, 382–394. doi: 10.1044/1092-4388(2011/10-0248)

Chen, M. Y. (1995). Acoustic parameters of nasalized vowels in hearing-impaired and normal-hearing speakers. *J. Acoust. Soc. Am.* 98, 2443–2453. doi: 10.1121/1.414399

Chen, M. Y. (1997). Acoustic correlates of English and French nasalized vowels. *J. Acoust. Soc. Am.* 102, 2360–2370. doi: 10.1121/1.419620

Chen, T., and Guestrin, C. (2016). "XGBoost: a scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (New York, NY: Association for Computing Machinery), 785–794.

Connor, C. M., Craig, H. K., Raudenbush, S. W., Heavner, K., and Zwolan, T. A. (2006). The age at which young deaf children receive cochlear implants and their vocabulary and speech-production growth: is there an added value for early implantation? *Ear Hear* 27, 628–644. doi: 10.1097/01.aud.0000240640.59205.42

Delattre, P. (1954). Les Attributs Acoustiques De La Na-Salité Vocalique Et Consonantique. *Studia Linguist.* 8, 103–109. doi: 10.1111/j.1467-9582.1954.tb00507.x

Delattre, P., and Monnot, M. (1968). The role of duration in the identification of french nasal vowels. *Int. Rev. Appl. Li.* 6:267. doi: 10.1515/iral.1968.6.1-4.267

Delvaux, V. (2002). *Contrôle et connaissance phonétique : Les voyelles nasales du français*. Belgique: Université Libre de Bruxelles.

Delvaux, V., Metens, T., and Soquet, A. (2002). *Propriétés acoustiques et articulatoires des voyelles nasales du français*. XXIVèmes Journées d'étude sur la parole, Nancy, France.

Duchesne, L., Sutton, A., Bergeron, F., and Trudeau, N. (2010). Le développement lexical précoce des enfants porteurs d'un implant cochléaire. *Can. J. Speech-Lang. Pa.* 34.

Faes, J., and Gillis, S. (2016). Word initial fricative production in children with cochlear implants and their normally hearing peers matched on lexicon size. *Clin. Linguist. Phon.* 30, 959–982. doi: 10.1080/02699206.2016.1213882

Fagniart, S., Delvaux, V., Harmegnies, B., Huberlant, A., Huet, K., Piccaluga, M., et al. (2024). Nasal/oral vowel perception in french-speaking children with cochlear implants and children with typical hearing. *J. Speech Lang. Hear. Res.* 67, 1243–1267. doi: 10.1044/2024_JSLHR-23-00274

Fox, J., and Weisberg, S. (2018). *An R Companion to Applied Regression*. Thousand Oaks CA: SAGE Publications, Inc.

Giezen, M. R., Escudero, P., and Baker, A. (2010). Use of Acoustic Cues by Children With Cochlear Implants. *J. Speech Lang. Hear. Res.* 53, 1440–1457. doi: 10.1044/1092-4388(2010/09-0252)

Grandon, B., and Vilain, A. (2020). Development of fricative production in French-speaking school-aged children using cochlear implants and children with normal hearing. *J. Commun. Disord.* 86, 105996. doi: 10.1016/j.jcomdis.2020.105996

Grandon, B., Vilain, A., Lœvenbruck, H., Schmerber, S., and Truy, E. (2017). Realisation of voicing by French-speaking CI children after long-term implant use: An acoustic study. *Clin. Ling. Phonet.* 31, 598–611. doi: 10.1080/02699206.2017.1302511

Hedlung, G., and Rose, Y. (2020). Phon.

Hedrick, M., Bahng, J., von Hapsburg, D., and Younger, M. S. (2011). Weighting of cues for fricative place of articulation perception by children wearing cochlear implants. *Int. J. Audiol.* 50, 540–547. doi: 10.3109/14992027.2010.549515

Horga, D., and Liker, M. (2006). Voice and pronunciation of cochlear implant speakers. *Clin. Ling. Phonet.* 20, 211–217. doi: 10.1080/02699200400027015

House, A. S., and Stevens, K. N. (1956). Analog studies of the nasalization of vowels. *J. Speech Hear. Disord.* 21, 218–232. doi: 10.1044/jshd.2102.218

Jakobson, R. (1968). *Child Language: Aphasia and Phonological Universals*. Berlin: Walter de Gruyter.

Kassambara, A., and Mundt, F. (2020). *factoextra: Extract and Visualize the Results of Multivariate Data Analyses*.

Kim, J., and Chin, S. B. (2008). Fortition and lenition patterns in the acquisition of obstruents by children with cochlear implants. *Clin. Ling. Phonet.* 22, 233–251. doi: 10.1080/02699200701869925

Lane, H., Matthies, M., Perkell, J., Vick, J., and Zandipour, M. (2001). The effects of changes in hearing status in cochlear implant users on the acoustic vowel space and CV coarticulation. *J. Speech Lang. Hear. Res.* 44, 552–563. doi: 10.1044/1092-4388(2001/043)

Le, S., Josse, J., and Husson, F. (2008). FactoMineR: an R package for multivariate analysis. *J. Stat. Softw.* 25, 1–18. doi: 10.18637/jss.v025.i01

Lenth, R. V., Bolker, B., Buerkner, P., Giné-Vázquez, I., Herve, M., Jung, M., et al. (2024). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. R package version 1.10.3099. Available at: https://rvlenth.github.io/emmeans/; https://rvlenth.github.io/emmeans/

Leybaert, J., and LaSasso, C. J. (2010). Cued speech for enhancing speech perception and first language development of children with cochlear implants. *Trends Amplif* 14, 96–112. doi: 10.1177/1084713810375567

Lisker, L., and Abramson, A. S. (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word* 20, 384–422. doi: 10.1080/00437956.1964.11659830

Loizou, P. C. (2006). Speech processing in vocoder-centric cochlear implants. *Adv. Otorhinolaryngol.* 64, 109–143. doi: 10.1159/000094648

Luckhurst, J. A., Lauback, C. W., and Unterstein VanSkiver, A. P. (2013). Differences in spoken lexical skills: preschool children with cochlear implants and children with typical hearing. *TVR* 113, 29–42. doi: 10.17955/tvr.113.1.729

Machart, L., Loevenbruck, H., Vilain, A., Meloni, G., and Puissant, C. (2021). "Speech production in children with cochlear implants: contribution of French Cued Speech," in *ICPLA 2021—International Clinical Phonetics and Linguistics Association (ICPLA)*. Glasgow.

Maeda, S. (1993). "Acoustics of vowel nasalization and articulatory shifts in French nasal vowels," in *Nasals, Nasalization, and the Velum*, Vol. 5, eds. M. K. Huffman and R. A. Krakow (San Diego, CA: Elsevier), 147–167.

Maner-Idrissi, G., L., Rouxel, G., Pajon, C., Dardier, V., Gavornikova-Baligand, Z., et al. (2009). Cochlear implant and lexical diversity development in deaf children: intra- and interindividual differences. *Curr. Psyc. Lett.* 25:4910. doi: 10.4000/cpl.4910

Manrique, M., Cervera-Paz, F. J., Huarte, A., and Molina, M. (2004). Advantages of cochlear implantation in prelingual deaf children before 2 years of age when compared with later implantation. *Laryngoscope* 114, 1462–1469. doi: 10.1097/00005537-200408000-00027

Medina, V., Loundon, N., Busquet, D., Petroff, N., and Serniclaes, W. (2004). *Perception catégorielle des sons de parole chez des enfants avec implants cochléaires*. 25ème journées d'Etudes sur la Parole JEP-TALN-RECITAL, Fès, Maroc.

Mildner, V., and Liker, M. (2008). Fricatives, affricates, and vowels in Croatian children with cochlear implants. *Clin. Ling. Phonet.* 22, 845–856. doi: 10.1080/02699200802130557

Moreno-Torres, I., and Torres, S. (2008). From 1-word to 2-words with cochlear implant and cued speech: a case study. *Clin. Ling. Phonet.* 22, 491–508. doi: 10.1080/02699200801899145

Nittrouer, S., Muir, M., Tietgens, K., Moberly, A. C., and Lowenstein, J. H. (2018). Development of phonological, lexical, and syntactic abilities in children with cochlear implants across the elementary grades. *J. Speech Lang. Hear. Res.* 61, 2561–2577. doi: 10.1044/2018_JSLHR-H-18-0047

Peng, Z. E., Hess, C., Saffran, J. R., Edwards, J. R., and Litovsky, R. Y. (2019). Assessing fine-grained speech discrimination in young children with bilateral cochlear implants. *Otol. Neurotol.* 40, e191–e197. doi: 10.1097/MAO.0000000000002115

Philippart De Foy, M., Delvaux, V., Huet, K., Monnier, M., Piccaluga, M., and Harmegnies, B. (2018). *Un protocole de recueil de productions orales chez l'enfant préscolaire : une étude préliminaire auprès d'enfants bilingues*. XXXIIe Journées d'Études sur la Parole, Aix-en-Provence, France.

R Core Team (2020). *R: A Language and Environment for Statistical Computing (version 4.0.2) [Computer Program]*. Available at: http://www.R-project.org

Rees, R., and Bladel, J. (2013). Effects of English Cued Speech on speech perception, phonological awareness and literacy: a case study of a 9-year-old deaf boy using a cochlear implant. *Deafness Educ. Int.* 15, 182–200. doi: 10.1179/1557069X13Y.0000000025

Reidy, P. F., Kristensen, K., Winn, M. B., Litovsky, R. Y., and Edwards, J. R. (2017). The acoustics of word-initial fricatives and their effect on word-level intelligibility in children with bilateral cochlear implants. *Ear Hear.* 38, 42–56. doi: 10.1097/AUD.0000000000000349

Rinaldi, P., Baruffaldi, F., Burdo, S., and Caselli, M. C. (2013). Linguistic and pragmatic skills in toddlers with cochlear implant. *Int. J. Lang. Comm. Dis.* 48, 715–725. doi: 10.1111/1460-6984.12046

Sfakianaki, A., Nicolaidis, K., and Kafentzis, G. P. (2024). Temporal, spectral and amplitude characteristics of the Greek fricative /s/ in hearing-impaired and normal-hearing speech. *Clin. Ling. Phonet.* 25, 1–27. doi: 10.1080/02699206.2023.2301308

Shadle, C. H., Chen, W.-R., Koenig, L. L., and Preston, J. L. (2023). Refining and extending measures for fricative spectra, with special attention to the high-frequency rangea). *J. Acoust. Soc. Am.* 154, 1932–1944. doi: 10.1121/10.0021075

Sharma, S. D., Cushing, S. L., Papsin, B. C., and Gordon, K. A. (2020). Hearing and speech benefits of cochlear implantation in children: a review of the literature. *Int. J. Pediatr. Otorhinolaryngol.* 133:109984. doi: 10.1016/j.ijporl.2020.109984

Stackhouse, J., and Wells, B. (1993). Psycholinguistic assessment of developmental speech disorders. *Eur. J. Disord. Commun.* 28, 331–348. doi: 10.3109/13682829309041469

Styler, W. (2017). On the acoustical features of vowel nasality in English and French. *J. Acoust. Soc. Am.* 142, 2469–2482. doi: 10.1121/1.5008854

Tamati, T. N., Pisoni, D. B., and Moberly, A. C. (2022). Speech and language outcomes in adults and children with cochlear implants. *Annu. Rev. Linguist.* 8, 299–319. doi: 10.1146/annurev-linguistics-031220-011554

Teveny, S., and Yamaguchi, N. (2023). Phoneme acquisition in French-speaking children with moderate hearing loss. *Clin. Ling. Phonet.* 37, 632–654. doi: 10.1080/02699206.2022.2074309

Todd, A. E., Edwards, J. R., and Litovsky, R. Y. (2011). Production of contrast between sibilant fricatives by children with cochlear implants. *J. Acoust. Soc. Am*. 130, 3969–3979. doi: 10.1121/1.3652852

Uchanski, R. M., and Geers, A. E. (2003). Acoustic characteristics of the speech of young cochlear implant users: a comparison with normal-hearing age-mates. *Ear Hear*. 24, 90S−105S. doi: 10.1097/01.AUD.0000051744.24290.C1

Van Bogaert, L., Machart, L., Gerber, S., Lœvenbruck, H., Vilain, A., and Consortium EULALIES. (2023). Speech rehabilitation in children with cochlear implants using a multisensory (French Cued Speech) or a hearing-focused (Auditory Verbal Therapy) approach. *Front. Hum. Neurosci*. 17:1152516. doi: 10.3389/fnhum.2023.1152516

Warner-Czyz, A. D., and Davis, B. L. (2008). The emergence of segmental accuracy in young cochlear implant recipients. *Cochlear Implants Int*. 9, 143–166. doi: 10.1179/cim.2008.9.3.143

Yang, J., and Xu, L. (2023). Acoustic characteristics of sibilant fricativs and affricates in Mandarin-speaking children with cochlear implants. *J. Acoust. Soc. Am*. 153, 3501–3512. doi: 10.1121/10.0019803

Young, G. A., and Killen, D. H. (2002). Receptive and expressive language skills of children with five years of experience using a cochlear implant. *Ann. Otol. Rhinol. Laryngol*. 111, 802–810. doi: 10.1177/000348940211100908