Interspeech 2018
2-6 September 2018, Hyderabad

Prosodic Focus Acquisition in French Early Cochlear Implanted Children

Chadi Farah [1,2,3], Stephane Roman[1] & Mariapaola D'Imperio [2]

[1] Pediatric Otolaryngology department, La Timone Children's Hospital (APHM), Marseille, France
[2] Aix-Marseille Univ, CNRS, Laboratoire Parole Langage, Aix-en-Provence, France
[3] Faculty of medicine, Saint-Joseph University, Beirut, Lebanon

chadifarah86@hotmail.com, stephane.roman@ap-hm.fr, mariapaola.dimperio@lpl-aix.fr

Abstract

Cochlear implanted (CI) children display an array of speech production and perception problems. No study has evaluated the specific use of prosody regarding information structure in the discourse, in French speaking early CI children.

This study aims to evaluate prosody production in these children, to determine whether they show prosodic effect on word duration.

We conducted a cross-sectional study of 10 prelingually hearing impaired French speaking children (4-7 years old), without comorbidities, CI before the age of 18 months between 2009 and 2012. The speech production task consisted in playing a computer-based semi-structured game, where children interacted with their caregiver. Results were interpreted according to both chronological age and hearing age (HA).

In our series, 6- and 7-year old children (HA<6.2 years) showed stronger lengthening of the focused word in the corrective narrow focus condition than in the contrastive narrow focus which in turn was stronger than in broad focus condition. Only 7-year old children adopted a strategy similar to that of adults, lengthening the end-phrase adjective to preserve the typical phrasing pattern of French.

This study shows for the first time that early CI children are able to acquire important intonation structure features comparable to adult patterns.

Index Terms: prosody, intonation, language acquisition, cochlear implants, prosodic focus, children, French.

1. Introduction

The acquisition of adult-like prosodic structure by children is a crucial element in the dynamics of communication and interaction [1]–[3]. This is a process that takes place gradually from birth to the adolescence for typically developing and normally hearing children. Though cochlear implantation (CI) has dramatically changed the life of children with profound to severe hearing loss, CI children still display an array of speech production and perception problems, both at the segmental and at the prosodic level. This is mainly due to the poverty of the acoustic input especially at the level of the fundamental frequency (F0) transmission [4], [5]. The major problem in speech perception remains the limited capabilities of the currently used devices in coding the F0 variations [6], [7]. As for speech production, rate, fluency, loudness and laryngeal quality are the major problems in this population [8]. With greater hearing experience, the performance of CI children in speech perception and production can improve [6], [7].

Despite a certain amount of work on prosody production in CI children [7]–[11], none of the existing studies have evaluated the acquisition of information structure cued by prosody (i.e., focus vs. background information), especially in the French speaking early CI population.

In studies of English speaking CI children, authors found that they have mainly difficulties in regulating rate and loudness. Moreover, they are less interactive, less talkative and more silent than their NH peers. Intonation performance was poorer in CI children than their NH peers in terms of using appropriate contour especially in using phrase-final pitch rise to signal requests. [6]–[11]

The aim of this study is hence to evaluate prosody production in prelingually hearing-impaired (HI) children with early CI, to determine whether they show adult-like prosodic effects of contrastive and corrective focus on word duration. Given that a previous study on normally hearing (NH) children has shown that French children start using prosodic structure at the phrasal level around 5, mainly through the use of relative syllable duration [12], we explored how and if focus structure could be prosodically encoded by this population. We know in fact that children can prosodically encode focus in stress-accent languages, though we know that this process is quite different in French adults [13]. In French, it appears that focus is mainly signaled through relative duration of the stressed syllable within the Accentual Phrase (AP), instead of a mere displacement of the nuclear pitch accent. Hence, we hypothesized that corrective focus items would induce longer word duration than contrastive focus ones, and that both would be longer than non-focused items, independent of their position in the phrase. Since we also recorded the caregivers' speech patterns, we also compared children's productions to adults' productions.

2. Methods

2.1 Participants

Ten pre-lingually HI French-speaking children aged 4 to 7 years old, participated in our study. They all received CI before the age of 18 months, between 2009 and 2012 at *La Timone* Children's Hospital in Marseille, France. No anomaly or malformations were associated to their HI. CI children's caregivers signed an informed consent before surgery, allowing the medical team to use any data related to CI for research while keeping it anonymous, and consented to participate in this study.

Mean age at testing was 6.06 years, with 3 children in the 4-year-old group, 2 in the 5-year-old group, 2 in the 6-year-old group and 3 in the 7-year-old group. There were 6 girls and 4 boys. All of the children, except three, received a bilateral CI either simultaneously or one to two years after the first CI. The mean age at the first CI intervention was 15.88 months. In the adult group, 6 participants were included, who were all females in their early thirties.

10.21437/Interspeech.2018-1320

2.2 Procedure

A speech production task was performed by each CI child. This task was appropriate to elicit spontaneous, while still controlled, prosodic focus productions. It was provided by an ongoing research work on NH children [12]. Specifically, the task consisted in playing a computer-based semi-structured game, where children interacted with the caregiver in the presence of the experimenter. We elicited, using visual elements, sentences containing words that were contrasted in the discourse. For example, in the sentence *Prends le* ***bonnet***$_F$ *orange* ('take the orange **hat**$_F$') the word in bold and followed by an "F" marking indicates the element under focus. Focused elements could be either a disyllabic object noun (N), familiar to the children, or a disyllabic adjective (A), which is "violet" or "orange", within a noun phrase (NP).

Recordings were performed as part of the yearly speech therapy follow-up at *La Timone* Children's Hospital while prosodic focus production was performed within regular speech evaluation tests. Recordings were performed in a quiet room while using a high definition microphone and recording device. Recordings started with a training phase in order to familiarize the children with the task, making sure that they understood it, as well as to make them more comfortable during the whole recording session. After the training phase, 12 trials per condition were performed in a random order. Focus conditions were either broad focus (BF), contrastive narrow focus on the noun (FN) or on the adjective (FA) and corrective narrow focus on noun (CFN) or on the adjective (CFA). Caregivers interacted with the children in order to produce the same words in the same conditions following the children's production. There were 408 produced NP included in the analysis and 19 rejected in the CI group. In the adults' group, 356 NP were included in the analysis while 10 were excluded.

2.3 Measures

Phonemic and syllabic segmentation was automatically carried out through SPPAS® [14] and manually corrected by the main investigator. Segmental durations were calculated using Praat® scripts [15].

Prosodic characteristics of the target utterances were labelled. The duration of each word was compared in the different focus conditions and levels, and between the children and adults' groups. The analysis was performed both according to chronological age of CI children and to actual hearing age. Hearing age corresponds to the hearing experience of the CI child and is calculated as the time elapsed between the first CI external device fitting (usually three weeks after the surgery) and the date of the recording. Statistical analysis was conducted on R Studio® software [16] using mixed effects logistic regression.

3. Results

In the CI group, the 3 children aged 4 could not be included in the analysis because they did not produce any utterance. Specifically, the first child in this group had problems in concentration and she could not even finish the training phase. This child had a hearing age of 2.8 years and had severe mastoiditis after the CI surgery which delayed the external processor fitting. The second child (hearing age of 3 years) could not proceed with the test because he had difficulties in concentration and did not want to perform the tasks. Finally, the third child (hearing age of 3.34 years) had difficulties in producing comprehensible words and this was judged by the speech therapist as being the result of bilingual input at home (French and Arabic).

3.1 Duration analysis of children productions

Figure 1 shows the distribution of the duration (in ms) of NP elements (Noun and Adjective), according to focus conditions for the entire CI group (excluding 4-year-old children).

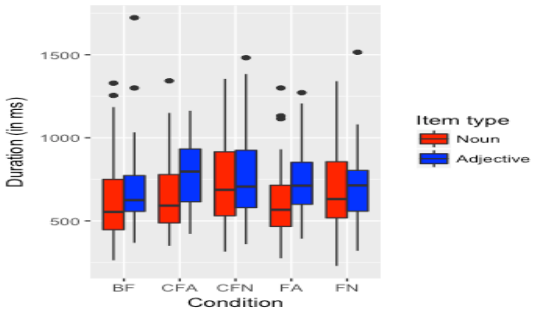

Figure 1: *Duration of Noun and Adjective according to focus condition in CI children (BF: Broad focus; CFA: Corrective narrow focus on adjective; CFN: Corrective narrow focus on noun; FA: Contrastive narrow focus on adjective; FN: Contrastive narrow focus on noun).*

Notice that, as expected by the structure of the French AP, the adjective has a tendency to be longer than the Noun, independent of the focus condition. This is because it is the last word in the AP, hence being marked by preboundary lengthening [17].

In order to evaluate the statistical significance of our results, we have first analyzed the effect of specific focus conditions on the duration of the NP elements, without taking into account the age factor. To do so, we have used linear mixed effect models, with random effect factors being participants and sentences (produced NPs), and fixed factors being Focus Condition (with 5 levels: "BF", "FN", "FA", "CFN" "CFA") and word position in the NP ("Position" with 2 levels: "1N" for noun and "2A" for adjective), as well as their interactions. Our dependent variable was duration (in ms) of the whole word. We chose BF as our baseline condition and N as the reference word position ("1N"). The retained model was: *Duration ~ Condition*Position + (1 | Participant) + (1 | Sentence)*

As expected, we found that noun duration was significantly longer in FN than BF utterances (β=72.82, t=2.744, p<.01), and in CFN than BF (β=116.48, t=4.320, p<.001). On the other hand, duration in CFN was longer than in FN with a tendency towards significance (β=43.7, t=1.809, p=0.07). Regarding Adjective duration, we found that it was significantly longer in FA than BF (β=56.44, t=2.081, p<.05), in CFA than BF (β=113.14, t=4.194, p<.001), and in CFA than in FA (β=56.7, t=2.366, p<.05).

We hence evaluated the effect of the interaction between focus condition and the type of the NP element on word duration. A duration was significantly longer than N duration in BF condition (β=61.98, t=2.584, p<.01), with an estimated difference of 61.98 ms. The interaction between focus condition and word type was also statistically significant in FA (β=68.14, t=2.003, p<.05) and in CFA (β=68.69, t=2.050, p<.05) conditions. A duration was almost 130 ms longer than N duration in FA and CFA cases, due to the presence of narrow focus. The difference between the duration of A and N in FN was only 2 ms with a marginally significant interaction (β=-60.23, t=-1.781, p=0.07527). This is expected given that here it

is the Noun that is focused. In CFN (β=-30.48, t=-0.896, p=0.37070), the interaction was not statistically significant though A duration was 31.5 ms longer than N duration. Note that greater amount of lengthening in the Adjective is expected given its position within the AP.

3.2 Duration analysis of adult's productions

Figure 2 shows the distribution of the duration (in ms) of NP elements according to focus conditions in the adult caregivers' group.

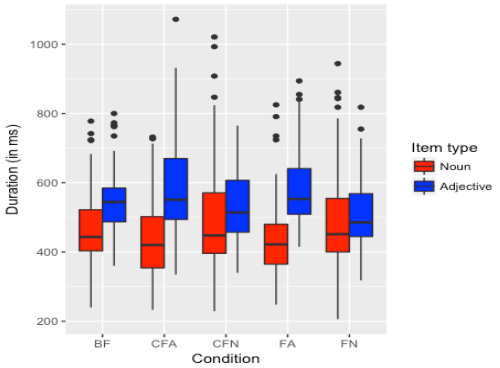

Figure 2: *Duration of Noun and Adjective according to focus condition in caregivers. (BF: Broad focus; CFA: Corrective narrow focus on adjective; CFN: Corrective narrow focus on noun; FA: Contrastive narrow focus on adjective; FN: Contrastive narrow focus on noun).*

In order to test the statistical significance of these results, we performed the same statistical analyses as for the children data. Results showed that N duration was longer in CFN than in FN, though the difference was not statistically significant. N duration in FN was marginally longer than in BF. As for the duration of N in CFN, it was on the other hand significantly longer than in BF (β=37.79, t=2.377, p<.05).

As for Adjectives, adults lengthened them significantly more in CFA relative to BF cases (β=49.71, t=3.116, p<.01). They also lengthened them in CFA more than in FA and in FA more than in BF, though those results did not reach statistical significance (Table 1).

Table 1: *Duration of noun and adjective in each related focus condition in adults (6 participants). (BF: Broad focus; CFA: Corrective narrow focus on adjective; CFN: Corrective narrow focus on noun; FA: Contrastive narrow focus on adjective; FN: Contrastive narrow focus on noun; A: Adjective; N: Noun; *: p<.05; **: p<.01)*

Item type	Focus conditions	Estimate (β)	Standard error	T value	P value
N	FN vs BF	27.08	15.55	1.741	0.082
	CFN vs BF	37.79	15.90	2.377	0.018 *
	CFN vs FN	10.711	12.810	0.836	0.403
A	FA vs BF	29.70	16.06	1.850	0.065
	CFA vs BF	49.71	15.95	3.116	0.002 **
	CFA vs FA	20.007	12.482	1.603	0.109

When analyzing the interaction between focus condition and word type, we found that A duration was significantly longer than N duration in BF (β=77.37, t=6.114, p<.001). This effect was present in all the conditions and was statistically significant. A duration was 146 ms longer than N duration in FA condition with a significant interaction (β=68.63, t=3.835, p<.001) and 155.6 ms longer in CFA, which was also significant (β=78.23, t=4.474, p<.001). Moreover, A duration was longer than N duration in FN and CFN. The difference between A and N in FN was almost 14 ms while in CFN it was almost 19 ms. The interaction between condition and word type was statistically significant in FN (β=-63.50, t=-3.548, p<.001) and in CFN (β=-58.62, t=-3.264, p<.01).

3.3. Prosodic focus in CI children compared to adults

In order to compare between the performances of CI children and adults, we compared the pattern of prosodic focus in each CI age group to the adults' productions. Figure 3 shows the distribution of the duration (in ms) of NP elements, according to focus condition for each CI age group.

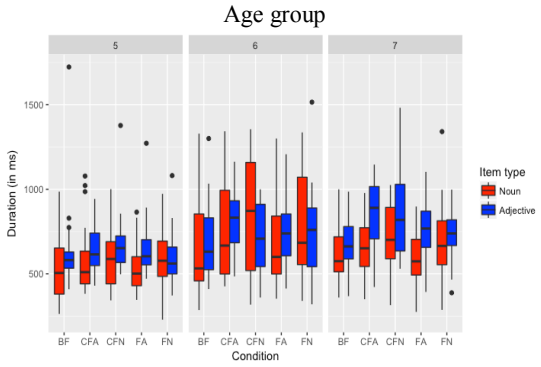

Figure 3: *Duration of Noun and Adjective according to focus condition in each age group of cochlear implanted children. (BF: Broad focus; CFA: Corrective narrow focus on adjective; CFN: Corrective narrow focus on noun; FA: Contrastive narrow focus on adjective; FN: Contrastive narrow focus on noun).*

Adult patterns of speech production were adopted by 6- and 7-year-old CI children when comparing word duration across conditions, given that they lengthened N duration in CFN relative to BF, and A duration in CFA vs BF. In other words, target word duration was longer in corrective narrow focus than in broad focus conditions. As for the duration of A vs N, only 7-year-olds adopted the same adult pattern of phrasing in all focus conditions and levels in lengthening A duration more than N. Their hearing age was no more than 6 years and 2 months. However, no statistical comparison can be done between children and adults due to the small number of participants in each CI age group, so that these results can only be considered as trends.

4. Discussion

This study is a first attempt to evaluate prosodic focus production in early CI French speaking children. We found that CI children lengthen significantly the duration of the noun in corrective narrow focus (CFN) relative to broad focus (BF) utterances , as well as in contrastive narrow focus (FN) than in BF. The difference between corrective and contrastive narrow focus did not reach significance. CI children also showed significantly longer duration for the adjective in corrective narrow focus (CFA) than in contrastive narrow focus (FA) and

BF utterances, as well as longer durations in FA than in BF conditions.

When analyzing individual CI age groups, 5-year-old children start to lengthen the duration of the target focus word across focus conditions. In the 6-year-old group of CI children, lengthening of adjective duration was greater than for the noun in all conditions except in CFN, in which the duration of the noun was longer than adjective duration. Seven-year-old children were able to lengthen the duration of both adjective and noun in corrective narrow focus more than in contrastive narrow focus and broad focus, and, in turn, in contrastive narrow focus more than in broad focus. They also lengthened adjective duration more than noun duration in all conditions. Note that age group differences are only trends because statistical significance could not be tested given that each age group was composed of only 2 to 3 participants.

In order to determine if French CI children of our study follow the adult pattern of prosodic focus lengthening, we compared them first to their caregivers' productions. In our series, we found that 5-year-old CI children start to produce the same pattern as adults by lengthening the duration of the adjective relative to the noun in all conditions except in FN, while 6-year-old CI children start this trend in all conditions except in CFN where the duration of the noun was longer than that of the adjective. The 7-year-old CI children, on the other hand, were found to better adopt the adult pattern, since they lengthened the duration of the phrase-final adjective more than the noun in all focus conditions and levels, similar to the adults. They also produced focus by means of duration lengthening of the target focused word (Noun or Adjective) between different focus conditions. Note that children in the 7-year-old group had a hearing age of no more than 6 years and 2 months.

Next, we wanted to see if early CI French children perform as their age match NH peers. We compared their performance by age group to the results of 4- and 5-year-old NH French children doing the same production task in the work of Esteve-Gibert et al [12]. In their paper, authors found that 4- and 5-year-old French NH children were unable to produce adult-like patterns of prosodic focus, while constantly lengthening the adjective in all conditions. This is due to implementing AP structure and preboundary lengthening given that in those utterances the adjective was always AP-final. The oldest child in the series was 5 years and 9 months old [12]. Compared to the NH children, only 7-year-old CI children were able to correctly lengthen the duration of the adjective relative to the noun in all focus conditions and levels. If we take into consideration hearing age, these children are at the beginning of the NH 6-year-old range. These results hence show that the pattern of lengthening the phrase-final word is mastered later in CI than in NH children.

In a study on conversational speech, Lenden and Flipsen [8] found abnormalities in word and sentence stress in six CI children aged 3 to 6 years. Phrasing, voice quality, and pitch were relatively unaffected. All CI children included in that study received an implant before the age of 3 and no analysis was performed according to different age groups nor to hearing age [8].

A recent study on phonetic production of prosody in disyllabic babble and first words in English CI and NH infants (between 6 months to 2 years of age), showed that CI infants use weaker acoustic prosodic cues. Specifically, CI infants mainly show difficulties in F0 production in line with the limited pitch perception through the CI device. Duration differences between the 2 groups were on the other hand due to greater articulatory difficulties for CI infants. NH infants were able to prolong phrase-final words and syllables while CI infants showed difficulties in this [5].

Evidence from the literature shows that the acquisition of word focus is prone to be more effortful for CI children, though not impossible. Though spectral and temporal resolution of the implant devices does not permit adequate F0 perception or changes in intensity, durational characteristics of syllables seem to be available to CI users [5].

In order to facilitate the acquisition of prosodic cues to focus, age at CI is an important factor to take into consideration due to the neuroplastic and neurolingistic dynamics in the development of the child. It is in fact optimally beneficial to offer CI during the critical period of cortex neuroplasticity in which speech and language are acquired and mastered [18]–[20]. Cochlear implantation performed before the age of 2 and optimally before 12 to 18 months has shown better results in terms of hearing and language development [21].

It is important to mention the limitations of our study. First, the small sample sizes can limit the extrapolation of the results to a larger population. The number of participants in each CI age group was in fact limited to 2 or 3 subjects. Second, the absence of a pediatric NH group performing the same task can limit our conclusions. Third, the use of a whole-word-based analysis of prosodic cues to focus needs to be followed up by a syllable-based analysis.

5. Conclusions

This is the first study testing the acquisition of prosodic cues in marking focus in pre-lingually hearing impaired French speaking children implanted with a CI before the age of 18 months. We found that CI children start acquiring adult patterns of prosodic focus and phrasing at the chronological age of 7 years (hence at hearing age of 6 years and 2 months). Specifically, they employ lengthening of the focused word in corrective narrow focus to a higher degree than in contrastive narrow focus which in turn shows longer duration than in the broad focus condition.

This study can be considered as a step towards understanding the rehabilitation needs of CI children for speech and language development. Independent of the specific difficulties in these children, a constant language stimulation by their environment is crucial in order to offer them a better social integration. The role of caregivers in the rehabilitation was not studied in this work, though it is fundamental in helping CI children gain a normal interactive life.

6. Acknowledgments

We would like to thank Nuria Estève-Gibert for providing us with the production task used in her research on normal hearing children in order to use it on cochlear implanted children in our series.

We would like to thank the department of Pediatric Otolaryngology at *La Timone* Children's Hospital, Marseille, France, and the head of the department Jean-Michel Triglia, for the logistic help in order to facilitate testing cochlear implanted children.

7. References

[1] M. D'Imperio, E. Dittinger, and M. Besson, "Prosodie et intonation : notions de base et données neuro-psycholinguistiques," in *Traité de neurolinguistique*, S. Pinto and M. Sato, Eds. De Boeck, 2016, pp. 133–145.

[2] M. E. Beckman, "The Parsing of Prosody," *Lang. Cogn. Process.*, vol. 11, no. 1–2, pp. 17–68, Apr. 1996.

[3] M. Rooth, "Association with focus," *Massachusetts Inst. Technol. PhD Diss.*, 1985.

[4] R. Torppa, A. Faulkner, J. Järvikivi, and M. Vainio, "Acquisition of focus by normal hearing and Cochlear Implanted children : The role of musical experience," *Speech Prosody*, no. November 2016, pp. 1–4, 2010.

[5] M. Pettinato, I. De Clerck, J. Verhoeven, and S. Gillis, "Expansion of Prosodic Abilities at the Transition From Babble to Words," *Ear Hear.*, vol. 38, no. 4, pp. 475–486, 2017.

[6] S.-C. Peng, J. B. Tomblin, L. J. Spencer, and R. R. Hurtig, "Imitative production of rising speech intonation in pediatric cochlear implant recipients.," *J. Speech. Lang. Hear. Res.*, vol. 50, no. 5, pp. 1210–27, Oct. 2007.

[7] S.-C. Peng, J. B. Tomblin, and C. W. Turner, "Production and Perception of Speech Intonation in Pediatric Cochlear Implant Recipients and Individuals with Normal Hearing," *Ear Hear.*, vol. 29, no. 3, pp. 336–351, Jun. 2008.

[8] J. M. Lenden and P. Flipsen, "Prosody and voice characteristics of children with cochlear implants.," *J. Commun. Disord.*, vol. 40, no. 1, pp. 66–81, Jan. 2007.

[9] S. B. Chin, T. R. Bergeson, and J. Phan, "Speech intelligibility and prosody production in children with cochlear implants," *J. Commun. Disord.*, vol. 45, no. 5, pp. 355–366, Sep. 2012.

[10] C. M. Holt, I. Yuen, and K. Demuth, "Discourse Strategies and the Production of Prosody by Prelingually Deaf Adolescent Cochlear Implant Users.," *Ear Hear.*, vol. 38, no. 2, pp. e101–e108, 2017.

[11] B. Lyxell, M. Wass, B. Sahlén, C. Samuelsson, L. Asker-Arnason, T. Ibertsson, E. Mäki-Torkko, B. Larsby, and M. Hällgren, "Cognitive development, reading and prosodic skills in children with cochlear implants.," *Scand. J. Psychol.*, vol. 50, no. 5, pp. 463–74, Oct. 2009.

[12] N. Esteve-Gibert, H. Lœvenbruck, M. Dohen, and M. D'Imperio, "The use of prosody and gestures for the production of contrastive focus in French-speaking 4 and 5 year olds," *Work. "Abstraction, Divers. Speech Dyn. Munich (Germany). May 3-5*, 2017.

[13] M. Dohen and H. Loevenbruck, "Pre-focal Rephrasing, Focal Enhancement and Post-focal Deaccentuation in French," in *Proceedings of the 8th International Conference on Spoken Language Processing*, 2004, pp. 2–5.

[14] B. Bigi, "SPPAS - Multi-lingual Approaches to the Automatic Annotation of Speech," *"the Phonetician" - Int. Soc. Phonetic Sci. ISSN 0741-6164*, no. 111–112, pp. 54–69, 2015.

[15] D. Boersma, P ; Weenink, "Praat, a system for doing phonetics by computer," *Version 6.0.23, retrieved 12 December 2016 from http://www.praat.org/*. 2016.

[16] RStudio Team, "RStudio: Integrated Development for R. RStudio, Inc., Boston, MA," 2016. [Online]. Available: http://www.rstudio.com/.

[17] M. D'Imperio, J. German, and A. Michelas, "A multi-level approach to focus, phrasing and intonation in French," in *Prosody and Meaning*, G. Elordieta and P. Prieto, Eds. Berlin: De Gruyter Mouton, 2012, pp. 11–35.

[18] T. M. De Melo, J. D. Lara, M. TM, and L. JD, "Auditory and oral language abilities in children with cochlear implants: a case study," *J. Soc. Bras. Fonoaudiol.*, vol. 24, no. 4, pp. 390–394, 2012.

[19] A. Sharma, M. F. Dorman, and A. J. Spahr, "A sensitive period for the development of the central auditory system in children with cochlear implants: implications for age of implantation.," *Ear Hear.*, vol. 23, no. 6, pp. 532–9, Dec. 2002.

[20] A. Sharma, P. M. Gilley, M. F. Dorman, and R. Baldwin, "Deprivation-induced cortical reorganization in children with cochlear implants.," *Int. J. Audiol.*, vol. 46, no. 9, pp. 494–9, Sep. 2007.

[21] H. Bruijnzeel, F. Ziylan, I. Stegeman, V. Topsakal, and W. Grolman, "A Systematic Review to Define the Speech and Language Benefit of Early (<12 Months) Pediatric Cochlear Implantation.," *Audiol. Neurootol.*, vol. 21, no. 2, pp. 113–26, 2016.