ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

Prosodic and Visual Cues Facilitate Irony Comprehension by Mandarin-Speaking Children With Cochlear Implants

Qianxi Yu,[a] Honglan Li,[a] Shanpeng Li,[a] and Ping Tang[a]

[a] School of Foreign Studies, Nanjing University of Science and Technology, China

ARTICLE INFO

Article History:
Received November 10, 2023
Revision received January 21, 2024
Accepted March 13, 2024

Editor-in-Chief: Julie A. Washington
Editor: James R. Booth

https://doi.org/10.1044/2024_JSLHR-23-00701

ABSTRACT

Purpose: This study investigated irony comprehension by Mandarin-speaking children with cochlear implants, focusing on how prosodic and visual cues contribute to their comprehension, and whether second-order Theory of Mind is required for using these cues.
Method: We tested 52 Mandarin-speaking children with cochlear implants (aged 3–7 years) and 52 age- and gender-matched children with normal hearing. All children completed a Theory of Mind test and a story comprehension test. Ironic stories were presented in three conditions, each providing different cues: (a) context-only, (b) context and prosody, and (c) context, prosody, and visual cues. Comparisons were conducted on the accuracy of story understanding across the three conditions to examine the role of prosodic and visual cues.
Results: The results showed that, compared to the context-only condition, the additional prosodic and visual cues both improved the accuracy of irony comprehension for children with cochlear implants, similar to their normal-hearing peers. Furthermore, such improvements were observed for all children, regardless of whether they passed the second-order Theory of Mind test or not.
Conclusions: This study is the first to demonstrate the benefits of prosodic and visual cues in irony comprehension, without reliance on second-order Theory of Mind, for Mandarin-speaking children with cochlear implants. It implies potential insights for utilizing prosodic and visual cues in intervention strategies to promote irony comprehension.

Cochlear implants (CIs) emerge as an effective tool for children with severe to profound hearing loss to regain access to hearing and engage in spoken communication (Hopyan et al., 2011; O'Donoghue et al., 2000). While numerous studies have reported the benefits of CIs in enhancing auditory perception and basic linguistic skills by hard-of-hearing children (Duchesne & Marschark, 2019; Schramm et al., 2010), research into their complex social-pragmatic abilities remains limited (Parola et al., 2023; Tuohimaa et al., 2023). Despite this, emerging studies have reported that children with CIs show greater difficulties in social-pragmatic performance, especially in understanding verbal irony, which requires inferring the intended meaning from the discrepancy between discourse context and literal meaning (Agostinelli et al., 2021; Panzeri et al., 2021). Understanding irony, compared to other forms of figurative language that also depend on context such as metaphor, is particularly challenging as it requires a recursive mental reasoning ability, that is, second-order Theory of Mind (ToM; Colston & Gibbs, 2002). Nevertheless, in daily speech, ironic expressions are also accompanied by unique prosody and facial expressions, which have been found to facilitate normal-hearing (NH) children's irony comprehension, even for those who face challenges in this regard when using the contextual cue (Capelli et al., 1990; González-Fuente, 2017; Laval & Bert-Erboul, 2005). The present study thus explored whether children with CIs could

Correspondence to Ping Tang: ping.tang@njust.edu.cn. *Disclosure:* *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

similarly benefit from prosodic and visual cues to understand irony and whether their ability to use these cues requires the development of second-order ToM.

Irony is a frequently encountered figure of speech in social-pragmatic communications (Gibbs & Colston, 2012). Ironic criticism, which often uses literally positive content to convey negative criticism, is the most common type of irony in daily communications (Capelli et al., 1990; Filippova & Astington, 2008, 2010; Gibbs, 2000; Hancock et al., 2000). For instance, A says, “You are such a great chef!” to B who has just made an awful meal. In this example, listeners need to infer the intended meaning of A (i.e., B is a bad chef) not only from the uttered sentence but more importantly from its incongruity with the negative context (i.e., B has just made an awful meal; Deliens et al., 2018; Gerrig & Goldvarg, 2000; Pexman & Olineck, 2002). Over the past decades, cross-language studies have explored irony understanding by NH children, primarily focusing on their ability to use narrative contexts to infer the intended meaning of irony (e.g., English: Hancock et al., 2000; Finnish: Loukusa & Leinonen, 2008; Italian: Panzeri et al., 2020; Mandarin: Zhang & Zhang, 2006). Research consistently documented that understanding irony is more challenging for both NH children and children with CIs than interpreting literal criticisms and literal praises, where the context is consistent with the literal meaning (Hancock et al., 2000; Panzeri et al., 2021; Pexman et al., 2011).

The difficulty in understanding irony has been attributed to the complexity involved in inferring the intended meaning, which requires the ability of ToM (Colston & Gibbs, 2002; Filippova & Astington, 2008; Happé, 1993; Winner & Leekam, 1991). According to the complexity of mental reasoning, ToM can be further classified into first-order ToM, which involves the awareness of different beliefs of another person (Wimmer & Perner, 1983), and second-order ToM, which involves recursive inferences where one grasps others’ thoughts about beliefs (Perner & Wimmer, 1985). While first-order ToM might be sufficient for children to understand other forms of figurative language such as metaphor, irony comprehension based on contextual information demands second-order ToM (Colston & Gibbs, 2002; Happé, 1993). Specifically, to infer the speaker’s ironic intent, children must recognize not only the contextual incongruity, which implies the speaker’s real belief, but also that the speaker believes the target listener is also aware of this incongruity. Such recursive belief evaluation is similar to the cognitive inference involved in second-order ToM, which is often framed as, “A thinks that B thinks that . . . ” (Baron-Cohen, 1989; Perner & Wimmer, 1985).

NH children typically develop ToM during their preschool years, particularly between the ages of 3 and 7 years (see Wellman, 2011, for a review), while children with CIs usually score lower in ToM tasks as compared to their age-matched NH peers (Liu et al., 2018; Panzeri et al., 2021; Peterson, 2004; Pluta et al., 2021). This is because these children often exhibit atypical patterns in two contributing factors to ToM development: generally language development and mind-related conversational input (Peterson, 2020). First, children with CIs often experience more gradual language development compared to their NH peers, stemming from limited exposure to spoken language caused by pre-implantation auditory deprivation and the extra time needed postimplantation to adapt to their devices and acquire various linguistic skills (Bavin et al., 2018; Connor et al., 2006; Duchesne et al., 2009; Houston et al., 2020). Second, these children have been reported to be exposed less to mental state–related conversation compared to their NH peers, as hearing parents often use less complex mind-related language with their child with CIs during early childhood, considering their paced linguistic progress (Moeller & Shick, 2006; Peterson, 2020; Wellman, 2014).

Given the gradual developmental trajectory of ToM in children with CIs, they may face great challenges in decoding the intended meaning of irony expressions based on the context information. For example, Panzeri et al. (2021) explored the comprehension of ironic stories of 4- to 12-year-old children with CIs based on narrated context. Their results showed that while these children understood literal criticism and literal praise with over 95% accuracy, their accuracy of understanding ironic stories was only 53% (slightly above the chance level), considerably lower than the 86% accuracy of their age-matched NH peers. Notably, while these children understood what happened in the context, they still tended to stick to the literal interpretation of irony, indicating a gap in their ability to rely on the contextual cue to infer the intended meaning of irony.

Beyond the difficulties in context-based inference, irony understanding in daily speech scenarios may be particularly challenging for children with CIs as it also relies on prosody perception, an aspect that CI users may not effectively grasp (Chatterjee & Peng, 2008; Cooper et al., 2008; Fuller et al., 2014; Gaudrain & Başkent, 2018; Nakata et al., 2012; Peng et al., 2008). In daily speech, ironic utterances are often expressed through unique prosodic patterns, including a slower speech rate and distinct fundamental frequency ($F0$) patterns (Anolli et al., 2002; Cheang & Pell, 2008, 2009; Gu et al., 2011; S. Li & Gu, 2021; S. Li et al., 2020). From a perception perspective, NH children benefit from these prosodic cues in their comprehension of irony (Capelli et al., 1990; Laval & Bert-Erboul, 2005; J. P. Li et al., 2013; Milosky & Ford, 1997; Zhang & Zhang, 2006). For instance, Mandarin-speaking 6-year-olds’ accuracy of irony understanding improved from 30% when only contextual information was provided to 66% when both contextual and prosodic cues were provided (Zhang & Zhang, 2006). However, children with CIs potentially face challenges in perceiving these cues, because *F*0 information, a major prosodic vehicle of ironic expressions, is not effectively transmitted via CIs, as a consequence of the restricted number of channels available to convey the broad range of speech frequencies (Limb & Roy, 2014; Ren et al., 2022; Vandali & van Hoesel, 2012). Previous research has documented that CI users, including both children and adults, have difficulties in *F*0 perception, such as lexical tone recognition (Peng et al., 2004; W. Wang et al., 2011), emotional speech perception (Good et al., 2017; Hopyan-Misakyan et al., 2009; H. Wang et al., 2019), and music melody appreciation (e.g., Cooper et al., 2008). This suggests that children with CIs might only have limited access to ironic prosody, that is, primarily the durational property rather than the *F*0 property. Therefore, these children may not effectively benefit from prosodic cues in irony comprehension like their NH peers.

Nevertheless, irony productions in daily communications are also accompanied by specific facial expressions, including rolling eyes, expressive lip movement, and rapid eye flashes (Aguert, 2022; Attardo et al., 2003; S. Li et al., 2022; Rockwell, 2001). These cues could potentially benefit children with CIs in irony comprehension. Notably, research in NH children has suggested that the availability of visual cues could effectively enhance their irony understanding, indicating an audio-visual benefit in irony comprehension. For example, González-Fuente (2017) reported that Catalan-speaking NH 5-year-olds were not able to decode the target meaning of irony based on contextual information alone (around 0% correct in the context-only condition), while their accuracy improved to 38% when both prosodic and visual cues were provided. Notably, children with hearing loss showed a strong visual speech perception ability contributed by the cross-modal plasticity after early auditory deprivation (see Kral & Sharma, 2023, for a review). Many studies have demonstrated such visual benefits in various aspects, including lexical tone perception (e.g., Tang et al., 2023), emotion perception (Hopyan-Misakyan et al., 2009; Most & Michaelis, 2012), and word and sentence comprehension (Bergeson et al., 2005; Lachs et al., 2001; Taitelbaum-Swead & Fostick, 2017). Therefore, children with CIs may use visual cues to enhance their irony comprehension. However, such a hypothesis remained to be examined with empirical evidence.

Moving beyond the focus on how children with CIs could potentially use prosodic and visual cues for irony comprehension, another question is yet to be addressed: Are these cues exclusively accessible for children who have developed second-order ToM? Exploring this question would deepen our understanding of the prerequisite for using multimodal cues in irony comprehension by children with CIs. In contrast with relying on the contextual cue, which involves complex second-order ToM-supported inference, using prosodic and visual cues may be less demanding in mental reasoning. Evidence from NH children has shown that those who face difficulties in context-based irony understanding could effectively use prosodic and/or visual cues to facilitate their comprehension (Capelli et al., 1990; González-Fuente, 2017; Laval & Bert-Erboul, 2005; Rankin et al., 2009; Zhang & Zhang, 2006). Such facilitation might be related to the fact that young children at an early age (around 3–4 years) were able to use prosodic and visual cues to detect the valence (positive and negative) of emotion (Baltaxe, 1991; Nelson & Russel, 2011; Sauter et al., 2013; Widen, 2013; Widen & Russell, 2008), much earlier than the full development of second-order ToM (around 6–7 years; see Miller, 2009, for a review). Therefore, the prosodic and visual cues in ironic expressions could help young children to better detect the speaker's negative meaning by providing direct perceptual information of negative emotions, rather than complex mental reasoning involved in context-based inference (Capelli et al., 1990; Nakassis & Snedeker, 2002). Therefore, it is possible that children with CIs who may have not fully developed second-order ToM could still use prosodic and/or visual cues for irony comprehension. However, this hypothesis requires empirical evidence to verify.

The current study therefore investigated the role of prosodic and visual cues in the comprehension of ironic speech by Mandarin-speaking children with CIs. We first asked whether prosodic and/or visual cues improve the irony comprehension performance of children with CIs (Research Question 1 [RQ1]). If yes, we further asked whether children who have and have not developed second-order ToM would use these cues to facilitate their irony comprehension in a similar way (Research Question 2 [RQ2]).

Based on previous evidence that children with CIs face challenges in prosody perception (Good et al., 2017; Hopyan-Misakyan et al., 2009; Peng et al., 2004; H. Wang et al., 2019; W. Wang et al., 2011), we hypothesized that relative to conditions where only the contextual cue is provided, the additional prosodic cues would not improve the accuracy of irony comprehension in children with CIs (Hypothesis 1 [H1]). However, we hypothesized that the further inclusion of visual cues, relative to the contextual and prosodic cues, would improve their irony comprehension (Hypothesis 2 [H2]), considering their strong ability in visual speech perception (Hopyan-Misakyan et al., 2009; Most & Michaelis, 2012).

Based on previous studies showing that prosodic and visual cues might be less demanding than the ToM-required contextual cue (Capelli et al., 1990; González-Fuente, 2017;

Laval & Bert-Erboul, 2005), we hypothesized that for children who passed and did not pass the second-order ToM test, they would exhibit similar patterns in using prosodic and/or visual cues to facilitate their irony comprehension (Hypothesis 3 [H3]).

Method

Participants

A total of fifty-two 3- to 7-year-old Mandarin-speaking children with CIs (24 females, M_{age} = 5.09 years, *SD* = 0.99) and 52 NH controls matched by gender and age (age gap < 3 months) were tested. Correlation analysis showed a significant and positive correlation between the chronological ages of children with CIs and NH controls, $r(50) = 0.99$, $p < .001$. The CI group received their implants between the ages of 0.78 and 6.28 years (M_{age} = 2.79 years, *SD* = 1.44) with CI experience ranging from 0.14 to 4.92 years (*M* = 2.32 years, *SD* = 1.30). All children were recruited from rehabilitation centers (CI group) and kindergartens (NH group) in northern China cities, where standard Mandarin Chinese is used in teaching practices and daily communication. According to reports from teachers and official records from kindergartens and rehabilitation centers, none of the children participating in this study had any intellectual issues, psychogenic disorders, or visual abnormalities, and those with CIs only experienced hearing loss prior to their implantation. Permission for conducting this research was granted by the Ethics Committee of Nanjing University of Science and Technology, ensuring full compliance with the informed consent procedures.

Materials

A total of 36 "Context scenario + Target remark" structured short stories were created for three types of attitudes, namely, irony, criticism, and praise (12 stories for each). Note that irony in this study refers to ironic criticism, as it is the most common form of irony in daily communication (Capelli et al., 1990; Dews & Winner, 1997; Hancock et al., 2000), whereas criticism and praise served as control attitudes to ensure that children understand these basic attitudes. Following Panzeri et al. (2021), we used the ceiling accuracy of understanding control attitudes to ensure that the materials used were well within their comprehension capabilities. This approach guarantees that any challenges in understanding irony are not due to the complexity of the materials, but rather reflect the children's intrinsic abilities.

All stories started with a contextual scenario illustrating an interaction between a boy (Xiaoming) and a girl (Xiaohong) in events familiar to young children, for example, a running competition, a football match, and so forth, followed by a target remark (either ironic, criticizing, or praising) from the girl toward the boy. The contexts for irony and criticism were both negative, for example, "Xiaoming lost the running competition," whereas the literal meaning of the target remark from Xiaohong was either positive, "You ran so fast" (irony), or negative, "You ran so slowly" (criticism). On the contrary, the context for praise was always positive, for example, "Xiaoming won the running competition," ending with a positive remark, for example, "You ran so fast."

The contextual scenario of each story was audio-recorded by a female speaker using neutral prosody, whereas the target remark was audio-visually recorded by the same speaker in affective and neutral styles with corresponding prosody and facial expressions (see Figure 1). Acoustic analysis (Appendix A) and perceptual evaluation by 10 adults (Appendix B) were conducted on target remarks, and the results showed that all utterances carried the desired prosody and facial expressions.

Procedure

All children were tested individually in quiet rooms at rehabilitation centers or kindergartens. Each child participated in a ToM test first, followed by the formal story comprehension test.

For the ToM test, classical false-belief tasks (Gopnik & Astington, 1988; Miller, 2013; Perner & Wimmer, 1985;

Figure 1. Prosody and facial expression styles of target remarks.

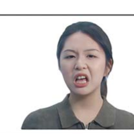
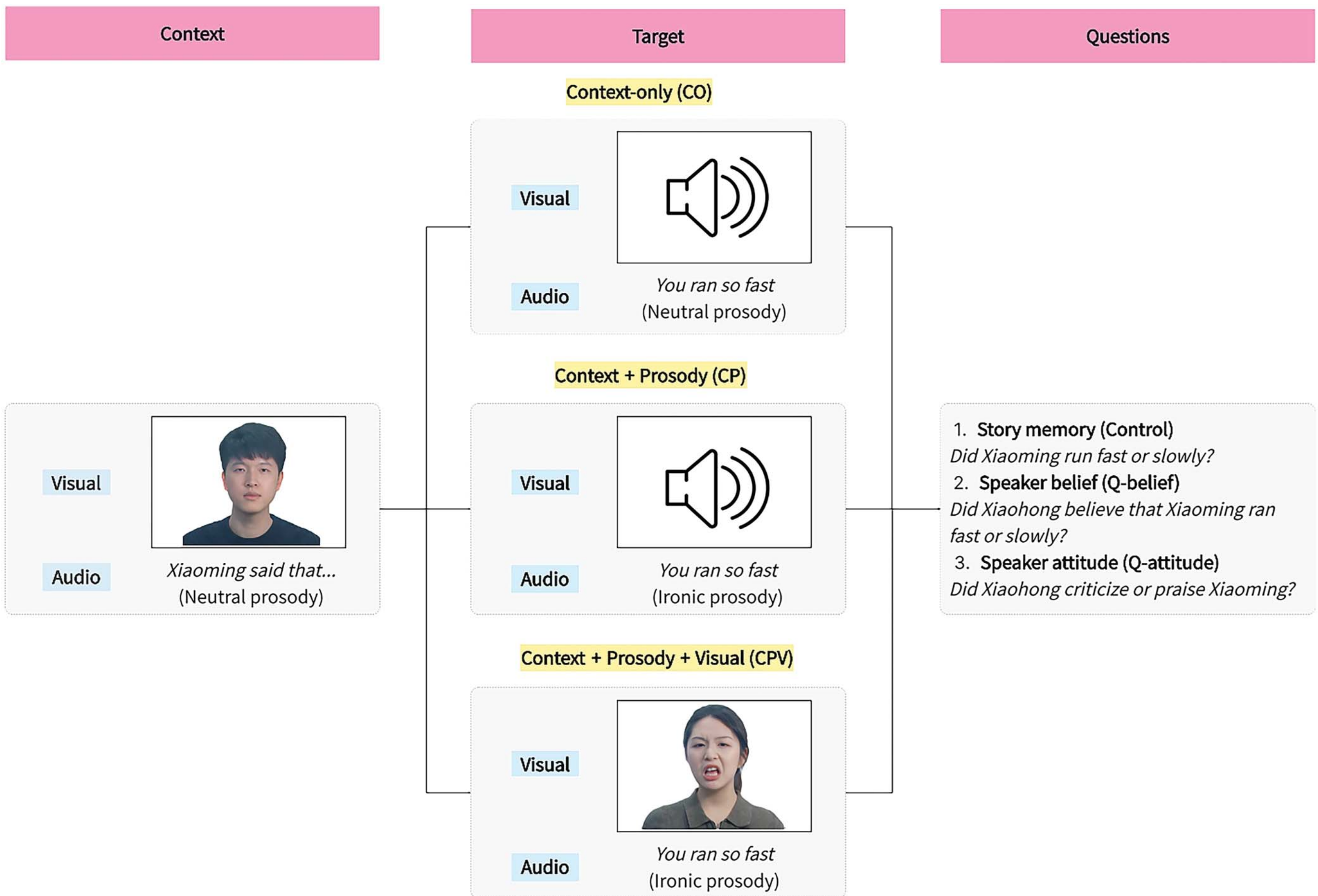
Style	Prosody	Facial expressions
Affective	Ironic	
	Criticizing	
	Praising	
Neutral	Neutral	

Figure 2. Procedure for presenting an irony story across three conditions.



Wimmer & Perner, 1983) were adopted to examine children's first-order and second-order ToM abilities. The first-order ToM test served as a baseline and only those who passed it continued to the second-order ToM test. All stories were translated into Chinese, and prerecorded by a female Mandarin speaker. After each story, control questions testing children's understanding of the story were first asked, followed by test questions examining children's ToM understanding (see Li & Leung, 2020, for ToM test materials).

The formal story comprehension test included 39 trials, equally distributed in three blocked conditions: context-only (CO), context + prosody (CP), and context + prosody + visual (CPV). All conditions started with a familiarization trial, followed by 12 test trials (four trials for each attitude) with a fully randomized order across participants. The procedure of the familiarization and test trials was identical, while only responses in test trials were included in further analysis. In each trial, the audio of the contextual scenario of a story was first played, with an image of the boy character Xiaoming being presented on the screen. Then, the target remark was presented, with different presentation styles across CO, CP, and CPV conditions (see Figure 2 for an example[1]). In the CO condition, for example, only the audio of the target remark was played, with neutral prosody. In the CP condition, again, the audio of the target remark was played, but with affective prosody. In the CPV condition, the target remark was presented as a video, providing the speaker's affective prosody and dynamic facial expressions. Therefore, the CO condition minimally provided prosodic and visual information about the target attitude, while the CP condition provided additional prosodic information, and the CPV condition provided both prosodic and visual information.

Each trial was followed by three questions. Q1 assessed children's understanding of the story content, for example, "Did Xiaoming run fast or slowly?" Q2 (Q-belief)

[1]Due to space limitation within Figure 2, the detailed content of context in the example is presented here: "Xiaoming said that he ran so fast and insisted on competing with Xiaohong. Xiaohong ran to the finish line early and Xiaoming couldn't catch up with Xiaohong no matter how hard he chased. Xiaohong said to Xiaoming."

evaluated children's understanding of the speaker's belief, for example, "Did Xiaohong believe that Xiaoming ran fast or slowly?" Q3 (Q-attitude) evaluated children's understanding of the speaker's attitude, for example, "Did Xiaohong criticize or praise Xiaoming?" Q1 served as a control question and only those who had correctly answered this question proceeded to Q-belief and Q-attitude. The other two questions accessed two different components of irony comprehension: the speaker's real belief which is opposite to the literal content of the remark (Q-belief), and the speaker's attitude, which is negative (Q-attitude). Children's responses to all questions were recorded by the experimenter using the software PsychoPy (Peirce et al., 2019).

Data Coding and Analysis

To test H1 and H2, whether prosodic or visual cues would facilitate children with CIs to comprehend irony better, we compared children's comprehension accuracy of stories across CO, CP, and CPV conditions. Only trials where children correctly understood the content of the story (i.e., Q1) were included in the analysis, leading to 1,777 trials of the NH group (irony: 589 trials; criticism: 598 trials; praise: 590 trials) and 1,739 trials of the CI group (irony: 574 trials; criticism: 587 trials; praise: 578 trials). Each child's responses to Q-belief and Q-attitude in each trial were binarily coded as "Correct" or "Incorrect."

All data analyses were conducted using the R program (R Core Team, 2022). For the two control attitudes, namely, criticism and praise, a series of one-tailed *t* tests were performed on NH and CI groups' accuracy of target questions to examine if they were at the ceiling (95% accuracy) in understanding the two basic attitudes (Helms et al., 2004; Panzeri et al., 2021). For irony, two generalized linear mixed-effect models were performed to compare the accuracy of target questions across conditions and groups, using the *glmer* function from the "lme4" package (Bates et al., 2015). The *mixed* function from the "afex" package was applied to identify statistical significance through likelihood-ratio tests (Singmann et al., 2023). For significant findings, Tukey-HSD post hoc comparisons were conducted to evaluate irony comprehension differences across three conditions between two groups, using the *emmeans* function from the "emmeans" package (Lenth, 2022).

To test H3, whether children who have and have not developed second-order ToM would exhibit similar patterns in using prosodic and/or visual cues in irony comprehension, children were first divided into two subgroups according to their second-order ToM score: "Pass second-ToM" (full score of 2) and "Non-pass second-ToM" (score of 1 or 0). The "Pass second-ToM" group included 21 NH children and eight children with CIs, whereas the "Non-pass second-ToM" group included 21 NH children and 17 children with CIs. An additional 10 NH children and 27 children with CIs were excluded from the analysis because they did not pass the first-order ToM test (NH: eight children; CI: 25 children) or did not correctly answer the control questions of the second-order ToM test (NH: two children; CI: two children). Then, the effect of prosodic and visual cues on irony comprehension within each group was examined (similar to the first analysis).

Results

Analysis 1: Irony Comprehension Across Three Conditions

Figure 3 illustrates the mean accuracy of Q-belief and Q-attitude for irony, criticism, and praise. For the two control attitudes, criticism and praise, one-tailed *t* tests compared children's accuracy to the 95% reference. The results showed that the accuracy of both groups was not significantly below 95% in any condition, suggesting that both groups were at the ceiling in understanding basic criticism and praise (Appendix C).

For irony stories, two generalized linear mixed-effects models were performed on children's accuracy of target questions, with two fixed factors: "Group" (NH and CI) and "Condition" (CO, CP, and CPV), and two random factors: "Participant" and "Item." Both models reported significant main effects of "Group," Q-belief: $\chi^2(1) = 31.43$, $p < .001$; Q-attitude: $\chi^2(1) = 27.28$, $p < .001$; and "Condition," Q-belief: $\chi^2(2) = 195.08$, $p < .001$; Q-attitude: $\chi^2(2) = 224.55$, $p < .001$, while the interaction of "Group × Condition" was not significant. Tukey-HSD post hoc analyses on the main effects of "Group" revealed that both Q-belief and Q-attitude accuracy of the CI groups were significantly lower than that of the NH group (Appendix C). Nevertheless, the post hoc analyses of the "Condition" revealed that both the CI and NH groups showed significant improvements in accuracy from CO to CP condition, and further to CPV condition (Appendix C).

Analysis 2: Irony Comprehension in Relation to Second-Order ToM

Figure 4 depicts the mean accuracy of irony comprehension of children who passed and did not pass the second-order ToM test. Two generalized linear mixed-effect models were performed on the accuracy of target questions in ironic stories, with three fixed factors "Group" (NH and CI), "Condition" (CO, CP, CPV), and "ToM-pass" (Pass second-ToM and Non-pass second-ToM) and two random factors "Participant" and "Item."

Figure 3. Mean accuracy of Q-belief and Q-attitude across attitudes and conditions. NH = normal hearing; CI = cochlear implant; CO = context only; CP = context + prosody; CPV = context + prosody + visual.

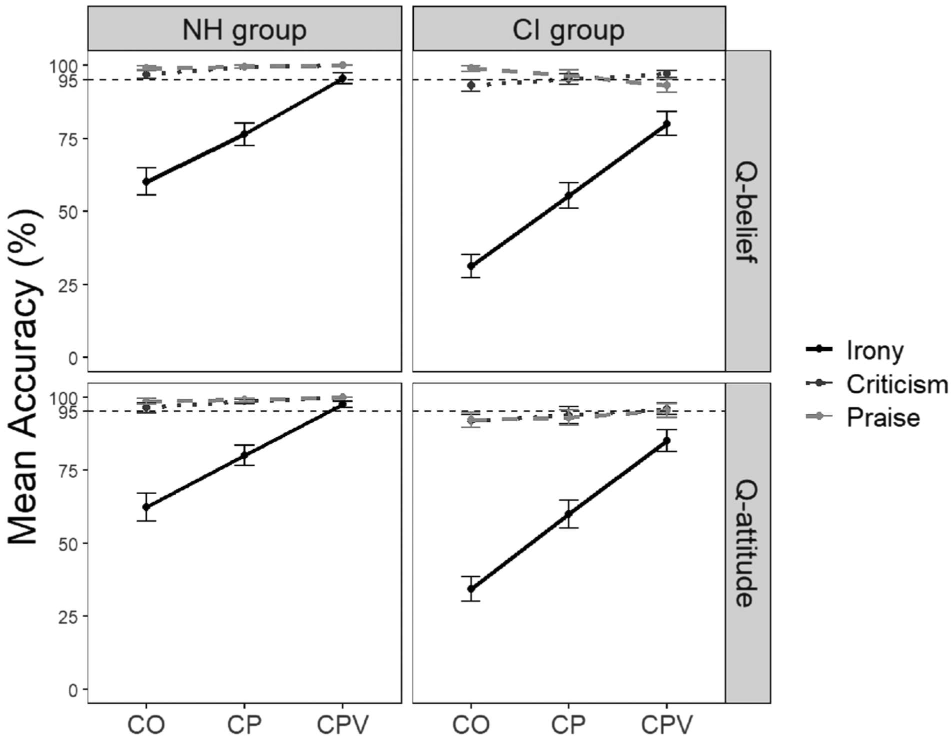
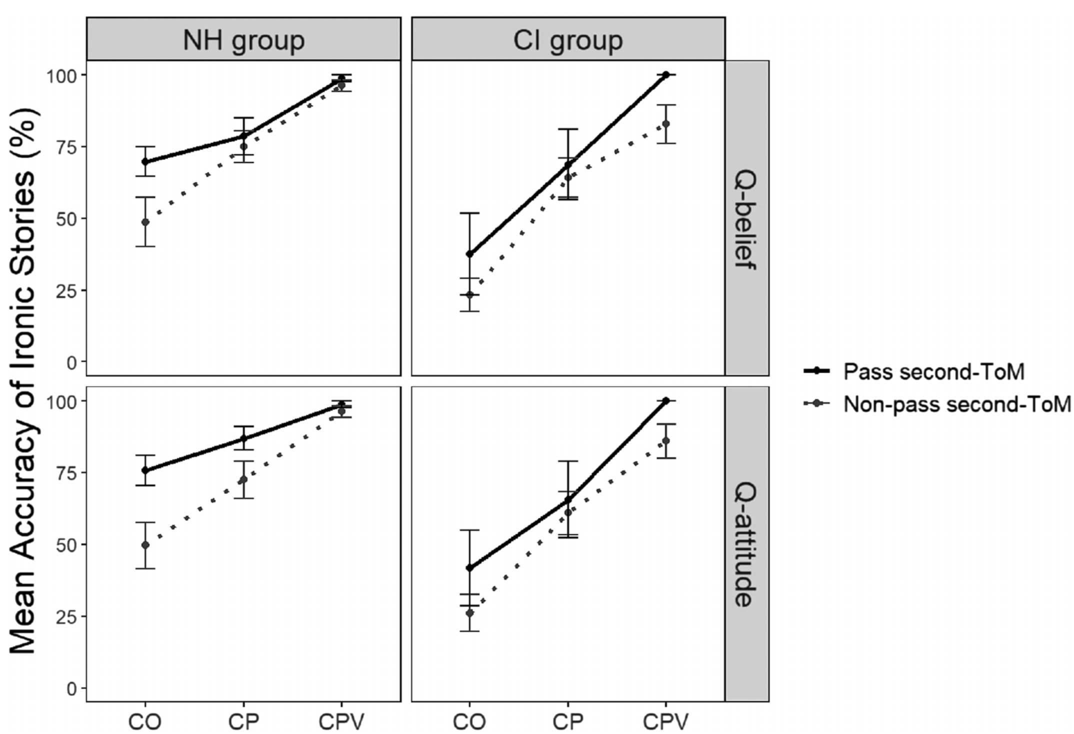


Figure 4. Mean accuracy of Q-belief and Q-attitude in ironic stories by children who passed and who did not pass the second-order Theory of Mind (ToM) test. NH = normal hearing; CI = cochlear implant; CO = context only; CP = context + prosody; CPV = context + prosody + visual.



The results of both models showed significant main effects of “Condition,” Q-belief: $\chi^2(2) = 171.65$, $p < .001$; Q-attitude: $\chi^2(2) = 155.05$, $p < .001$; and “ToM-pass,” Q-belief: $\chi^2(1) = 9.39$, $p = .002$; Q-attitude: $\chi^2(1) = 12.11$, $p < .001$, while the interaction of “Condition × ToM-pass” was only significant in the model of Q-belief, $\chi^2(2) = 7.14$, $p = .028$, but not Q-attitude, $\chi^2(2) = 4.09$, $p = .13$. Despite this, post hoc analyses on the “Condition × ToM-pass” interaction for Q-belief and the main effect of “Condition” on Q-attitude indicated that both the “Pass second-ToM” and “Non-pass second-ToM” groups significantly improved in accuracy from the CO to CP and further to the CPV condition (Appendix D).

Discussion

Our results showed that similar to NH children, children with CIs were able to use both prosodic and facial cues to facilitate their irony comprehension. Furthermore, the benefits from these cues were consistent across groups, no matter whether they passed the second-order ToM test or not.

The benefit from prosodic cues on irony comprehension did not support H1 that children with CIs would not use prosodic cues to enhance irony comprehension due to their limited access to the *F*0 information. A plausible explanation is that these children might resort to other more easily accessible cues of ironic speech, such as speech rate and intensity cues, to facilitate their comprehension (note that our acoustic analysis in Appendix A did confirm that the irony and neutral praise utterances in our stimuli differed in *F*0, speech rate, and intensity cues). This explanation complied with previous research showing that CI users (both children and adults) could comprehensively use acoustic/prosodic cues to perform various tasks involving prosody perception, including emotion recognition (Caldwell et al., 2015; Giannantonio et al., 2015; Hopyan et al., 2011, 2016), question/statement discrimination (Marx et al., 2015; Peng et al., 2009), musical appreciation (Caldwell et al., 2015; Hopyan et al., 2011; Innes-Brown et al., 2013; Kong et al., 2004; Phillips-Silver et al., 2015), and speech act ambiguities resolution (e.g., Zhou et al., 2012). Specifically, Hegarty and Faulkner (2013) suggested the important role of duration for children with CIs in perceiving intonation when pitch information is inaccessible. Similarly, Hopyan et al. (2016) reported that children with CIs relied on tempo in their discrimination of emotions in music. Such preserved ability to use acoustic/prosodic cues in children with CIs also corroborated our findings that these children also benefited from prosody in their irony comprehension.

The benefit from visual cues on irony comprehension supported H2 that children with CIs were able to use visual information to enhance their understanding of irony. This finding suggested a strong visual speech comprehension ability of these children. Such audio-visual advantage has been consistently reported in previous literature focusing on other basic linguistic skills, including lexical tone perception (e.g., Tang et al., 2023), emotion perception (Hopyan-Misakyan et al., 2009; Most & Michaelis, 2012), and word recognition and sentence comprehension (Bergeson et al., 2003, 2005; Geers & Brenner, 1994; Kirk et al., 2007; Lachs et al., 2001). Our results therefore extended prior literature by revealing this audio-visual benefit in the social-pragmatic abilities of children with CIs, with irony comprehension as a test case.

Our results supported H3 that children, both with and without second-order ToM, could effectively use prosodic and visual cues in irony comprehension. This aligned with previous findings that NH children can more readily use prosodic and visual cues to comprehend irony, as compared to relying on the contextual cue, which demands higher-order cognitive reasoning (González-Fuente, 2017; Laval & Bert-Erboul, 2005; Rankin et al., 2009). As suggested by previous research, prosodic and visual cues may directly prompt the negative emotions of the ironic speaker, potentially allowing listeners to detect the negative intention of the speaker (Capelli et al., 1990; Laval & Bert-Erboul, 2005; Nakassis & Snedeker, 2002). Therefore, using these cues to understand irony may not depend on second-order ToM, which is required in context-based inference.

Our study for the first time revealed that Mandarin-speaking children with CIs were able to use prosodic and visual cues to enhance their comprehension of irony. This extended the documented benefits from these multimodal cues in enhancing irony understanding from the NH children to the CI population. Furthermore, our results revealed that the ability to use prosodic and visual cues is unlikely to be contingent on second-order ToM. Therefore, the less ToM-demanding prosodic and visual cues might potentially provide children with CIs a more accessible approach to understanding this complex figurative language of irony, as well as a potential opportunity to leverage their preserved abilities to enhance their irony comprehension. Our findings also call for teachers and parents to be more aware of the importance of these cues in speech comprehension and consider integrating these cues into their teaching practices.

Some limitations should be acknowledged in this study. First, our materials exclusively focused on ironic criticism, the most common type of irony in daily communications, which has also been extensively studied in previous literature (Capelli et al., 1990; Hancock et al., 2000). While this focus aimed to connect our results more effectively to the broader literature with implications for irony comprehension in daily communications, it would limit the scope of our findings as these may not apply to other forms of irony. Future studies may include more types of irony and potentially other types of social affects, thus broadening the scope of the investigation and enriching our understanding of the development of social abilities in children with CIs. Secondly, this study exclusively used female voices for stimuli recording, potentially limiting the generalizability of current findings. Although both the perceptual rating and the acoustic analysis showed that our stimuli were validly perceived as irony and acoustically aligned with Mandarin ironic criticism (Appendix A; Gu et al., 2011; S. Li et al., 2020), male and female speakers might adopt different strategies in irony expressions, as previously observed in emotional speech (Chatterjee et al., 2015; Lin et al., 2022). Therefore, it calls for future studies to examine gender differences in Mandarin ironic expressions and their potential effects (if any) on children's irony comprehension.

In conclusion, although children with CIs face difficulties in irony comprehension using contextual information, they could gain considerable benefits from prosodic and visual cues, even for those without advanced ToM skills. These findings revealed the abilities of children with CIs in integrating prosodic and visual cues into speech comprehension, with new evidence from the social-pragmatic ability of irony understanding. Furthermore, it also implied how multimodal cues could compensate for the difficulty involved in ToM-supported inference in irony comprehension for children with CIs. These results highlighted the accessibility of audio-visual cues in irony comprehension for children with CIs, without reliance on advanced mental reasoning. Our results therefore highlighted the importance of multimodal cues in speech comprehension by children with CIs, also calling for more investigation on the role of multimodal cues in other aspects of language abilities for this population.

Author Contributions

Qianxi Yu: Conceptualization (Lead), Methodology (Lead), Investigation (Lead), Data curation (Lead), Formal analysis (Lead), Writing – original draft (Lead). **Honglan Li:** Methodology (Supporting), Writing – original draft (Supporting). **Shanpeng Li:** Conceptualization (Supporting), Formal analysis (Supporting). **Ping Tang:** Conceptualization (Supporting), Methodology (Supporting), Writing – original draft (Supporting), Funding acquisition (Lead).

Data Availability Statement

The data for this study are available from the corresponding author upon reasonable request.

Ethics Approval Statement

This study was approved by the Ethics Committee of Nanjing University of Science and Technology, China. Informed consent to participate in this study was provided by the guardian of each child.

Acknowledgments

This research is funded by The National Social Science Fund of China (20CYY012). Thanks to all the children, parents, and teachers who participated in this study.

References

Agostinelli, A., Barp, S., Gambalonga, M., Sorrentino, F., Brotto, D., Trevisi, P., & Montino, S. (2021). Pragmatic competence in children with cochlear implant: A new tool to assess the understanding and production of ironic sentences. *Audiologia e Foniatria, 6*(1), 9–14. https://audiologiaefoniatria.padovauniversitypress.it/2021/1/4

Aguert, M. (2022). Paraverbal expression of verbal irony: Vocal cues matter and facial cues even more. *Journal of Nonverbal Behavior, 46*(1), 45–70. https://doi.org/10.1007/s10919-021-00385-z

Anolli, L., Ciceri, R., & Infantino, M. G. (2002). From "blame by praise" to "praise by blame": Analysis of vocal patterns in ironic communication. *International Journal of Psychology, 37*(5), 266–276. https://doi.org/10.1080/00207590244000106

Attardo, S., Eisterhold, J., Hay, J., & Poggi, I. (2003). Multimodal markers of irony and sarcasm. *Humor, 16*(2), 243–260. https://doi.org/10.1515/humr.2003.012

Baltaxe, C. A. (1991). Vocal communication of affect and its perception in three- to four-year-old children. *Perceptual and Motor Skills, 72*(Suppl. 3), 1187–1202. https://doi.org/10.2466/pms.1991.72.3c.1187

Baron-Cohen, S. (1989). The autistic child's Theory of Mind: A case of specific developmental delay. *The Journal of Child Psychology and Psychiatry, 30*(2), 285–297. https://doi.org/10.1111/j.1469-7610.1989.tb00241.x

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bavin, E. L., Sarant, J., Leigh, G., Prendergast, L., Busby, P., & Peterson, C. (2018). Children with cochlear implants in infancy: Predictors of early vocabulary. *International Journal of Language & Communication Disorders, 53*(4), 788–798. https://doi.org/10.1111/1460-6984.12383

Bergeson, T. R., Pisoni, D. B., & Davis, R. A. (2003). A longitudinal study of audiovisual speech perception by children with hearing loss who have cochlear implants. *Volta Review, 103*(4), 347–370.

Bergeson, T. R., Pisoni, D. B., & Davis, R. A. (2005). Development of audiovisual comprehension skills in prelingually deaf children with cochlear implants. *Ear and Hearing, 26*(2), 149–164. https://doi.org/10.1097/00003446-200504000-00004

Caldwell, M., Rankin, S. K., Jiradejvong, P., Carver, C., & Limb, C. J. (2015). Cochlear implant users rely on tempo rather than on pitch information during perception of musical emotion. *Cochlear Implants International, 16*(Suppl. 3), S114–S120. https://doi.org/10.1179/1467010015Z.000000000265

Capelli, C. A., Nakagawa, N., & Madden, C. M. (1990). How children understand sarcasm: The role of context and intonation. *Child Development, 61*(6), 1824–1841. https://doi.org/10.2307/1130840

Chatterjee, M., & Peng, S.-C. (2008). Processing F0 with cochlear implants: Modulation frequency discrimination and speech intonation recognition. *Hearing Research, 235*(1–2), 143–156. https://doi.org/10.1016/j.heares.2007.11.004

Chatterjee, M., Zion, D. J., Deroche, M. L., Burianek, B. A., Limb, C. J., Goren, A. P., Kuikarni, A. M., & Christensen, J. A. (2015). Voice emotion recognition by cochlear-implanted children and their normally-hearing peers. *Hearing Research, 322,* 151–162. https://doi.org/10.1016/j.heares.2014.10.003

Cheang, H. S., & Pell, M. D. (2008). The sound of sarcasm. *Speech Communication, 50*(5), 366–381. https://doi.org/10.1016/j.specom.2007.11.003

Cheang, H. S., & Pell, M. D. (2009). Acoustic markers of sarcasm in Cantonese and English. *The Journal of the Acoustical Society of America, 126*(3), 1394–1405. https://doi.org/10.1121/1.3177275

Colston, H. L., & Gibbs, R. W., Jr. (2002). Are irony and metaphor understood differently? *Metaphor and Symbol, 17*(1), 57–80. https://doi.org/10.1207/S15327868MS1701_5

Connor, C. M., Craig, H. K., Raudenbush, S. W., Heavner, K., & 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 and Hearing, 27*(6), 628–644. https://doi.org/10.1097/01.aud.0000240640.59205.42

Cooper, W. B., Tobey, E., & Loizou, P. C. (2008). Music perception by cochlear implant and normal hearing listeners as measured by the Montreal Battery for Evaluation of Amusia. *Ear and Hearing, 29*(4), 618–626. https://doi.org/10.1097/AUD.0b013e318174e787

Deliens, G., Antoniou, K., Clin, E., Ostashchenko, E., & Kissine, M. (2018). Context, facial expression and prosody in irony processing. *Journal of Memory and Language, 99,* 35–48. https://doi.org/10.1016/j.jml.2017.10.001

Dews, S., & Winner, E. (1997). Attributing meaning to deliberately false utterances: The case of irony. In C. Mandell & A. McCabe (Eds.), *The problem of meaning: Behavioral and cognitive perspectives* (Advances in Psychology, Vol. 122, pp. 377–414). Elsevier Science. https://doi.org/10.1016/S0166-4115(97)80142-2

Duchesne, L., & Marschark, M. (2019). Effects of age at cochlear implantation on vocabulary and grammar: A review of the evidence. *American Journal of Speech-Language Pathology, 28*(4), 1673–1691. https://doi.org/10.1044/2019_AJSLP-18-0161

Duchesne, L., Sutton, A., & Bergeron, F. (2009). Language achievement in children who received cochlear implants between 1 and 2 years of age: Group trends and individual patterns. *Journal of Deaf Studies and Deaf Education, 14*(4), 465–485. https://doi.org/10.1093/deafed/enp010

Filippova, E., & Astington, J. W. (2008). Further development in social reasoning revealed in discourse irony understanding. *Child Development, 79*(1), 126–138. https://doi.org/10.1111/j.1467-8624.2007.01115.x

Filippova, E., & Astington, J. W. (2010). Children's understanding of social-cognitive and social-communicative aspects of discourse irony. *Child Development, 81*(3), 913–928. https://doi.org/10.1111/j.1467-8624.2010.01442.x

Fuller, C. D., Gaudrain, E., Clarke, J. N., Galvin, J. J., Fu, Q.-J., Free, R. H., & Başkent, D. (2014). Gender categorization is abnormal in cochlear implant users. *Journal of the Association for Research in Otolaryngology, 15*(6), 1037–1048. https://doi.org/10.1007/s10162-014-0483-7

Gaudrain, E., & Başkent, D. (2018). Discrimination of voice pitch and vocal-tract length in cochlear implant users. *Ear and Hearing, 39*(2), 226–237. https://doi.org/10.1097/AUD.0000000000000480

Geers, A., & Brenner, C. (1994). Speech perception results: Audition and lipreading enhancement. *Volta Review, 96*(5), 97–108.

Gerrig, R. J., & Goldvarg, Y. (2000). Additive effects in the perception of sarcasm: Situational disparity and echoic mention. *Metaphor and Symbol, 15*(4), 197–208. https://doi.org/10.1207/S15327868MS1504_1

Giannantonio, S., Polonenko, M. J., Papsin, B. C., Paludetti, G., & Gordon, K. A. (2015). Experience changes how emotion in music is judged: Evidence from children listening with bilateral cochlear implants, bimodal devices, and normal hearing. *PLOS ONE, 10*(8), Article e0136685. https://doi.org/10.1371/journal.pone.0136685

Gibbs, R. W., Jr. (2000). Irony talk among friends. *Metaphor and Symbol, 15*(1–2), 5–27. https://doi.org/10.1080/10926488.2000.9678862

Gibbs, R. W., Jr., & Colston, H. L. (2012). *Interpreting figurative meaning*. Cambridge University Press. https://doi.org/10.1017/CBO9781139168779

González-Fuente, S. (2017). *Audiovisual prosody and verbal irony* [Unpublished doctoral dissertation, Universitat Pompeu Fabra].

Good, A., Gordon, K. A., Papsin, B. C., Nespoli, G., Hopyan, T., Peretz, I., & Russo, F. A. (2017). Benefits of music training for perception of emotional speech prosody in deaf children with cochlear implants. *Ear and Hearing, 38*(4), 455–464. https://doi.org/10.1097/AUD.0000000000000402

Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development, 59*(1), 26–37. https://doi.org/10.2307/1130386

Gu, W., Zhang, T., & Fujisaki, H. (2011, August 27–31). *Prosodic analysis and perception of Mandarin utterances conveying attitudes* [Paper presentation]. Proceedings of Twelfth Annual Conference of the International Speech Communication Association, Florence, Italy.

Hancock, J. T., Dunham, P. J., & Purdy, K. (2000). Children's comprehension of critical and complimentary forms of verbal irony. *Journal of Cognition and Development, 1*(2), 227–248. https://doi.org/10.1207/S15327647JCD010204

Happé, F. G. (1993). Communicative competence and Theory of Mind in autism: A test of relevance theory. *Cognition, 48*(2), 101–119. https://doi.org/10.1016/0010-0277(93)90026-R

Hegarty, L., & Faulkner, A. (2013). The perception of stress and intonation in children with a cochlear implant and a hearing aid. *Cochlear Implants International, 14*(Suppl. 4), 35–39. https://doi.org/10.1179/1467010013Z.000000000132

Helms, J., Weichbold, V., Baumann, U., von Specht, H., Schön, F., Müller, J., Esser, B., Ziese, M., Anderson, I., & D'Haese, P. (2004). Analysis of ceiling effects occurring with speech recognition tests in adult cochlear-implanted patients. *ORL, 66*(3), 130–135. https://doi.org/10.1159/000079332

Hopyan, T., Gordon, K. A., & Papsin, B. C. (2011). Identifying emotions in music through electrical hearing in deaf children using cochlear implants. *Cochlear Implants International, 12*(1), 21–26. https://doi.org/10.1179/146701010X12677899497399

Hopyan, T., Manno, F. A., III, Papsin, B. C., & Gordon, K. A. (2016). Sad and happy emotion discrimination in music by children with cochlear implants. *Child Neuropsychology, 22*(3), 366–380. https://doi.org/10.1080/09297049.2014.992400

Hopyan-Misakyan, T. M., Gordon, K. A., Dennis, M., & Papsin, B. C. (2009). Recognition of affective speech prosody and facial affect in deaf children with unilateral right cochlear implants. *Child Neuropsychology, 15*(2), 136–146. https://doi.org/10.1080/09297040802403682

Houston, D. M., Chen, C. H., Monroy, C., & Castellanos, I. (2020). How early auditory experience affects children's ability to learn spoken words. In M. Marschark & H. Knoors (Eds.), *The Oxford handbook of deaf studies in learning and cognition* (pp. 123–137). Oxford University Press.

Innes-Brown, H., Marozeau, J. P., Storey, C. M., & Blamey, P. J. (2013). Tone, rhythm, and timbre perception in school-age children using cochlear implants and hearing aids. *Journal of the American Academy of Audiology, 24*(9), 789–806. https://doi.org/10.3766/jaaa.24.9.4

Kirk, K. I., Hay-McCutcheon, M. J., Frush Holt, R., Gao, S., Qi, R., & Gerlain, B. L. (2007). Audiovisual spoken word recognition by children with cochlear implants. *Audiological Medicine, 5*(4), 250–261. https://doi.org/10.1080/16513860701673892

Kong, Y. Y., Cruz, R., Jones, J. A., & Zeng, F.-G. (2004). Music perception with temporal cues in acoustic and electric hearing. *Ear and Hearing, 25*(2), 173–185. https://doi.org/10.1097/01.AUD.0000120365.97792.2F

Kral, A., & Sharma, A. (2023). Crossmodal plasticity in hearing loss. *Trends in Neurosciences, 46*(5), 377–393. https://doi.org/10.1016/j.tins.2023.02.004

Lachs, L., Pisoni, D. B., & Kirk, K. I. (2001). Use of audiovisual information in speech perception by prelingually deaf children with cochlear implants: A first report. *Ear and Hearing, 22*(3), 236–251. https://doi.org/10.1097/00003446-200106000-00007

Laval, V., & Bert-Erboul, A. (2005). French-speaking children's understanding of sarcasm: The role of intonation and context. *Journal of Speech, Language, and Hearing Research, 48*(3), 610–620. https://doi.org/10.1044/1092-4388(2005/042)

Lenth, R. (2022). *emmeans: Estimated marginal means, aka least-squares means.* R package (Version 1.7.4-1). https://CRAN.R-project.org/package=emmeans

Li, H., & Leung, M. T. (2020). Relations between verb factivity and first-order and second-order false belief understanding: Evidence from Mandarin-speaking typically developing children and children with autism spectrum disorders. *Clinical Linguistics & Phonetics, 34*(1–2), 185–200. https://doi.org/10.1080/02699206.2019.1628810

Li, J. P., Law, T., Lam, G. Y., & To, C. K. (2013). Role of sentence-final particles and prosody in irony comprehension in Cantonese-speaking children with and without autism spectrum disorders. *Clinical Linguistics & Phonetics, 27*(1), 18–32. https://doi.org/10.3109/02699206.2012.734893

Li, S., & Gu, W. (2021, January 24–27). *Prosodic profiles of the Mandarin speech conveying ironic compliment* [Paper presentation]. 12th International Symposium on Chinese Spoken Language Processing, Hong Kong, China.

Li, S., Chen, A., Chen, Y., & Tang, P. (2022). The role of auditory and visual cues in the interpretation of Mandarin ironic speech. *Journal of Pragmatics, 201,* 3–14. https://doi.org/10.1016/j.pragma.2022.09.007

Li, S., Gu, W., Liu, L., & Tang, P. (2020). The role of voice quality in Mandarin sarcastic speech: An acoustic and electroglottographic study. *Journal of Speech, Language, and Hearing Research, 63*(8), 2578–2588. https://doi.org/10.1044/2020_JSLHR-19-00166

Limb, C. J., & Roy, A. T. (2014). Technological, biological, and acoustical constraints to music perception in cochlear implant users. *Hearing Research, 308,* 13–26. https://doi.org/10.1016/j.heares.2013.04.009

Lin, Y. S., Wu, C. M., Limb, C. J., Lu, H. P., Feng, I. J., Peng, S. C., Deroche, M. L. D., & Chatterjee, M. (2022). Voice emotion recognition by Mandarin-speaking pediatric cochlear implant users in Taiwan. *Laryngoscope Investigative Otolaryngology, 7*(1), 250–258. https://doi.org/10.1002/lio2.732

Liu, M., Wu, L., Wu, W., Li, G., Cai, T., & Liu, J. (2018). The relationships among verbal ability, executive function, and Theory of Mind in young children with cochlear implants. *International Journal of Audiology, 57*(12), 881–888. https://doi.org/10.1080/14992027.2018.1498982

Loukusa, S., & Leinonen, E. (2008). Development of comprehension of ironic utterances in 3- to 9-year-old Finnish-speaking children. *Psychology of Language and Communication, 12*(1), 55–69. https://doi.org/10.2478/v10057-008-0003-0

Marx, M., James, C., Foxton, J., Capber, A., Fraysse, B., Barone, P., & Deguine, O. (2015). Speech prosody perception in cochlear implant users with and without residual hearing. *Ear and Hearing, 36*(2), 239–248. https://doi.org/10.1097/AUD.0000000000000105

Miller, S. A. (2009). Children's understanding of second-order mental states. *Psychological Bulletin, 135*(5), 749–773. https://doi.org/10.1037/a0016854

Miller, S. A. (2013). Effects of deception on children's understanding of second-order false belief. *Infant and Child Development, 22*(4), 422–429. https://doi.org/10.1002/icd.1799

Milosky, L. M., & Ford, J. A. (1997). The role of prosody in children's inferences of ironic intent. *Discourse Processes, 23*(1), 47–61. https://doi.org/10.1080/01638539709544981

Moeller, M. P., & Schick, B. (2006). Relations between maternal input and Theory of Mind understanding in deaf children. *Child Development, 77*(3), 751–766. https://doi.org/10.1111/j.1467-8624.2006.00901.x

Most, T., & Michaelis, H. (2012). Auditory, visual, and auditory-visual perceptions of emotions by young children with hearing loss versus children with normal hearing. *Journal of Speech, Language, and Hearing Research, 55*(4), 1148–1162. https://doi.org/10.1044/1092-4388(2011/11-0060)

Nakassis, C., & Snedeker, J. (2002). Beyond sarcasm: Intonation and context as relational cues in children's recognition of irony. In *Proceedings of the Twenty-Sixth Boston University Conference on Language Development* (pp. 429–440). Cascadilla Press.

Nakata, T., Trehub, S. E., & Kanda, Y. (2012). Effect of cochlear implants on children's perception and production of speech prosody. *The Journal of the Acoustical Society of America, 131*(2), 1307–1314. https://doi.org/10.1121/1.3672697

Nelson, N. L., & Russell, J. A. (2011). Preschoolers' use of dynamic facial, bodily, and vocal cues to emotion. *Journal of Experimental Child Psychology, 110*(1), 52–61. https://doi.org/10.1016/j.jecp.2011.03.014

O'Donoghue, G. M., Nikolopoulos, T. P., & Archbold, S. M. (2000). Determinants of speech perception in children after cochlear implantation. *The Lancet, 356*(9228), 466–468. https://doi.org/10.1016/S0140-6736(00)02555-1

Panzeri, F., Cavicchiolo, S., Giustolisi, B., Di Berardino, F., Ajmone, P. F., Vizziello, P., Donnini, V., & Zanetti, D. (2021). Irony comprehension in children with cochlear implants: The role of language competence, Theory of Mind, and prosody recognition. *Journal of Speech, Language, and Hearing Research, 64*(8), 3212–3229. https://doi.org/10.1044/2021_JSLHR-20-00671

Panzeri, F., Giustolisi, B., & Zampini, L. (2020). The comprehension of ironic criticisms and ironic compliments in individuals with Down syndrome: Adding another piece to the puzzle. *Journal of Pragmatics, 156,* 223–234. https://doi.org/10.1016/j.pragma.2019.08.009

Parola, A., Hilviu, D., Vivaldo, S., Marini, A., Diego, D. I., Consolino, P., & Bosco, F. M. (2023). Development of communicative-pragmatic abilities in children with early cochlear implants. *Journal of Child Language*. Advance online publication. https://doi.org/10.1017/S0305000923000405

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods, 51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Peng, S.-C., Lu, N., & Chatterjee, M. (2009). Effects of cooperating and conflicting cues on speech intonation recognition by cochlear implant users and normal hearing listeners. *Audiology and Neurotology, 14*(5), 327–337. https://doi.org/10.1159/000212112

Peng, S.-C., Tomblin, J. B., Cheung, H., Lin, Y.-S., & Wang, L.-S. (2004). Perception and production of Mandarin tones in prelingually deaf children with cochlear implants. *Ear and Hearing, 25*(3), 251–264. https://doi.org/10.1097/01.AUD.0000130797.73809.40

Peng, S. C., Tomblin, J. B., & Turner, C. W. (2008). Production and perception of speech intonation in pediatric cochlear implant recipients and individuals with normal hearing. *Ear and Hearing, 29*(3), 336–351. https://doi.org/10.1097/AUD.0b013e318168d94d

Perner, J., & Wimmer, H. (1985). "John *thinks* that Mary *thinks* that…" Attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology, 39*(3), 437–471. https://doi.org/10.1016/0022-0965(85)90051-7

Peterson, C. C. (2004). Theory-of-mind development in oral deaf children with cochlear implants or conventional hearing aids. *The Journal of Child Psychology and Psychiatry, 45*(6), 1096–1106. https://doi.org/10.1111/j.1469-7610.2004.t01-1-00302.x

Peterson, C. C. (2020). Theory of Mind and conversation in deaf and hearing children. In H. Knoors & M. Marschark (Eds.), *The Oxford handbook of deaf studies in learning and cognition* (pp. 213–231). Oxford University Press.

Pexman, P. M., & Olineck, K. M. (2002). Does sarcasm always sting? Investigating the impact of ironic insults and ironic compliments. *Discourse Processes, 33*(3), 199–217. https://doi.org/10.1207/S15326950DP3303_1

Pexman, P. M., Rostad, K. R., McMorris, C. A., Climie, E. A., Stowkowy, J., & Glenwright, M. R. (2011). Processing of ironic language in children with high-functioning autism spectrum disorder. *Journal of Autism and Developmental Disorders, 41*(8), 1097–1112. https://doi.org/10.1007/s10803-010-1131-7

Phillips-Silver, J., Toiviainen, P., Gosselin, N., Turgeon, C., Lepore, F., & Peretz, I. (2015). Cochlear implant users move in time to the beat of drum music. *Hearing Research, 321,* 25–34. https://doi.org/10.1016/j.heares.2014.12.007

Pluta, A., Krysztofiak, M., Zgoda, M., Wysocka, J., Golec, K., Wójcik, J., Włodarczyk, E., & Haman, M. (2021). False belief understanding in deaf children with cochlear implants. *The Journal of Deaf Studies and Deaf Education, 26*(4), 511–521. https://doi.org/10.1093/deafed/enab015

Rankin, K. P., Salazar, A., Gorno-Tempini, M. L., Sollberger, M., Wilson, S. M., Pavlic, D., Stanley, C. M., Glenn, S., Weiner, M. W., & Miller, B. L. (2009). Detecting sarcasm from paralinguistic cues: Anatomic and cognitive correlates in neurodegenerative disease. *NeuroImage, 47*(4), 2005–2015. https://doi.org/10.1016/j.neuroimage.2009.05.077

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Ren, L., Zhang, Y., Zhang, J., Qin, Y., Zhang, Z., Chen, Z., Wei, C., & Liu, Y. (2022). Voice emotion recognition by Mandarin-speaking children with cochlear implants. *Ear and Hearing, 43*(1), 165–180. https://doi.org/10.1097/AUD.0000000000001085

Rockwell, P. (2001). Facial expression and sarcasm. *Perceptual and Motor Skills, 93*(1), 47–50. https://doi.org/10.2466/pms.2001.93.1.47

Sauter, D. A., Panattoni, C., & Happé, F. (2013). Children's recognition of emotions from vocal cues. *British Journal of Developmental Psychology, 31*(1), 97–113. https://doi.org/10.1111/j.2044-835X.2012.02081.x

Schramm, B., Bohnert, A., & Keilmann, A. (2010). Auditory, speech and language development in young children with cochlear implants compared with children with normal hearing. *International Journal of Pediatric Otorhinolaryngology, 74*(7), 812–819. https://doi.org/10.1016/j.ijporl.2010.04.008

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. (2023). *afex: Analysis of factorial experiments*. R package (Version 1.3-0). https://CRAN.R-project.org/package=afex

Taitelbaum-Swead, R., & Fostick, L. (2017). Audio-visual speech perception in noise: Implanted children and young adults versus normal hearing peers. *International Journal of Pediatric Otorhinolaryngology, 92,* 146–150. https://doi.org/10.1016/j.ijporl.2016.11.022

Tang, P., Shen, Y., Feng, Y., & Li, S. (2023, November 2–5). *Audiovisual perception of Mandarin tones by children with CIs* [Paper presentation]. 48th Annual Boston University Conference on Language Development, Boston, MA, United States.

Tuohimaa, K., Loukusa, S., Löppönen, H., Välimaa, T., & Kunnari, S. (2023). Development of social-pragmatic understanding in children with congenital hearing loss and typical hearing between the ages of 4 and 6 years. *Journal of Speech, Language, and Hearing Research, 66*(7), 2503–2520. https://doi.org/10.1044/2023_JSLHR-22-00700

Vandali, A. E., & van Hoesel, R. J. (2012). Enhancement of temporal cues to pitch in cochlear implants: Effects on pitch ranking. *The Journal of the Acoustical Society of America, 132*(1), 392–402. https://doi.org/10.1121/1.4718452

Wang, H., Wang, Y., & Hu, Y. (2019). Emotional understanding in children with a cochlear implant. *Journal of Deaf Studies and Deaf Education, 24*(2), 65–73. https://doi.org/10.1093/deafed/eny031

Wang, W., Zhou, N., & Xu, L. (2011). Musical pitch and lexical tone perception with cochlear implants. *International Journal of Audiology, 50*(4), 270–278. https://doi.org/10.3109/14992027.2010.542490

Wellman, H. M. (2011). Developing a Theory of Mind. In U. Goswami (Ed.), *The Wiley-Blackwell handbook of childhood cognitive development* (pp. 258–284). Blackwell.

Wellman, H. M. (2014). *Making minds: How Theory of Mind develops*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199334919.001.0001

Widen, S. C. (2013). Children's interpretation of facial expressions: The long path from valence-based to specific discrete categories. *Emotion Review, 5*(1), 72–77. https://doi.org/10.1177/1754073912451492

Widen, S. C., & Russell, J. A. (2008). Young children's understanding of others' emotions. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (pp. 348–363). The Guilford Press.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*(1), 103–128. https://doi.org/10.1016/0010-0277(83)90004-5

Winner, E., & Leekam, S. (1991). Distinguishing irony from deception: Understanding the speaker's second-order intention. *British Journal of Developmental Psychology, 9*(2), 257–270. https://doi.org/10.1111/j.2044-835X.1991.tb00875.x

Zhang, M., & Zhang, J. J. (2006). The influence of shared factors on the learning of the two sides of the referential communication. *Acta Psychologica Sinica, 49*(2), 197–205. https://doi.org/10.3724/SP.J.1041.2017.00197

Zhou, P., Su, Y. E., Crain, S., Gao, L., & Zhan, L. (2012). Children's use of phonological information in ambiguity resolution: A view from Mandarin Chinese. *Journal of Child Language, 39*(4), 687–730. https://doi.org/10.1017/S0305000911000249

Appendix A (p. 1 of 2)

Acoustic Characteristics of the Stimuli Used in This Study

To better understand the acoustic features of our stimuli, we provided acoustic features (mean fundamental frequency [*F*0], *F*0 range, speech rate, mean intensity, and intensity range) for three types of stimuli (irony, affective praise, and neutral praise). The three attitudes have identical lexical content but differ in prosody style (see Table A1).

Table A1. Mean values and standard deviations of mean fundamental frequency (*F*0), *F*0 range, speech rate, mean intensity and intensity range of irony, affective praise, and neutral praise.

Type	Mean *F*0 (Hz)	*F*0 range (max/min)	Speech rate (syllable/s)	Mean intensity (dB)	Intensity range (max–min)
Irony	357.79 (34.12)	2.86 (0.74)	3.36 (0.42)	65.67 (1.64)	34.19 (4.81)
Affective praise	383.50 (24.75)	2.53 (0.57)	4.86 (0.30)	65.41 (1.58)	31.07 (5.35)
Neutral praise	254.21 (7.01)	1.80 (0.42)	4.39 (0.36)	56.88 (2.16)	29.70 (4.39)

We also compared the acoustic difference of irony, affective praise, and neutral praise in the five parameters, using separate linear mixed-effects models. Each model included one of the five parameters as the dependent variable, with a fixed factor “Type” (irony, affective praise, and neutral praise) and a random factor “Item.” The results showed a significant main effect of “Type” on all five parameters (see Table A2).

Table A2. Results of the linear mixed-effects models (main effects of “Type”) on five acoustic parameters.

Parameter	*df1*	*df2*	*F*	*p*
*F*0	2	26	146.73	$< .001^{***}$
*F*0 range	2	26	24.99	$< .001^{***}$
Speech rate	2	26	68.61	$< .001^{***}$
Intensity	2	26	460.25	$< .001^{***}$
Intensity range	2	26	4.28	$.025^{*}$

Note. *F*0 = fundamental frequency.

$^{*}p < .05$. $^{***}p < .001$.

Tukey-HSD post hoc comparison analyses on the main effect of “Type” in each model identified the specific acoustic differences among the three types of stimuli. For *F*0-related cues, irony had a significantly lower *F*0 than affective praise, while both irony and affective praise had a significantly higher *F*0 than neutral praise (see Table A3). Furthermore, irony and affective praise exhibited a significantly greater *F*0 range than neutral praise (see Table A4). Moreover, irony was produced with a slower speech rate compared to affective praise and neutral praise, with affective praise being faster than neutral praise (see Table A5). For intensity cues, irony showed a higher mean intensity (see Table A6), and a larger intensity range than neutral praise (see Table A7).

Table A3. Post hoc comparisons of mean fundamental frequency (*F*0) across irony, affective praise, and neutral praise.

Comparison	β	*SE*	*df*	*t*	*p*
Irony - Affective praise	–25.7	7.99	34.2	–3.218	$.008^{**}$
Irony - Neutral praise	103.6	7.99	34.2	12.963	$< .001^{***}$
Affective praise - Neutral praise	129.3	7.99	34.2	16.18	$< .001^{***}$

$^{**}p < .01$. $^{***}p < .001$.

Table A4. Post hoc comparisons of fundamental frequency (*F*0) range across irony, affective praise, and neutral praise.

Comparison	β	*SE*	*df*	*t*	*p*
Irony - Affective praise	0.324	0.154	25.2	2.109	.108
Irony - Neutral praise	1.061	0.154	25.2	6.898	< .001***
Affective praise - Neutral praise	0.737	0.154	25.2	4.79	< .001***

****p* < .001.

Table A5. Post hoc comparisons of speech rate across irony, affective praise, and neutral praise.

Comparison	β	*SE*	*df*	*t*	*p*
Irony - Affective praise	−1.47	0.128	37.8	−11.496	< .001***
Irony - Neutral praise	−0.984	0.128	37.8	−7.698	< .001***
Affective praise - Neutral praise	0.485	0.128	37.8	3.798	.002**

***p* < .01. ****p* < .001.

Table A6. Post hoc comparisons of mean intensity across irony, affective praise, and neutral praise.

Comparison	β	*SE*	*df*	*t*	*p*
Irony - Affective praise	0.269	0.33	17.9	0.815	.699
Irony - Neutral praise	8.799	0.33	17.9	26.673	< .001***
Affective praise - Neutral praise	8.53	0.33	17.9	25.858	< .001***

****p* < .001.

Table A7. Post hoc comparisons of intensity range across irony, affective praise, and neutral praise.

Comparison	β	*SE*	*df*	*t*	*p*
Irony - Affective praise	3.12	1.57	34.1	1.985	.131
Irony - Neutral praise	4.49	1.57	34.1	2.855	.019*
Affective praise - Neutral praise	1.37	1.57	34.1	0.871	.662

**p* < .05.

Therefore, the acoustic analysis of our stimuli suggested that the verbal irony used in this study was acoustically consistent with that reported in previous studies (Gu et al., 2011; Li & Gu, 2021). Furthermore, the observed acoustic differences between irony and neutral praise in all tested parameters also confirmed that our irony stimuli carry distinct prosodic cues differentiating them from neutral statements, suggesting the validity of our irony stimuli.

Appendix B

Perceptual Rating Results

Table B1. The mean rating score of selected stimuli of ironic criticism.

Type	Modality	Mean score	*SD*
Irony	Audio-only	6.02	0.36
	Visual-only	6.21	1.12
	Audio-visual	6.41	0.69

Note. The rating scale ranges from 1 (*sincere*) to 7 (*ironic*).

Table B2. Mean rating score of affective and neutral stimuli of criticism and praise.

Type	Modality	Mean score	*SD*
Affective criticism	Audio-only	6.33	1.01
	Visual-only	6.01	1.04
	Audio-visual	6.37	0.72
Neutral criticism	Audio-only	1.12	0.32
	Visual-only	1.58	0.82
	Audio-visual	1.41	0.7
Affective praise	Audio-only	6.06	1.02
	Visual-only	6.03	1.07
	Audio-visual	6.2	0.9
Neutral praise	Audio-only	1.12	0.35
	Visual-only	1.26	0.5
	Audio-visual	1.22	0.48

Note. The rating scale ranges from 1 (*neutral*) to 7 (*affective*).

Appendix C (p. 1 of 2)

Statistical Results for Analysis 1

Table C1. Results of *t* tests comparing Q-belief accuracy of criticism and praise with the ceiling level (0.95).

Group	Type	Condition	*M*	*t*	*df*	*p*
CI	Criticism	CO	0.930	−0.961	49	.341
		CP	0.952	0.090	49	.929
		CPV	0.970	1.723	49	.091
	Praise	CO	0.988	4.646	49	< .001***
		CP	0.965	0.827	49	.413
		CPV	0.930	−0.865	49	.391
NH	Criticism	CO	0.968	1.285	49	.205
		CP	0.995	9.000	49	< .001***
		CPV	1.000	NULL	49	NULL
	Praise	CO	0.990	5.716	49	< .001***
		CP	0.995	9.000	49	< .001***
		CPV	1.000	NULL	49	NULL

Note. NULL indicates that all participants in this condition got an accuracy of 1. CI = cochlear implant; CO = context-only; CP = context + prosody; CPV = context + prosody + visual; NH = normal hearing.

***p < .001.

Table C2. Results of *t* tests comparing Q-attitude accuracy of criticism and praise with the ceiling level (0.95).

Group	Type	Condition	*M*	*t*	*df*	*p*
CI	Criticism	CO	0.918	−1.354	49	.182
		CP	0.938	−0.398	49	.693
		CPV	0.960	0.516	49	.608
	Praise	CO	0.922	−1.108	49	.273
		CP	0.930	−0.785	49	.436
		CPV	0.953	0.141	49	.889
NH	Criticism	CO	0.963	0.807	49	.424
		CP	0.985	4.127	49	< .001***
		CPV	1.000	NULL	49	NULL
	Praise	CO	0.985	3.156	49	.003**
		CP	0.990	5.716	49	< .001***
		CPV	1.000	NULL	49	NULL

Note. NULL indicates that all participants in this condition got an accuracy of 1. CI = cochlear implant; CO = context-only; CP = context + prosody; CPV = context + prosody + visual; NH = normal hearing.

**p < .01. ***p < .001.

Appendix C (p. 2 of 2)

Statistical Results for Analysis 1

Table C3. Post hoc comparison results for main effects of "Group" on the accuracy of Q-belief and Q-attitude.

Comparison	Question	β	*SE*	*df*	z	*p*
CI - NH	Q-belief	–0.24	0.04	581	–5.92	< .001***
	Q-attitude	–0.23	0.04	581	–5.41	< .001***

Note. CI = cochlear implant; NH = normal hearing.

****p* < .001.

Table C4. Post hoc comparison results for main effects of "Condition" on the accuracy of Q-belief and Q-attitude.

Comparison	Question	β	*SE*	z	*p*
CO - CP	Q-belief	–0.233	0.0363	–6.414	< .001***
	Q-attitude	–0.253	0.0367	–6.888	< .001***
CO - CPV	Q-belief	–0.454	0.034	–13.368	< .001***
	Q-attitude	–0.465	0.0351	–13.224	< .001***
CP - CPV	Q-belief	–0.221	0.0308	–7.191	< .001***
	Q-attitude	–0.212	0.03	–7.08	< .001***

Note. CO = context-only; CP = context + prosody; CPV = context + prosody + visual.

****p* < .001.

Appendix D

Statistical Results for Analysis 2

Table D1. Post hoc results of the "Condition × ToM-pass" interaction on Q-belief accuracy.

ToM-pass	Comparison	β	*SE*	*z*	*p*
Pass 2nd-ToM	CO - CP	−0.246	0.0749	−3.279	.003**
	CO - CPV	−0.457	0.0729	−6.264	< .001***
	CP - CPV	−0.211	0.0625	−3.374	.002**
Non-pass 2nd-ToM	CO - CP	−0.384	0.0606	−6.335	< .001***
	CO - CPV	−0.576	0.0555	−10.386	< .001***
	CP - CPV	−0.192	0.0516	−3.725	< .001***

Note. ToM = Theory of Mind; CO = context-only; CP = context + prosody; CPV = context + prosody + visual.

**$p < .01$. ***$p < .001$.

Table D2. Post hoc comparison results of the main effect of "Condition" on Q-attitude accuracy.

Comparison	β	*SE*	*z*	*p*
CO - CP	−0.268	0.0463	−5.784	< .001***
CO - CPV	−0.487	0.0437	−11.156	< .001***
CP - CPV	−0.219	0.0395	−5.554	< .001***

Note. CO = context-only; CP = context + prosody; CPV = context + prosody + visual.

***$p < .001$.