



HAL
open science

Similar gaze behavior during dialogue perception in Congenitally Deaf children with Cochlear Implants and normal hearing children

Céline Hidalgo, Christelle Zielinski, Sophie Chen, Stéphane Roman, Eric Truy,
Daniele Schön

► To cite this version:

Céline Hidalgo, Christelle Zielinski, Sophie Chen, Stéphane Roman, Eric Truy, et al.. Similar gaze behavior during dialogue perception in Congenitally Deaf children with Cochlear Implants and normal hearing children. *International Journal of Language and Communication Disorders*, 2024, Online ahead of print. 10.1111/1460-6984.13094 . hal-04648806

HAL Id: hal-04648806

<https://hal.science/hal-04648806v1>

Submitted on 15 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 Similar gaze behavior during dialogue perception in Congenitally Deaf children with Cochlear
2 Implants and normal hearing children

3 Céline Hidalgo¹, Christelle Zielinski², Sophie Chen¹, Stéphane Roman^{1,3}, Eric Truy^{4,5,6,7}, Daniele
4 Schön^{1,2}

5 ¹ Aix Marseille Univ, Inserm, INS, Inst Neurosci Syst, Marseille, France;

6 ² Aix-Marseille Univ, Institute of Language, Communication and the Brain, France;

7 ³ Pediatric Otolaryngology Department, La Timone Children's Hospital (APHM), Marseille,
8 France;

9 ⁴ Service d'ORL et de Chirurgie cervico-faciale, Hôpital Edouard Herriot, CHU LYON, France;

10 ⁵ Inserm U1028, Lyon Neuroscience Research Center, Equipe IMPACT, Lyon, France;

11 ⁶ CNRS UMR5292, Lyon Neuroscience Research Center, Equipe IMPACT, Lyon, France;

12 ⁷ University Lyon 1, Lyon, France.

13 **Corresponding author:** Daniele Schön, daniele.schon@univ-amu.fr

14 **Professional Address:** Institut de Neurosciences des Systèmes, Faculté de Médecine, bd
15 Jean Moulin, 13005 Marseille, France

16 **Telephone:** 04 91 32 41 00

17 Other authors emails:

18 celinehidalgo13@gmail.com; christelle.zielinski@univ-amu.fr; sophie.chen@univ-amu.fr,
19 stephane.roman@ap-hm.fr; eric.truy@chu-lyon.fr

20

21

22

23 **Author contributions:** Conceptualization C.H. and D.S.; Data curation C.H., C.Z. and S.C.;
24 Formal Analysis C.H.; Funding acquisition D.S.; Project administration C.H.; Supervision D.S.;
25 Visualization C.H, C.Z.; Writing – original draft C.H., S.C., C.Z., S.R., E.T. and D.S.

26 **Acknowledgments:** We warmly thank all the children and their families as well as Anna
27 Montagnini for her contribution to data acquisition. Research supported by grants ANR-21-

28 CE28-0010 (DS), ANR-16-CONV-0002 (ILCB) and the Excellence Initiative of Aix-Marseille
29 University (A*MIDEX).

30 The authors do not have any **conflict of interest** to declare

31

32 **Abstract**

33 Perceptual and speech production abilities of children with cochlear implants (CI) are usually
34 tested by word and sentence repetition or naming tests. However, these tests are quite far apart
35 from daily life linguistic contexts. Here, we describe a way of investigating the link between
36 language comprehension and anticipatory verbal behaviour promoting the use of more complex
37 listening situations.

38 The setup consists in watching the audio-visual dialogue of two actors. Children's gaze-switches
39 from one speaker to the other serve as a proxy of their prediction abilities. Moreover, to better
40 understand the basis and the impact of anticipatory behaviour, we also measured children's
41 ability to understand the dialogue content, their speech perception and memory skills as well as
42 their rhythmic skills, that also require temporal predictions. Importantly, we compared children
43 with CI performances with those of an age-matched group of children with NH.

44 While children with CI revealed poorer speech perception and verbal working memory abilities
45 than NH children, there was no difference in gaze anticipatory behaviour. Interestingly, in
46 children with CI only, we found a significant correlation between dialogue comprehension,
47 perceptual skills and gaze anticipatory behaviour.

48 Our results extend to a dialogue context previous findings showing an absence of predictive
49 deficits in children with CI. The current design seems an interesting avenue to provide an
50 accurate and objective estimate of anticipatory language behaviour in a more ecological
51 linguistic context also with young children.

52

53 **Keywords:** dialogue; prediction; eye-tracking; gaze; comprehension; rhythmic skills

54 **Data availability statement:** In accordance with the Peer Reviewers' Openness Initiative
55 (<https://opennessinitiative.org>, Morey, Chambers, Etchells, Harris, Hoekstra, Lakens, et al.,
56 2016), all materials and scripts associated with this manuscript will be available on <https://osf.io>
57 following publication

58

59 **Conflicts of interest:** we do not have any conflict of interest to notify.

61 Introduction

62 Even when early implanted, children with cochlear implants (CI) show heterogeneous language
63 skills (van Wieringen & Wouters, 2015). Indeed, although they can be included in mainstream
64 schools and can, sometimes, achieve language scores close to those of children with normal-
65 hearing (NH) (Hayes et al., 2009), several children with CI still show deficits in vocabulary
66 knowledge (Lund, 2016) and morphological, syntactic or narrative skills (Boons et al., 2013) as
67 well as speech intelligibility (Habib et al., 2010), speech prosody (Chin et al., 2012) and turn-
68 taking (Kondaurova et al., 2020, 2022). The best predictors of such variability in speech and
69 language levels of children with CI seem to be the age of implantation, the child's residual
70 hearing and the parents socio-economic status (Niparko et al., 2010; Nicholas & Geers, 2018;
71 see Ruben, 2018 for a review), although other factors from the quality of the electrode-neuron
72 interface (Arjmandi et al., 2022) to the linguistic context (Lovcevic et al., 2022) seem to
73 contribute to the variability in children's speech outcomes. Nonetheless, these factors do not
74 explain the entire heterogeneity found in their language outcomes (Houston, 2022).

75 This lack of explanatory power may be partly linked to the lack of ecological validity of the tests.
76 Indeed, it has been shown that adding ecologically valid variables can increase the variance
77 accounted for executive functioning (Chaytor et al., 2006) and it is surely one of the great
78 challenges of hearing research in the years to come as it allows to assess linguistic skills in
79 close to real-life listening situations (Lunner et al., 2020; Keidser et al., 2020). Indeed,
80 perceptual and speech production abilities of children with CI are usually tested by simple word
81 and sentence repetition or naming tests (see e.g. Yoshinaga-Itano et al., 2010 or Svirski et al.,
82 2000, for language assessment in a longitudinal study or Dettman et al., 2016 for a multi-centre
83 study) but are rather faraway from the contexts they encounter in their daily lives.

84 Adopting more ecologically valid tests may also improve our understanding of language deficits
85 insofar as different processes may be at work in speech perception tests using isolated words or
86 short sentences compared to listening to continuous speech. Indeed, understanding a narrative
87 requires more complex processing than understanding words or sentences out of context
88 (Chambers & Juan, 2008, Xu et al., 2005). Moreover, in conversational contexts,
89 comprehension is even less of a low-level process but strongly involves prediction/anticipation
90 capacities at different levels of representations (Carlile & Keidser, 2020; Friston et al., 2020,
91 Pickering & Gambi, 2018). Importantly, predictions are considered to become particularly
92 relevant during degraded listening conditions of speech (Sohoglu & Davis, 2020; Strauß et al.,
93 2013) and may possibly play an important role in speech perception of cochlear implant users.
94 Thus, assessing language skills in an interactive linguistic context may improve our
95 understanding of the language outcome variability by modeling the variability of the
96 anticipatory/predictive behavior.

97 One way to indirectly measure predictions and their impact on comprehension was proposed by
98 Grisoni and colleagues (2021), but requires recording EEG during sentence processing as well
99 as controlled linguistic material. Another way to measure predictions is to study turn-taking
100 behaviour in conversation. Indeed, because the gaps between turns are of the order of 200 ms,
101 while the latencies involved in language production are much longer (over 600 ms), this has
102 been interpreted in terms of predictive anticipatory behaviour (Levinson, 2016; Stivers et al.,
103 2009). Several studies have assessed conversational interactions in hearing impaired children
104 showing differences in terms of turn-taking (Kondaourova et al., 2020, 2022) and maternal
105 prosody (Dilley et al., 2020). However, while the study of turns in ecological conditions is
106 important and highly informative, it does not easily allow to disentangle the respective
107 contribution of auditory, multisensory, linguistic and predictive processes. Recently a new
108 approach was inspired by the work of Casillas and Frank (2017) who showed that a predictive
109 behaviour is present not only in turn-taking but also when simply watching other people talk.
110 The approach consists in showing pre-recorded audio-visual dialogues of two individuals and
111 measure the participant's gaze switches, a temporally-resolved behavioural proxy of internal
112 predictions. Interestingly, a recent study found that anticipatory gaze switches to the upcoming
113 speaker were associated with a better global comprehension of the dialogue content. This
114 relationship was sensitive to the difficulty level of speech perception: the more speech dialogues
115 were difficult to perceive (level of vocoding) the more gaze behaviour was predictive of
116 discourse understanding, assessed by asking general questions at the end of each dialogue
117 (Hidalgo et al., 2022).

118 The aim of this study was to compare the temporal predictive abilities of children with CI and NH
119 in a complex linguistic context: watching a dialogue and anticipating turns on the sole basis of
120 speech and language cues (i.e. removing in the stimuli visual anticipatory cues such as hand,
121 head and eye movements). At this aim, we have used a recently developed paradigm (Hidalgo
122 et al., 2022) wherein children fitted with an eye-tracker, watch videos of two individuals talking to
123 each other. Gaze-switches from one speaker to the other serve as a proxy of children's
124 prediction abilities. In order to assess the impact of anticipatory behaviour on the global
125 understanding of the dialogues we also asked children to answer specific questions about the
126 content of the dialogues. Finally, in order to assess whether anticipatory gaze behaviour is
127 related to other forms of temporal predictions, we also collected data using musical sensori-
128 motor rhythmic tasks with different levels of complexity. Indeed there is evidence that temporal
129 predictions may share common processes in speech and music (Schön & Morillon, 2018, Patel
130 & Morgan, 2017) and that, in children with CI, the ability to reproduce rhythms with a complex
131 hierarchical structure is related to their speech perception abilities (Hidalgo et al., 2021).

132

133 Materials and Methods

134 Participants

135 Two groups of children, aged between 8 and 12 years (mean = 10.5 years; SD = 1.3 months)
136 were selected. In the experimental group, 19 children with cochlear implants were recruited from
137 two ENT paediatric French centers: one in Lyon, (Edouard Herriot Hospital) and one in
138 Marseille, (La Timone Hospital). All children (nine girls) fitted with cochlear implants were
139 French native speakers and had severe-to-profound hearing loss without additional disorders.
140 17 children were fitted with bilateral cochlear implants. For six children the precise onset of
141 deafness was unknown, but for five of them it was before 5-month age. Two of them wore only
142 one CI without a contralateral hearing aid. All children were included in regular schools (see
143 demographics for more details in Supplementary Table 1). In the control group, fifteen children
144 with normal hearing were recruited in Lyon and Marseille. These children (eleven girls) were all
145 French native speakers without any known visual, speech or cognitive disorders. This group of
146 children passed a fast pure tone audiometry at 20dB HL at 250 Hz, 500Hz, 1000 Hz, 2000 Hz,
147 4000 Hz and 8000 Hz (1 sec tones) using a staircase procedure (Expyriment version 0.10.0,
148 Krause & Lindemann, 2014). The aim of the screening was not to estimate the precise hearing
149 thresholds but simply to ensure that none of them had non-diagnosed hearing disorders.
150 Children in the control group were age-matched with children from the experimental group
151 (Wilcoxon-test between groups, $W = 72$, $p > 0.5$). Also the educational level of the parents was
152 similar across groups ($W = 68$, $p > 0.5$).

153 Two children with normal hearing were excluded from the analysis for “loss” of eye-tracking data
154 (camera losing track of the pupil > 50% of the time). Three children with cochlear implants
155 were excluded due to a difficulty in answering questions on a video (similar to the ones used in
156 the experiment) during a training session. The reasons of such exclusion criterion was that
157 children with a poor understanding of the dialogues are not be able to perform the task, namely
158 staying focus and keep looking at the video-dialogues. Three others were excluded for loss of
159 eye-tracking data (> 50%) and one for being an outlier in terms of number of blinks in the eye-
160 tracking data, leaving thirteen and twelve children in each group.

161 All children and their parents were informed of the procedure of the study, which was approved
162 by the Sud Méditerranée V Ethics Committee (ID RCB : 2020-A02556-33).

163

164 Procedure

165 **Perceptual, cognitive and rhythmic assessments**

166 For all children, the tests took place in a sound booth in Marseille and a sound treated room in
167 Lyon hospitals in the ENT Paediatric Departments. The setup and the experimenter were
168 identical in both centers.

169 **Speech perception test**

170 Children listened to and repeated 21 non-words composed of 3 to 6 syllables (mean = 4.14
171 syllables, SD = 1.28, ex. *paloti*) extracted from the TERMO French test (Busquet &
172 Descourtieux, 2003). Repetition of the stimuli was allowed when the child was not focussed, but
173 occurred rarely (~5%). Scoring comprised an estimate of the accuracy for the global item (0 or
174 1, max score = 21) and the number of correct syllables reproduced, max score = 87).

175 **Span tests**

176 In order to assess the verbal working memory we used stimuli from the Wechsler Intelligence
177 Scale for Children IV test (Wechsler, 2005). Children listened to and repeated in the same order
178 (direct span) or in the reverse order (indirect span) series of digits of increasing length. The test
179 ended when the child made two consecutive errors at a given length. Scoring takes place at the
180 item level (correct or incorrect). These digits were scored by the experimenter as 0 (failure) or 1
181 (success). The maximum score children could get for direct and indirect span was 14 and 16.
182 Actual scores were transformed in % of correct responses with respect to the total score. The
183 stimuli of the speech perception and memory span tests were delivered orally by the
184 experimenter with lip-reading available.

185 **Rhythmic tests**

186 In order to assess the temporal predictive and sensorimotor synchronization abilities we used
187 different rhythmic tests. First children tapped as synchronously as possible along a metronome,
188 for one minute (beat = 666ms). Then, they repeated the same task with a musical excerpt
189 ("Chocolate" by the band The 1975 ; beat = 603ms). Finally we also assessed the ability to
190 reproduce complex rhythms. In this task, children listened to a complex rhythm that is repeated
191 eight times in a row. Children were asked to imitate the whole rhythmic pattern (and not the
192 beat) by tapping along the rhythm. We used eight different patterns. Importantly, the use of a
193 timbre with a short attack time and a short duration ensures a low spectral complexity of these
194 stimuli.

195 For the sensori-motor synchronization on the metronome and the musical beat we used custom
196 Matlab scripts to project taps in a circular space, wherein each tap is attributed a specific angle
197 relative to the expected beat time (= 0° on the circle). The resulting vector is calculated to
198 quantify sensori-motor synchronization to the beat (for details, see Sowiński & Dalla Bella,
199 2013). The length of this vector indicates the synchronization consistency (from 0 to 1; 0 =
200 uniform random distribution; 1 = perfect phase synchronization).

201 For the sensori-motor synchronization on complex rhythmic patterns, we considered only the
202 last three repetitions in the scoring; the first five were considered as the learning phase. Note
203 however that analyses were run on approximately ninety taps per child. First, we extracted the
204 temporal series of tapping from the audio recording. To this aim we used the SciPy function
205 find_peaks with a minimal distance of 100 ms and a prominence of 2 SD. We used an objective
206 procedure as in Hidalgo et al., 2021, to score children's performances. This measure is based
207 on computing a correlation between children's tap timing and the rhythms to be reproduced.
208 The score of each performance was defined as the maximum value of these correlations.

209 Children's speech productions (non-words, digits and sentences repetition) and tapping were
210 recorded with a digital recorder (Zoom , Handy Recorder, H4n). The recorder was placed in a
211 wooden box used for the tapping.

212 The speech from sentences and video-dialogues as well as rhythmic sounds were delivered
213 through loudspeakers (Behringer MS 16) which were located on both sides of the screen.

214 **Experimental task: prediction during dialogue perception**

215 **Stimuli**

216 The videos consisted of three audiovisual stimuli of ~4.5 minutes each, showing the head of a
217 man and a woman having a conversation. The themes of the pre-scripted dialogues differed
218 and were all original scripts (thus completely unknown to children), although inspired in their
219 style from novels and pieces of theater for children. Dialogues were as engaging as possible to
220 maintain children's attention during the whole duration of the videos. Dialogues contained on
221 average 59 turns (min = 46; max = 66). Turns lasted on average 3.7 s (SD = 2.7) and syllabic
222 rate was 5Hz. 60% of these turns were questions/responses adjacency pairs (min = 56 %; max
223 = 66%) in order to elicit predictive behaviours as suggested in Casillas & Frank, 2017 (Casillas
224 & Frank, 2017). We used Final Cut Pro X to set the gap duration at speakers' turn to 500ms in
225 order to have homogeneous turn conditions and allow anticipatory gaze behaviour (cf.
226 Foulsham et al., 2010; Casillas & Frank, 2017). Importantly predictive eye behaviour could
227 essentially be based on speech and linguistic cues because we controlled and removed all
228 possible visual cues such as head and eye movements as well as anticipatory mouth and lips

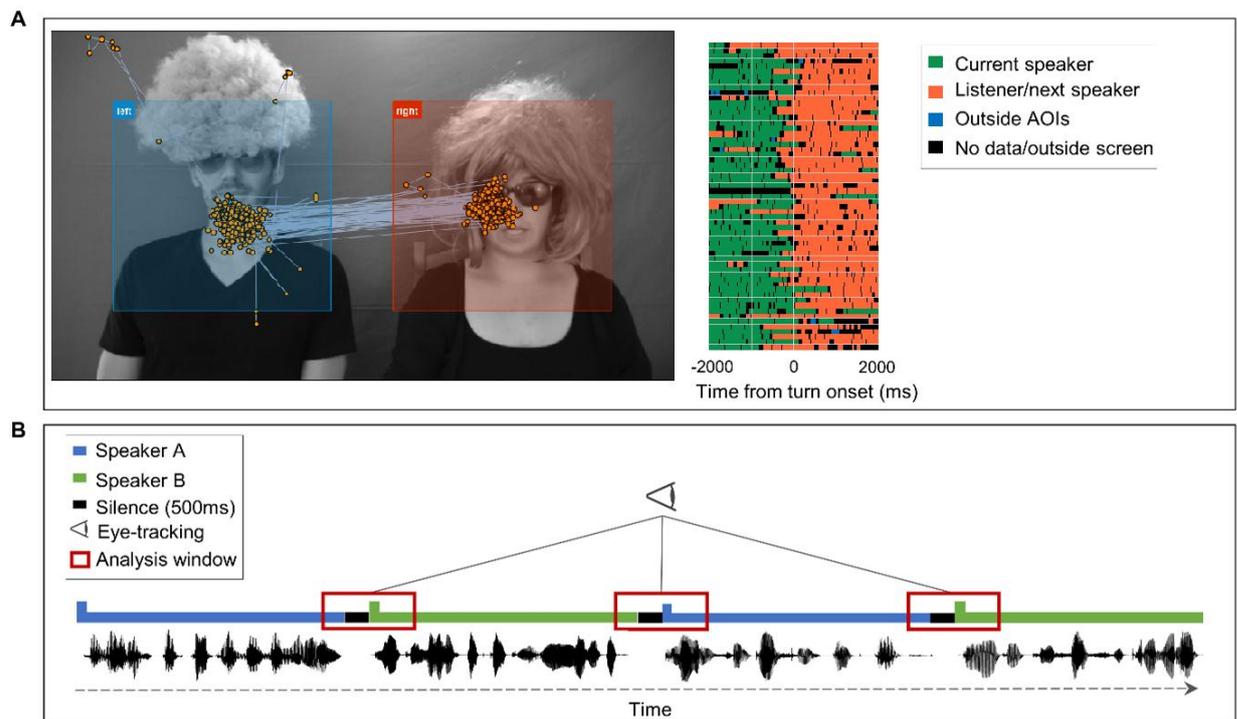
229 movements were absent in the present stimuli. To avoid head movement when reading the
230 dialogues, the speakers sat on a chair and their head was maintained still by a hard hat fixed to
231 the chair. To hide eye movements each speaker wore sunglasses and a wig to hide the hard
232 hat (Figure 1A left). Furthermore, in order to discard further possible spurious visual cues, the
233 non-speaking speaker was “frozen” via video editing whenever a head or lip movement
234 occurred before the turn. This happened in approximately ~20% of the turns. In order to avoid
235 sharp jumps in the images of the upcoming speaker such freezing took advantage of fading and
236 morphing technics (mMorphCut plug-in ; FINAL CUT pro X).

237 **Procedure**

238 Children were equipped with the monocular head-mounted Pupil core eye-tracker (Pupil Labs
239 GmbH, Berlin, Germany) with their chin rested on a chinrest, fixed at the table at ~60 cm from
240 the screen (24” with a resolution of 1920 x 1080px and a refresh rate of 100Hz). Pupil Capture
241 software (v 3.0.7) launched the eye-tracker calibration and recording. Before the calibration
242 procedure, we ensured that QRcodes stuck on the four screen’s corners (to detect the screen
243 surface) and children’s pupils were well detected by the eye-tracker world and eye cameras
244 respectively. Then we run a 5 target calibration followed by a 4 target test procedure. During the
245 task, children were instructed to listen and watch each video carefully in order to be able to
246 answer twelve questions after each dialogue. These were close questions and could be scored
247 as failure or success (0/1), allowing a maximum score of twelve per dialogue. The video and
248 associated questions were delivered using the OpenSesame software (Mathôt et al., 2012).

249 **Stimuli and eye-tracking synchronization**

250 A custom-made sync box based on an Arduino micro-controller ensured the synchronization
251 between the eye-tracking data and the auditory signal embedded in the video file. The stereo
252 sound went from the stimulation laptop to the sync box (audio cable). There, the stereo was split
253 and the first channel containing the speech signals of the dialogues sent directly to the
254 loudspeakers. The second channel signal went through the micro-controller. This channel
255 contained an audio spike placed at the onset and at the end of each dialogue. The Arduino
256 detected the audio trigger and sent a code to the serial port that was subsequently detected and
257 transmitted via a Python script to the Pupil Core network API, an API which supports fast and
258 reliable communication and real-time access to data. This ensured, for each dialogue, a precise
259 synchronization of the eye-tracking.



260

261 Figure 1 . Schematic illustration of the task and analyses. A. On the left, snapshot of one
 262 video illustrating the areas of interest (AOI) and the fixation points of a single participant during
 263 the whole video duration. On the right, gaze behaviour of the same participant showing the gaze
 264 switch for every turn (y axis) of a single dialogue with respect to turn onset (time zero). B.
 265 Schematic illustration of the turns in the dialogue, the controlled gap between turns (500ms) and
 266 the gaze switch window of analysis.

267 Eye-tracking data processing

268 Gaze data relative to the screen surface were then extracted using the Pupil Player software (v
 269 3.0.7) and gaze switches were detected using custom Matlab scripts. In this study we used the
 270 gaze switch from one speaker to the other as a proxy of anticipatory behaviour with respect to
 271 the upcoming turn onset (see Figure 1). We considered a temporal window around the turn
 272 allowing to keep both anticipatory and non-anticipatory gaze switches, resulting in a -1/+1s
 273 window around the turn onset. Note, nonetheless that, as expected, most shifts occurred in the
 274 inter-turn gap, that is before the turn onset and this has been interpreted as predictive behaviour
 275 insofar as the gaze switch does not take place in reaction to the upcoming turn (Foulsham et al.,
 276 2010; Keitel et al. 2013; Casillas and Frank 2017, see Figure 1A, right panel).

277 Moreover, we also applied three supplementary criteria to filter spurious gaze switches. First,
 278 gaze before and after the switch should fall within an area of interest (AOIs), defined as a
 279 stationary rectangle surrounding each face (see Figure 1A left). Second, the gaze switch should
 280 be preceded by at least 100ms fixation on the current speaker. Third, it should be followed by at

281 least 150ms fixation to the next speaker. Concerning the possible back and forth gaze
282 behaviour preceding turn (~5% in our data), we kept only the first gaze switch.

283 In short, we computed a gaze switch latency relative to the upcoming turn onset time. A positive
284 value indicates thus a gaze switch after the turn onset, while a negative value indicates a gaze
285 switch before the turn onset (Figure 1A right).

286 **Statistical modeling**

287 All statistical analysis was computed using R (R Core Team, 2021) and the lme4 package
288 (Bates et al., 2015). We computed Linear Mixed Models (LMM) with *group* as a predictor
289 variable, to evaluate the difference of performances between children with Cochlear Implant (CI)
290 and children with Normal-Hearing (NH) on: 1) the number of correct syllables reproduced during
291 the non-word repetition task, 2) the correlation's coefficients obtained from the complex rhythm
292 sequences, 3) the gaze switches during the video-dialogues. We modelled *subject* as a random
293 effect in all analyses. We added *item* as a supplementary random effect in the model for the
294 non-word repetition. Because the complex rhythm task may be sensitive to *musical practice*
295 (number of years of) and *working memory*, we modelled these two variables as a confounding
296 factor in the LMM.

297 To evaluate the differences between groups on 1) the comprehension score, 2) the direct span
298 and 3) the indirect span measures, we computed Generalized Linear Mixed Models because of
299 the binary nature of the variables (0 or 1), with *group* as a predictor variable and *subject* as
300 random effect. Statistical significance of the fixed effects was assessed by model comparison
301 using the Akaike Information Criterion, thus arbitrating between complexity and explanatory
302 power of the models. Normality and homoscedasticity of the residuals of all the models were
303 systematically visually inspected. Reported p values are Satterthwaite approximations obtained
304 with the lmerTest package (Kuznetsova et al., 2017).

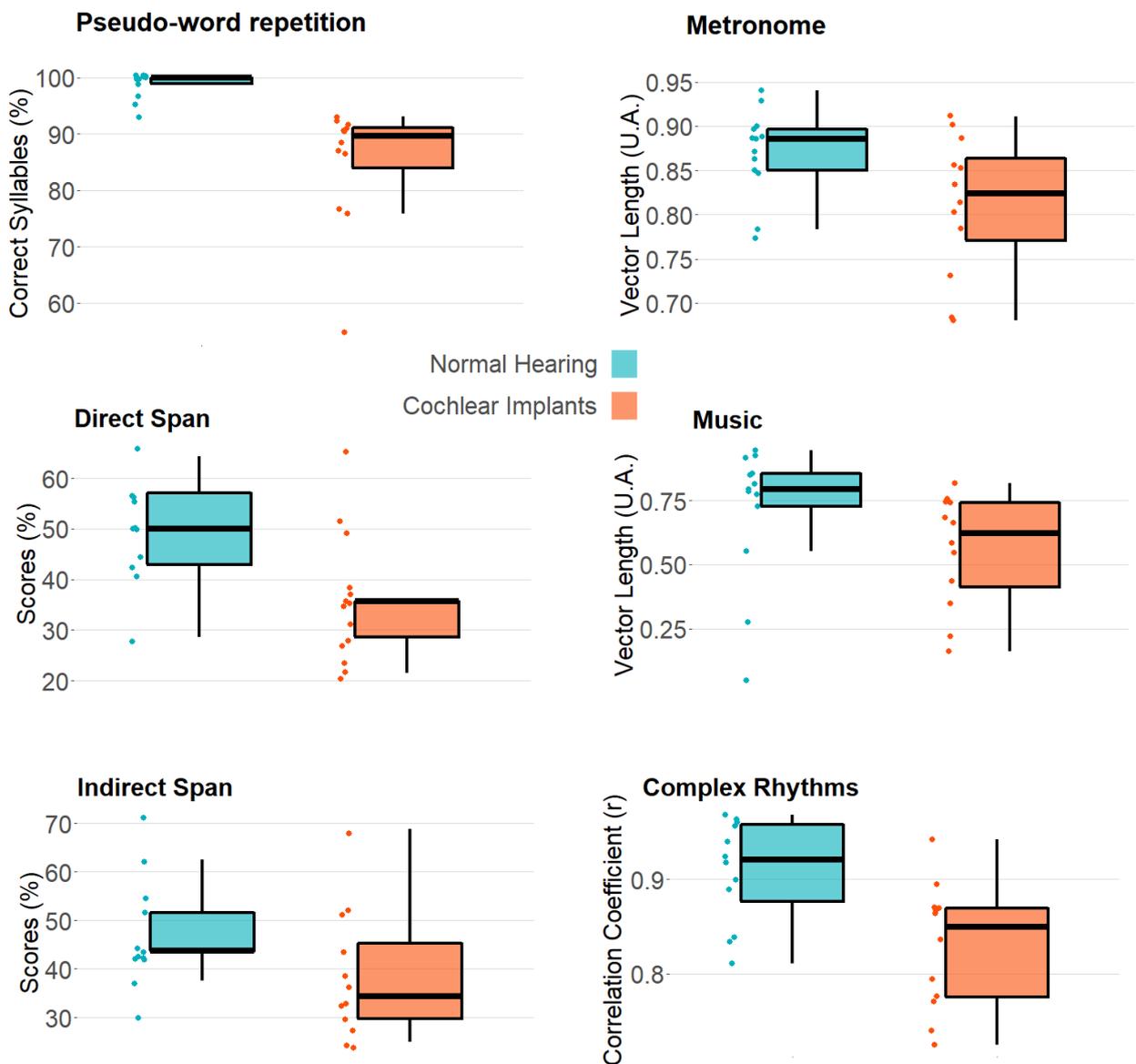
305 For the metronome and musical beat synchronisation tests that yield one summary statistics for
306 participant (vector length, that is the response consistency), we computed Linear Models, with
307 *group* as a predictor variable and *musical practice* as a confounding factor.

308 Finally, in order to analyse the relationship between comprehension level, non-word perception
309 skills, gaze switch latencies during videos and synchronization skills on complex rhythmic
310 sequences, we computed Spearman's coefficient correlations between these measures, for
311 each group separately (12 subjects in each group). When relevant, we computed a Linear
312 Model for the significant correlations also introducing the covariates of interest in the model. All
313 the assumptions of the models were tested using the Global Validation of Linear Model
314 Assumptions package and for each group comparison, we computed Cohen'd effect size.

315 **Results**

316 **Perceptual, cognitive and rhythmic performances**

317 In the non-word repetition task, CI children accuracy (number of correct syllables) was lower
318 compared to NH children ($\beta = -0.48$, $SE = 0.073$, $t = -6.55$, $p < 0.001$, Cohen's $d = 0.44$, see Fig.
319 2). Children with CI also showed poorer verbal memory performances than children with NH as
320 assessed by the direct span test ($\beta = -0.80$, $SE = 0.22$, $t = -3.58$, $p < 0.001$, Cohen's $d = 0.39$,
321 see Fig. 2). Groups also differed in terms of working memory capacities, as assessed by the
322 indirect span test, although the effect size was rather small ($\beta = -0.39$, $SE = 0.21$, $t = -1.86$ $p =$
323 0.063 , Cohen's $d = 0.19$, see Fig. 2). Sensori-motor synchronisation skills were also poorer in
324 children with CI compared to those of children with NH. They tapped less consistently on the
325 metronome beat ($\beta = -0.058$, $SE = 0.026$, $t = -2.26$, $p = 0.033$, Cohen's $d = 0.903$, see Fig. 2)
326 but this difference was no longer significant when controlling for musical practice ($p = 0.227$).
327 Concerning synchronisation consistency on the musical beat, we found no difference between
328 groups ($p = 0.12$) even when controlling for the musical practice ($p = 0.45$). However, children
329 with CI's synchronisation abilities were not as good as those of children with NH on complex
330 rhythmic sequences ($\beta = -0.079$, $SE = 0.025$, $t = -3.18$, $p = 0.004$, Cohen's $d = 0.864$, see Fig.
331 2) and this difference was still significant when controlling for working memory and musical
332 practice ($\beta = -0.077$, $SE = 0.025$, $t = -3.01$, $p = 0.006$).



333

334 Figure 2. Boxplots of the results of the speech repetition, direct and indirect span, tapping to a
 335 metronome, tapping to music and complex rhythm reproduction tasks for children with normal
 336 hearing and cochlear implants.

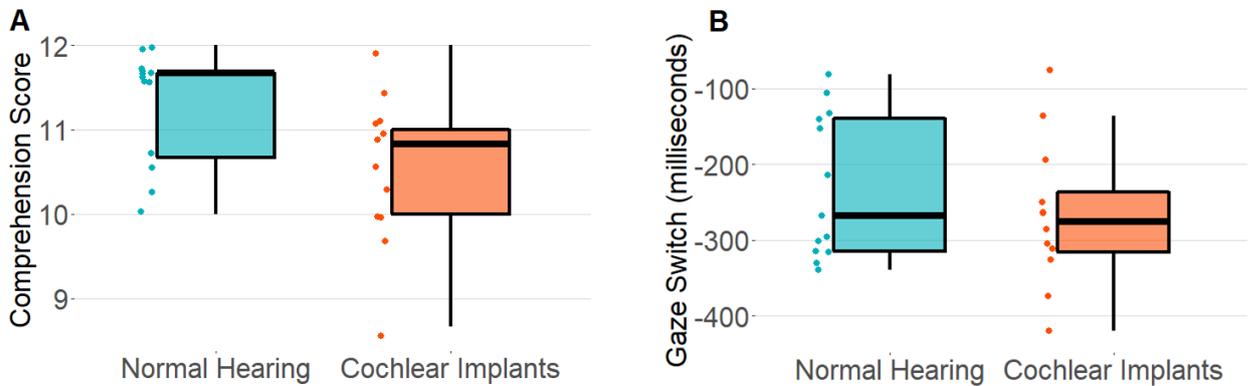
337 **Comprehension and temporal prediction performances**

338 Compared to the children with NH, dialogue comprehension scores were lower for children with
 339 CI ($\beta = -0.89$, $SE = 0.34$, $t = -2.61$, $p = 0.009$, Cohen's $d = 0.23$, see Fig. 3A). This was despite
 340 the fact that children with CI spent more time looking at the current speakers than their NH peer
 341 ($W = 274$, $p < 0.001$). When investigating the gaze behaviour in the two groups, all children
 342 showed gaze switches well before the onset of the upcoming turn. However, we did not find
 343 significant latency differences between children with CI and their NH peers, ($\beta = -36.28$, $SE =$
 344 38.54 , $t = -0.94$, $p = 0.36$) (see Fig.3B).

345

346

347



348

349 Figure 3. Boxplots of the results for the comprehension score and the gaze switch in the
350 experimental task wherein children are first watching two people having a conversation. The
351 comprehension score for each child is reported as the average of the correct answers across
352 the three dialogues (based on twelve questions each). The gaze switch for each child is
353 reported as the average time for gaze switch around the 177 turns (negative values indicate a
354 gaze switch before the onset of the next speaker).

355 Correlations

356 Correlations on separated groups revealed in children with CI a positive link between the
357 capacities to understand dialogues and their perceptual language skills, measured by pseudo-
358 words repetition ($r = 0.71$, $\beta = 0.20$, $SE = 0.061$, $t = 3.23$, $p = 0.009$). Gaze anticipatory
359 behaviour while watching the dialogues was also associated to better perceptual language skills
360 ($r = -0.65$; $\beta = -6.58$, $SE = 2.41$, $t = -2.73$, $p = 0.021$). Gaze behaviour was also associated to
361 comprehension scores ($r = -0.53$, $p = 0.074$), and this relation became even more apparent
362 when taking into account memory capacities of children with CI as a co-variate ($\beta = -27.63$, SE
363 $=10.63$, $t = -2.60$, $p = 0.028$). Sensorimotor synchronization abilities and in particular
364 performance in the complex rhythm reproduction task were also associated to gaze behaviour
365 ($r=-0.39$) and perceptual language skills ($r=0.43$) although these medium size correlations did
366 not reach significance. By contrast, we did not find any significant correlation between variables
367 in children with NH (all $ps > 0.05$) mostly due to a ceiling effect or smaller variability of their
368 performance in most of the skills assessed (cf Figure 2).

369 **Discussion**

370 In this study, we investigated speech-turn temporal predictions abilities in congenitally deaf
371 children with CI. To this end, we recorded eye movements and used gaze switch latencies as a
372 proxy of anticipatory behaviour while children were watching videos of two individuals during a
373 dialogue wherein all visual anticipatory cues were removed. Moreover, in order to better
374 understand the basis and the impact of anticipatory behaviour, we also measured children's
375 ability to understand the dialogue content, their speech perception and memory skills as well as
376 their rhythmic skills. Importantly, we compared children with CI performances with those of an
377 age-matched group of children with NH.

378 As expected, children with CI revealed poorer speech perception and verbal memory abilities
379 than NH children. They also showed poorer sensori-motor synchronisation skills in particular on
380 complex rhythmic sequences. On the experimental task, children with CI showed poorer
381 dialogue comprehension, spent more time looking at the current speaker but showed no
382 differences in gaze anticipatory behaviour compared to children with NH. Interestingly, in
383 children with CI only, we found a significant correlation between dialogue comprehension,
384 perceptual skills and gaze anticipatory behaviour.

385 A first relevant element in discussing these results is that the task we used, inspired by previous
386 work of Casillas and Frank (2017) and further refined and assessed in adults (Hidalgo et al.,
387 2022) allows to assess both language comprehension and anticipatory verbal behaviour in a
388 rather ecological setup. In the present experiment, the ecological potential of the setup was not
389 fully exploited because we chose to control and remove all the visual cues that are known to
390 play an important role in conversational contexts (hand, eye, head and mouth movements
391 preceding speech). Nonetheless, this has the advantage to allow interpreting the findings in
392 terms of auditory (speech and language) predictions. Thus, gaze behaviour in the current setup
393 is mostly determined on the basis of linguistic information that here comprised of both speech
394 and lip reading inputs.

395 Another relevant aspect is the fact that, while a few previous studies have used indirect
396 measures (e.g. manipulating the predictability of the context) to assess the ability of children
397 with CI to predict words in speech (Smiljanic and Sladen, 2013; Holt et al., 2016), here we use a
398 direct measure, eye movements, as a proxy of predictions in a rather complex language task.
399 The use of gaze latencies had already been successfully employed in a visual word completion
400 task, showing a reliable use of semantic informative cues to predict upcoming words in children
401 with hearing loss (Holt, et al., 2021; Blomquist et al., 2021). However, the current setup, by
402 tracking continuous gaze shift during a dialogue more closely resemble to a real-life situation
403 compared to sentence completion using a picture.

404 That all children showed average gaze switches well before the onset of the upcoming turn can
405 be interpreted in terms of predictive behavior. Even more so when considering that gaze switch
406 requires a cognitive process (decision), motor preparation and the saccade realization. The
407 average latency of a saccade made towards a sudden visual onset target is around 200 ms
408 (Hutton, 2000) and the latency goes up to 300ms for voluntary saccades (e.g. arrow cue,
409 Walker et al., 2000). In the present study the average gaze switch takes place 200ms after the
410 end of the turn (and 300ms before the upcoming turn). Therefore one can consider that the
411 preparation of the gaze switch begins before the current speaker has finished to speak. The
412 lack of group differences concerning linguistic predictive skills goes somewhat against the idea
413 that children with cochlear implants have difficulties extracting structure from and learning
414 sequential input patterns, possibly due to signal degradation and auditory deprivation in the first
415 years of life (Conway et al., 2009; Deocampo et al., 2018). It stands also in contrast with a
416 reduced use of contextual information (Conway et al., 2014; Eisenberg et al., 2002;
417 Stelmachowicz et al., 2000) and slow language processing (Burkholder & Pisoni, 2003; Pisoni
418 et al., 2011; Holt et al., 2016) among children with hearing loss. However, recently Pesnot
419 Lerousseau and colleagues developed a modified serial reaction time task where cochlear
420 implanted children and normal hearing children had to react to auditory sequences that embed
421 multiple statistical regularities. No differences between cochlear implanted children and their
422 normal hearing peers were visible at the group level (Pesnot Lerousseau et al., 2022). These
423 results are more in line with the present findings as they indicate that auditory statistical learning
424 is preserved in congenitally deaf children with cochlear implants, at least at a group level.
425 Importantly, we extend these findings to a rather complex linguistic context, here watching the
426 movies of two people having a conversation wherein multiple levels of predictions are required
427 to efficiently anticipate turns. Assessing predictions and statistical learning in a complex
428 linguistic context may partly explain the discrepancy with previous findings describing impaired
429 statistical learning in children with CI.

430 The fact that children with CI spent more time looking at the current speaker points to a larger
431 reliance on the visual modality. This probably allows compensating the poorer speech
432 integration and comprehension that may otherwise reduce their anticipatory behaviour
433 (Bergeson and Pisoni, 2004; Peterson et al., 2010; Holt et al., 2023), in particular in the current
434 context wherein visual anticipatory cues were removed from the videos. Recently we showed
435 that, using the same setup in normal hearing adults, speech degradation increases the time
436 spent on the current speaker and reduces the anticipatory gaze behaviour (Hidalgo et al.,
437 2022). Here, we only observed a greater time spent on the current speaker in children with CI
438 compared to children with NH, but no differences at the level of anticipatory gaze behaviour. It is
439 possible that the extra time children with CI spent on the current speaker compensated for their
440 poorer audiovisual integration and understanding of the dialogues. Indeed, improving

441 audiovisual integration may improve syntactic and semantic processing that play an important
442 role to anticipate turns (Lewis et al., 2017; Blomquist et al., 2021; Holt et al., 2023). Moreover,
443 children may also rely on supra-segmental features such as coarse prosodic patterns (e.g.
444 questions and answers) that may be partly preserved (Carter et al., 2002), although they may
445 depend on other variables such as musical environment (Torppa et al., 2014).

446 While clear group differences are absent in gaze predictive behaviour, we did find that temporal
447 predictions during conversation correlate with both speech integration abilities (non-word
448 repetition) and dialogue comprehension. More precisely, the more the anticipatory gaze
449 behaviour, the better speech integration and dialogue comprehension in children with CI.
450 Interestingly, in the normal hearing adult experiment, the relation between gaze anticipatory
451 behaviour and dialogue comprehension was maximal in the strongly speech degraded condition
452 and almost absent in the normal speech condition (Hidalgo et al., 2022). These results under
453 the strongly degraded and normal conditions resemble to the results we obtained with children
454 with CI and with NH, respectively. Indeed, previous work showed that adverse listening
455 conditions increase the uncertainty of the speech signal, thus urging listeners to exploit prior
456 knowledge or expectations to constrain perception (Smiljanic & Sladen, 2013; Sohoglu & Davis,
457 2020). This possibly explains the lack of correlation between gaze anticipatory behaviour and
458 dialogue comprehension in children with normal hearing who did not experience degraded
459 speech in the present experiment.

460 Our study has a certain number of limitations. First, while our design goes indeed in the
461 direction of ecological studies, it cannot yet be considered as a real-life listening situation and
462 further efforts will be required in order to be able to accurately track speech perception and
463 anticipatory behaviour in natural conditions (Lunner et al., 2020). In particular, testing in a quiet
464 booth is not representative of real-life conditions in terms of background noise, which may
465 explain the rather good performance in terms of comprehension and in particular in anticipatory
466 gaze behaviour. Second, while we do reproduce previous findings showing impaired speech
467 perception (Caldwell & Nittrouer, 2013), comprehension (Niparko et al., 2010; Caselli et al.,
468 2012), working memory (Nittrouer et al., 2013; AuBuchon et al., 2015) and temporal processing
469 (Good et al., 2017; Polonenko et al., 2017; Hidalgo et al., 2021), our paradigm does not allow to
470 understand how children with CI manage to have a “normal” anticipatory gaze behaviour for turn
471 taking. As stated above they may more heavily rely upon audiovisual integration, coarse
472 prosodic patterns or lexico-semantic or syntactic structures. Further studies may manipulate for
473 instance the prosodic content of conversational speech, such as the continuation contour
474 (Portes and Bertrand, 2005), to assess the respective contribution of lexical-semantic and
475 prosodic processing in anticipating turns. Nonetheless, considering that in our stimuli we
476 purposely removed most visual cues that typically play a role in turn taking (Auer, 2022), the fact
477 that children with CI have a similar anticipatory gaze behaviour compared to children seems to

478 point to the fact that they can build “good enough” predictions of turns. This is relevant when
479 trying to understand the difficulties they encounter in conversational context (Crowe &
480 Dammeyer, 2021). On the other side, it is also true that one may have expected a greater
481 usage of anticipatory gaze behavior to compensate the speech signal degradation of the
482 implant. Further research may show that this strategy may be present in a population of greater
483 age.

484

485 To conclude, our results confirm and extend to a complex linguistic context previous findings
486 showing a lack of evidence for differences in predictive processing in children with CI (Hall et al.,
487 2018; von Koss Torkildsen et al., 2018; Holt et al., 2021, Blomquist et al., 2021, Pesnot
488 Lerousseau et al., 2022). Moreover, the present design, possibly using less controlled stimuli,
489 seems an interesting avenue to provide an accurate and objective estimates of anticipatory
490 behaviour in a more ecological and dynamic linguistic context (Keidser et al., 2020). Also note
491 that this type of design has been previously used with very young children, showing that they
492 spontaneously made anticipatory gaze switches by age two and continue improving through
493 age six (Casillas & Frank, 2017). Thus, it may be possible to use it to refine diagnosis at a rather
494 early age and soon after cochlear implantation.

495 Finally the link between anticipatory gaze behaviour and perceptual language skills and speech
496 comprehension is relevant in the context of language rehabilitation: establishing a link between
497 temporal prediction abilities and level of comprehension in an complex linguistic setup suggests
498 that rhythmic music stimulation may benefit language interaction by improving temporal
499 predictions (Hidalgo et al., 2017, 2019).

500

501

502 References

503 Arjmandi, M. K., Jahn, K. N., & Arenberg, J. G. (2022). Single-channel focused thresholds relate
504 to vowel identification in pediatric and adult cochlear implant listeners. *Trends in Hearing*, 26,
505 23312165221095364.

506 AuBuchon, A. M., Pisoni, D. B., & Kronenberger, W. G. (2015). Short-term and working memory
507 impairments in early-implanted, long-term cochlear implant users are independent of audibility
508 and speech production. *Ear and hearing*, 36(6), 733.

509 Auer, P. (2021). Turn-allocation and gaze: A multimodal revision of the “current-speaker-
510 selects-next” rule of the turn-taking system of conversation analysis. *Discourse Studies*, 23(2),
511 117-140.

512 Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H., Singmann, H., & Dai, B.
513 (2015). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1–7. 2014.

- 514 Bergeson, T. R., and Pisoni, D.B. (2004). "Audiovisual speech perception in deaf adults and
515 children following cochlear implantation," in *The Handbook of Multisensory Processes*, edited by
516 G. A. Calvert, C. Spence, and Stein B. E. (The MIT Press, Cambridge, MA:), pp. 153–176.
- 517 Blomquist, C., Newman, R. S., Huang, Y. T., & Edwards, J. (2021). Children with cochlear
518 implants use semantic prediction to facilitate spoken word recognition. *Journal of Speech,
519 Language, and Hearing Research*, 64(5), 1636-1649.
- 520 Boons, T., De Raeve, L., Langereis, M., Peeraer, L., Wouters, J., & Van Wieringen, A. (2013).
521 Expressive vocabulary, morphology, syntax and narrative skills in profoundly deaf children after
522 early cochlear implantation. *Research in Developmental Disabilities*, 34(6), 2008-2022.
- 523 Burkholder, R. A., & Pisoni, D. B. (2003). Speech timing and working memory in profoundly deaf
524 children after cochlear implantation. *Journal of experimental child psychology*, 85(1), 63-88.
- 525 Busquet, D., & Descourtieux, C. (2003). T.E.R.M.O. Tests d'Evaluation de la Réception du
526 Message Oral par l'enfant sourd à destination des professionnels de la surdité. Paris : Ortho
527 Edition.
- 528 Caldwell A. , & Nittrouer S. (2013). Speech perception in noise by children with cochlear
529 implants . *Journal of Speech, Language, and Hearing Research*, 56 , 13 – 30.
- 530 Carlile, S., & Keidser, G. (2020). Conversational interaction is the brain in action: Implications
531 for the evaluation of hearing and hearing interventions. *Ear and hearing*, 41, 56S-67S.
- 532 Carter, A. K., Dillon, C. M., & Pisoni, D. B. (2002). Imitation of nonwords by hearing impaired
533 children with cochlear implants: Suprasegmental analyses. *Clinical linguistics & phonetics*,
534 16(8), 619-638.
- 535 Caselli M. C. Rinaldi P. Varuzza C. Giuliani A. , & Burdo S. (2012). Cochlear implant in the
536 second year of life: Lexical and grammatical outcomes . *Journal of Speech, Language, and
537 Hearing Research*, 55 , 382 – 394.
- 538 Casillas, M., & Frank, M. C. (2017). The development of children's ability to track and predict
539 turn structure in conversation. *Journal of memory and language*, 92, 234-253.
- 540 Chambers, C. G., & San Juan, V. (2008). Perception and presupposition in real-time language
541 comprehension: Insights from anticipatory processing. *Cognition*, 108(1), 26-50.
- 542 Chaytor, N., Schmitter-Edgecombe, M., & Burr, R. (2006). Improving the ecological validity of
543 executive functioning assessment. *Archives of clinical neuropsychology*, 21(3), 217-227.
- 544 Chin, S. B., Bergeson, T. R., & Phan, J. (2012). Speech intelligibility and prosody production in
545 children with cochlear implants. *Journal of communication disorders*, 45(5), 355-366.
- 546 Conway, C. M., Deocampo, J. A., Walk, A. M., Anaya, E. M., & Pisoni, D. B. (2014). Deaf
547 children with cochlear implants do not appear to use sentence context to help recognize spoken
548 words. *Journal of Speech, Language, and Hearing Research*, 57(6), 2174-2190.
- 549 Conway, C. M., Pisoni, D. B., & Kronenberger, W. G. (2009). The importance of sound for
550 cognitive sequencing abilities: The auditory scaffolding hypothesis. *Current directions in
551 psychological science*, 18(5), 275-279.
- 552 Crowe, K., & Dammeyer, J. (2021). A review of the conversational pragmatic skills of children
553 with cochlear implants. *The Journal of Deaf Studies and Deaf Education*, 26(2), 171-186.
- 554 Deocampo, J. A., Smith, G. N., Kronenberger, W. G., Pisoni, D. B., & Conway, C. M. (2018).
555 The role of statistical learning in understanding and treating spoken language outcomes in deaf

556 children with cochlear implants. *Language, Speech, and Hearing Services in Schools*, 49(3S),
557 723-739.

558 Dettman, S. J., Dowell, R. C., Choo, D., Arnott, W., Abrahams, Y., Davis, A., ... & Briggs, R. J.
559 (2016). Long-term communication outcomes for children receiving cochlear implants younger
560 than 12 months: A multicenter study. *Otology & Neurotology*, 37(2), e82-e95.

561 Dilley, L., Lehet, M., Wieland, E. A., Arjmandi, M. K., Kondaurova, M., Wang, Y., ... & Bergeson,
562 T. (2020). Individual differences in mothers' spontaneous infant-directed speech predict
563 language attainment in children with cochlear implants. *Journal of Speech, Language, and*
564 *Hearing Research*, 63(7), 2453-2467.

565 Eisenberg, L. S., Martinez, A. S., Holowecky, S. R., & Pogorelsky, S. (2002). Recognition of
566 lexically controlled words and sentences by children with normal hearing and children with
567 cochlear implants. *Ear and Hearing*, 23(5), 450-462.

568 Foulsham, T., Cheng, J. T., Tracy, J. L., Henrich, J., & Kingstone, A. (2010). Gaze allocation in
569 a dynamic situation: Effects of social status and speaking. *Cognition*, 117(3), 319-331.

570 Friston, K. J., Parr, T., Yufik, Y., Sajid, N., Price, C. J., & Holmes, E. (2020). Generative models,
571 linguistic communication and active inference. *Neuroscience & Biobehavioral Reviews*, 118, 42-
572 64.

573 Good, A., Gordon, K. A., Papsin, B. C., Nespoli, G., Hopyan, T., Peretz, I., & Russo, F. A.
574 (2017). Benefits of music training for perception of emotional speech prosody in deaf children
575 with cochlear implants. *Ear and Hearing*, 38(4), 455.

576 Grisoni, L., Tomasello, R., & Pulvermüller, F. (2021). Correlated brain indexes of semantic
577 prediction and prediction error: brain localization and category specificity. *Cerebral Cortex*,
578 31(3), 1553-1568.

579 Habib, M. G., Waltzman, S. B., Tajudeen, B., & Svirsky, M. A. (2010). Speech production
580 intelligibility of early implanted pediatric cochlear implant users. *International journal of pediatric*
581 *otorhinolaryngology*, 74(8), 855-859.

582 Hall, M. L., Eigsti, I. M., Bortfeld, H., & Lillo- Martin, D. (2018). Auditory access, language
583 access, and implicit sequence learning in deaf children. *Developmental Science*, 21(3), e12575.

584 Hayes, H., Geers, A. E., Treiman, R., & Moog, J. S. (2009). Receptive vocabulary development
585 in deaf children with cochlear implants: Achievement in an intensive auditory-oral educational
586 setting. *Ear and hearing*, 30(1), 128-135.

587 Hidalgo, C., Falk, S., & Schön, D. (2017). Speak on time! Effects of a musical rhythmic training
588 on children with hearing loss. *Hearing research*, 351, 11-18.

589 Hidalgo, C., Mohamed, I., Zielinski, C., & Schön, D. (2022). The effect of speech degradation on
590 the ability to track and predict turn structure in conversation. *Cortex*, 151, 105-115.

591 Hidalgo, C., Pesnot-Lerousseau, J., Marquis, P., Roman, S., & Schön, D. (2019). Rhythmic
592 training improves temporal anticipation and adaptation abilities in children with hearing loss
593 during verbal interaction. *Journal of Speech, Language, and Hearing Research*, 62(9), 3234-
594 3247.

595 Hidalgo, C., Zécri, A., Pesnot-Lerousseau, J., Truy, E., Roman, S., Falk, S., ... & Schön, D.
596 (2021). Rhythmic abilities of children with hearing loss. *Ear and Hearing*, 42(2), 364-372.

597 Holt, C. M., Demuth, K., & Yuen, I. (2016). The use of prosodic cues in sentence processing by
598 prelingually deaf users of cochlear implants. *Ear and Hearing*, 37(4), e256-e262.

599 Holt, R., Bruggeman, L., & Demuth, K. (2021). Children with hearing loss can predict during
600 sentence processing. *Cognition*, 212, 104684

601 Holt, R., Bruggeman, L., & Demuth, K. (2023). Effects of hearing loss and audio-visual cues on
602 children's speech processing speed. *Speech Communication*, 146, 11-21.

603 Houston, D. M. (2022). A framework for understanding the relation between spoken language
604 input and outcomes for children with cochlear implants. *Child Development Perspectives*, 16(1),
605 60-66.

606 Hutton, S. B. (2008). Cognitive control of saccadic eye movements. *Brain and cognition*, 68(3),
607 327-340.

608 Keidser, G., Naylor, G., Brungart, D. S., Caduff, A., Campos, J., Carlile, S., ... & Smeds, K.
609 (2020). The quest for ecological validity in hearing science: What it is, why it matters, and how
610 to advance it. *Ear and hearing*, 41(Suppl 1), 5S.

611 Keitel, A., Prinz, W., Friederici, A. D., Von Hofsten, C., & Daum, M. M. (2013). Perception of
612 conversations: The importance of semantics and intonation in children's development. *Journal*
613 *of Experimental Child Psychology*, 116(2), 264-277.

614 Kondaurova, M. V., Smith, N. A., Zheng, Q., Reed, J., & Fagan, M. K. (2020). Vocal turn-taking
615 between mothers and their children with cochlear implants. *Ear and hearing*, 41(2), 362.

616 Kondaurova, M. V., Zheng, Q., VanDam, M., & Kinney, K. (2022). Vocal turn-taking in families
617 with children with and without hearing loss. *Ear and hearing*, 43(3), 883.

618 Krause, F. & Lindemann, O. (2014). Expyriment: A Python library for cognitive and
619 neuroscientific experiments. *Behavior Research Methods*, 46(2), 416-428.

620 Kuznetsova A, Brockhoff PB, Christensen RHB (2017). "lmerTest Package: Tests in Linear
621 Mixed Effects Models." *Journal of Statistical Software*, 82(13), 1–26.

622 Pesnot Lerousseau, J, Hidalgo, C., Roman, S., & Schön, D. (2022). Does auditory deprivation
623 impairs statistical learning in the auditory modality?. *Cognition*, 222, 105009.

624 Levinson, S. C. (2016). Turn-taking in human communication—origins and implications for
625 language processing. *Trends in cognitive sciences*, 20(1), 6-14.

626 Lewis, D. E., Kopun, J., McCreery, R., Brennan, M., Nishi, K., Cordrey, E., ... & Moeller, M. P.
627 (2017). Effect of context and hearing loss on time-gated word recognition in children. *Ear and*
628 *hearing*, 38(3), e180.

629 Lovcevic, I., Burnham, D., & Kalashnikova, M. (2022). Language development in infants with
630 hearing loss: Benefits of infant-directed speech. *Infant Behavior and Development*, 67, 101699.

631 Lund, E. (2016). Vocabulary knowledge of children with cochlear implants: A meta-analysis.
632 *Journal of deaf studies and deaf education*, 21(2), 107-121.

633 Lunner, T., Alickovic, E., Graverson, C., Ng, E. H. N., Wendt, D., & Keidser, G. (2020). Three
634 new outcome measures that tap into cognitive processes required for real-life communication.
635 *Ear and hearing*, 41(Suppl 1), 39S.

636 Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical
637 experiment builder for the social sciences. *Behavior research methods*, 44, 314-324.

638 Nicholas, J. G., & Geers, A. E. (2018). Sensitivity of expressive linguistic domains to surgery
639 age and audibility of speech in preschoolers with cochlear implants. *Cochlear implants*
640 *international*, 19(1), 26-37.

641 Niparko, J. K., Tobey, E. A., Thal, D. J., Eisenberg, L. S., Wang, N. Y., Quittner, A. L., ... &
642 CDaCI Investigative Team. (2010). Spoken language development in children following
643 cochlear implantation. *Jama*, 303(15), 1498-1506.

644 Nittrouer, S., Caldwell-Tarr, A., & Lowenstein, J. H. (2013). Working memory in children with
645 cochlear implants: Problems are in storage, not processing. *International Journal of Pediatric*
646 *Otorhinolaryngology*, 77(11), 1886-1898.

647 Patel, A. D., & Morgan, E. (2017). Exploring cognitive relations between prediction in language
648 and music. *Cognitive Science*, 41, 303-320.

649 Peterson, N. R., Pisoni, D. B., & Miyamoto, R. T. (2010). Cochlear implants and spoken
650 language processing abilities: Review and assessment of the literature. *Restorative neurology*
651 *and neuroscience*, 28(2), 237-250.

652 Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and
653 review. *Psychological bulletin*, 144(10), 1002.

654 Pisoni, D., Kronenberger, W., Roman, A., & Geers, A. (2011). Article 7: Measures of digit span
655 and verbal rehearsal speed in deaf children following more than 10 years of cochlear
656 implantation. *Ear and hearing*, 32(1), 60s.

657 Polonenko, M. J., Giannantonio, S., Papsin, B. C., Marsella, P., & Gordon, K. A. (2017). Music
658 perception improves in children with bilateral cochlear implants or bimodal devices. *The Journal*
659 *of the Acoustical Society of America*, 141(6), 4494-4507.

660 Portes, C., & Bertrand, R. (2005). Some cues about the interactional value of the «continuation»
661 contour in French. In *Actes, Discours et Prosodie comme Interface Complexe (IDP)* (pp. 1-14).
662 Proceedings of Discourse-Prosody Interface Symposium-IDP05.. Aix-en-Provence: Seridisc,
663 Belgique. 2006. Cederom.

664 R Core Team. (2020). R: A Language and Environment for Statistical Computing; R Foundation
665 for Statistical Computing: Vienna, Austria, 2021

666 Ruben, R. J. (2018). Language development in the pediatric cochlear implant patient.
667 *Laryngoscope investigative otolaryngology*, 3(3), 209-213.

668 Schön, D., & Morillon, B. (2018). The Oxford handbook of music and the brain.

669 Sohoglu, E., & Davis, M. H. (2020). Rapid computations of spectrotemporal prediction error
670 support perception of degraded speech. *Elife*, 9, e58077.

671 Sowiński, J., & Dalla Bella, S. (2013). Poor synchronization to the beat may result from deficient
672 auditory-motor mapping. *Neuropsychologia*, 51(10), 1952-1963.

673 Stelmachowicz, P. G., Hoover, B. M., Lewis, D. E., Kortekaas, R. W., & Pittman, A. L. (2000).
674 The relation between stimulus context, speech audibility, and perception for normal-hearing and
675 hearing-impaired children. *Journal of Speech, Language, and Hearing Research*, 43(4), 902-
676 914.

677 Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., ... & Levinson, S.
678 C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the*
679 *National Academy of Sciences*, 106(26), 10587-10592.

680 Strauß, A., Kotz, S. A., & Obleser, J. (2013). Narrowed expectancies under degraded speech:
681 revisiting the N400. *Journal of Cognitive Neuroscience*, 25(8), 1383-1395.

682 Svirsky, M. A., Robbins, A. M., Kirk, K. I., Pisoni, D. B., & Miyamoto, R. T. (2000). Language
683 development in profoundly deaf children with cochlear implants. *Psychological Science*, 11(2),
684 153-158.

685 Torppa, R., Faulkner, A., Huutilainen, M., Järvikivi, J., Lipsanen, J., Laasonen, M., & Vainio, M.
686 (2014). The perception of prosody and associated auditory cues in early-implanted children:
687 The role of auditory working memory and musical activities. *International Journal of Audiology*,
688 53(3), 182-191.

689 Van Wieringen, A., & Wouters, J. (2015). What can we expect of normally-developing children
690 implanted at a young age with respect to their auditory, linguistic and cognitive skills?. *Hearing*
691 *Research*, 322, 171-179.

692 von Koss Torkildsen, J., Arciuli, J., Haukedal, C. L., & Wie, O. B. (2018). Does a lack of auditory
693 experience affect sequential learning?. *Cognition*, 170, 123-129.

694 Xu, J., Kemeny, S., Park, G., Frattali, C., & Braun, A. (2005). Language in context: emergent
695 features of word, sentence, and narrative comprehension. *Neuroimage*, 25(3), 1002-1015.

696 Yoshinaga-Itano, C., Baca, R. L., & Sedey, A. L. (2010). Describing the trajectory of language
697 development in the presence of severe to profound hearing loss: A closer look at children with
698 cochlear implants versus hearing aids. *Otology & neurotology: official publication of the*
699 *American Otological Society, American Neurotology Society and European Academy of Otology*
700 *and Neurotology*, 31(8), 1268.

701 Walker, R., Walker, D. G., Husain, M., & Kennard, C. (2000). Control of voluntary and reflexive
702 saccades. *Experimental Brain Research*, 130, 540-544.

703 Wechsler, D. (2005). Wechsler intelligence scale for children (WISC IV)—translated. *ECPA*,
704 *Paris*.

705

706

707

Child	Age in years	Sex	Duration of use of the CI	Ear-implanted	Age at switch on first CI in months	Age at switch on second CI in months	Age of first hearing aid in months	Onset of deafness	Current hearing loss without CI(s)
2	11	F	9	bilateral	19	51	6	unknown	profound
4	11	M	10	left	17	NA	4	unknown	profound
5	9	M	8	bilateral	12	31	8	congenital	profound
6	9	F	8	bilateral	12	35	5	congenital	profound
11	11	F	9	right	17	NA	unknown	congenital	profound
12	9	M	8	bilateral	12	23	4	unknown	profound
13	10	M	7	bilateral	38	38	5	unknown	profound
15	7	F	6	bilateral	11	19	2	congenital	profound
18	12	M	10	bilateral	25	48	unknown	congenital	profound
21	8	M	7	bilateral	15	52	2	unknown	profound
22	9	F	7	bilateral	26	26	unknown	unknown	profound
24	9	F	7	bilateral	21	21	unknown	congenital	profound