Spoken Sentence Complexity and Grammar Use in Children with CIs

Michaela Socher1,2 and Elisabeth Ingo1,∗

1Department of Behavioural Sciences and Learning, Linköping University, Linköping, Sweden
2Fraunhofer Institute for Building Physic, Stuttgart, Germany

Abstract

This study investigated if the language profiles of prelingually deaf children with bilateral cochlear implants (CIs) and children with typical hearing (TH) matched on their quantitative score on clinical spoken expressive language tasks differed in terms of sentence complexity, sentence length, and severity of grammatical errors. No significant differences were found between the groups in terms of (1) proportion of simple, conjoined, and complex sentences; (2) mean length of utterance based on words and syllables; and (3) proportion of local and global grammatical errors. The results indicate that the quantitative scores on the clinical spoken expressive language tasks are related to similar spoken language profiles in children with CIs and children with TH. These findings suggest that these tests can be used for meaningful comparisons of expressive spoken language skills. However, more studies are needed on the real-life expressive language skills of children with CIs, as clinical tests often rely on one specific modality (in this study: spoken language) and might therefore not accurately represent the language skills of the children.

Children with cochlear implants (CIs) have been claimed to be at risk for delays in spoken expressive grammar ability (Boons et al., 2013a, 2013b; Inscue et al., 2009; Wie et al., 2020; but see Socher et al., 2020 for contradicting results). As many children with CIs are born to hearing parents (Mitchell & Karchmer, 2004) and grow up with oral language as their main communication mode, delays in spoken expressive grammar might have negative effects on their ability to communicate (but see Hall & De Anda, 2021 for information on different language access profiles). Not being able to communicate in an age-appropriate way is a risk factor for social problems with peers (Ross & Weinberg, 2006) and better language ability has been found to be associated with better quality of life for children with CIs (Haukedalet al., 2020).

Spoken expressive grammar comprises the ability to build both long and short sentences using different grammatical structures without making syntactical or morphological errors. Most studies on the spoken expressive grammar of children with CIs rely on one specific measure, for example, a closure task (the ability to use the correct word ending when for example building plural form: Wie et al., 2020), the mean length utterance (average sentence length; Nicholas & Geers, 2018; Samiee et al., 2020; Werfel et al., 2021), or an overall measure of expressive grammar such as the Formulated sentences task (Semel et al., 2004). The latter is a helpful clinical tool to investigate language delays and to evaluate which grammatical structures have been mastered by a child. However, in research, the standard procedure of scoring of the test is often used. This scoring procedure provides an estimate of the expressive grammar level of a child, but it cannot be used to investigate qualitative differences in expressive spoken grammar. The standard scoring procedure of the task might underestimate or mask differences in the language profiles of children with CIs and children with typical hearing (TH). To investigate if there are qualitative variations between the expressive spoken grammar of children with CIs and children with TH, it is important to evaluate the language profiles of the children in more detail and not rely on only quantitative global scores. Measures, such as the length and complexity of produced sentences, and the severity of possible grammatical errors should be taken into consideration.

The length and complexity of sentences are connected. Sentences with more complex grammatical structures tend to include more words. However, length and complexity are not a one-to-one match and different measures can be used for these two properties of a sentence. Sentence length is often measured using the mean length of utterance (MLU; Ariman-Rupp et al., 1976; Brown, 2013; Hickey, 1991) based on either word, syllable, or morpheme level. Sentence complexity can be measured by investigating how many simple, conjoined, and complex (including >1 sub-clause) sentences a child produces (Werfel et al., 2021; Yaruss, 1999).

Werfel et al. (2021) investigated if there are any indications for disrupted development of sentence complexity for deaf and hard of hearing (DHH) children with different hearing levels (~29% profound hearing loss) aged 4–5 years using spoken language. Their results indicated disrupted development for some complex sentence structures but showed that DHH children and children with TH matched on MLU in morphemes produced the same proportion of complex sentences. However, although the children in the study were matched on MLU, they were not matched on their overall spoken expressive language level and differed in terms of their vocabulary and grammar ability. The DHH children were compared with children with TH with higher scores on expressive vocabulary, higher scores on a test for morphosyntax, and higher scores on an omnibus spoken language test. Lexical skills and sentence complexity are associated (Hoff et al., 2018). Therefore, if children differ in their vocabulary skills, they...
would be expected to show significant differences in sentence complexity. This might explain the differences found by Werfel et al. (2021). However, it could also be argued that difficulties with longer and more complex sentences could be due to differences in verbal working memory ability. Previous research has indicated significant differences between children with CIs and children with TH in terms of verbal working memory ability (Akçakaya et al., 2019; Nittroer et al., 2017). Results by Willstedt-Svensson et al. (2004) show an association between verbal working memory and expressive language skills in children with CIs. If differences in verbal working memory ability would lead children with CIs to produce on average shorter sentences, this difference in the language profile would not be captured by the standard procedure of scoring for the Formulated Sentences Test (Semel et al., 2004).

Also, to evaluate spoken language profiles, it is not only important to take complexity and sentence length but also grammatical errors into consideration.

One way of analyzing grammatical errors is to divide them into local and global errors (Dulay & Burt, 1972; Shaffer, 2005). The division into local and global errors depends on the grammatical structure of the language. Local errors refer to errors that do not hinder the overall understanding of sentences (i.e., in English inconsistent tense, the wrong grammatical number, or preposition are considered local errors). These kinds of errors can, however, still influence how easily sentences can be comprehended by a recipient. Global errors refer to errors that affect the whole sentence structure and hinder understanding (i.e., in English a missing subject, object, or verb is considered a global error).

Listening with CIs makes the acquisition of some spoken language properties more challenging. Previous studies indicate that children with CIs have specific problems with morphology (Davies et al., 2020; Hansson et al., 2017). This is most likely due to difficulties in discriminating between different morphological markers due to the reduced auditory quality provided by a CI in comparison to TH. As Hammer (2010) suggests, problems perceiving grammatical morphemes might slow down a child’s language acquisition. This might lead to a higher proportion of local errors in sentences produced by children with CIs. However, such differences would not be captured by the standard procedure of scoring for the Formulated Sentences Test (Semel et al., 2004).

A study by Boons et al. (2013a) looked at differences in terms of grammatical error severity between children with CIs aged 5–13 years and age-matched peers with TH. According to the definition used by Boons et al. (2013a), more severe errors include: an incomplete sentence; a complete sentence with >2 deviations in grammar or semantics; a complete sentence that does not make sense; a complete sentence that does not contain the target word or phrase; or a complete sentence that is not related to the context of the picture. Boons et al. found that children with CIs aged 5–13 years make more severe grammatical errors in their sentences produced for the Formulated Sentences Test (Semel et al., 2004) in comparison with age-matched peers with TH. This was the case even though the groups do not differ significantly on their overall score on the Formulated Sentences Test (Semel et al., 2004). This could suggest that although the groups perform on a similar level on the Formulated Sentences Task in terms of the score received by the standard procedure of scoring, they still have different language profiles. The results could, however, also be explained by the fact that the compared groups in the study by Boons et al. were not matched. The groups compared by Boons et al. (2013a) did not differ significantly in their score on the Formulated Sentences Test (Semel et al., 2004). But, as laid out by Mervis and Klein-Tasman (2004), non-significant differences do not indicate that two groups are highly similar on a skill. In accordance with Frick’s categories (see Frick, 1995), the p-value should be >.05 or larger to consider groups definitely matched. The differences in terms of error severity found by Boons et al. (2013a) could therefore be explained by differences in the overall expressive spoken grammar skills of the compared groups.

In summary, there is a lack of knowledge concerning the expressive grammar skills of children with CIs. There is some indication that the standard procedure of scoring for the Formulated Sentences Test (Semel et al., 2004) may not accurately represent differences in the language profile of children with CIs and children with TH. The current study is an extension of a published study (Socher et al., 2020) comparing the expressive grammar skills of children with CIs and children with TH using the Swedish version of the Formulated Sentences Test (Semel et al., 2004). The current study is exploratory and investigates the language profiles in terms of sentence complexity, sentence length, and severity of grammatical errors of children with CIs and children with TH matched on their score on the Formulated Sentences Test (Semel et al., 2004) and expressive vocabulary (CELF-IV, Expressive vocabulary, Semel et al., 2004). The aim is to investigate if the language profiles of the two groups differ in any of the analyzed areas. The results will give important indications concerning whether the Formulated Sentences Test (Semel et al., 2004) may not accurately represent differences in the expressive spoken language profiles of children with CIs and children with TH.

Method
Participants

The data used for the current paper have been collected for a previous study (Socher et al., 2020) and the children have been tested on more language and cognitive tests than reported in the current study. The project was approved by the local ethics committee in Linköping, Sweden (Ref. 2015/308-31). The quantitative measure (scores) of the Formulated Sentences Test of the Clinical Evaluation of Language Fundamentals 4 (CELF-IV, Semel et al., 2004) and the Expressive Vocabulary task of the CELF-IV (Semel et al., 2004) have been used in the previous publication. The original dataset included 55 children. Seventeen of them had one or two CIs and 38 had TH. It has been found that the spoken language development of children with bilateral and unilateral CIs differs, with children with unilateral implants performing significantly more poorly on receptive and expressive language tests than children with bilateral CIs (Eskridge et al., 2021). Only prelingually deaf children (deafness detected before the child was 1 year of age) with bilateral CIs were included in the current study. The final sample included 12 children with CIs. Four of the children used oral language only, six used oral language and sign as support (~10% of the time according to parent report), two children were bilingual using both sign language and oral language (1 of them using both modalities simultaneously). Twelve children with TH were pair-matched to the children with CIs on their score on the Formulated Sentences Test (Semel et al., 2004).

If several matches were possible, the one better matched in terms of expressive vocabulary was used. For one child with CIs, no exact match could be found. A child with TH with 2 points less on the Formulated Sentences Test (Semel et al., 2004) was therefore used as a match. As the resulting groups were not definitely matched according to Fricks’s categories (p < .05, see Frick, 1995; Mervis and Klein-Tasman, 2004) on the expressive vocabulary score, two children with TH were replaced with children with TH matched.
on expressive vocabulary, but with a point difference of 2 and −1, respectively, on the Formulated Sentences Test (Semel et al., 2004). This was done to make sure the groups were comparable in overall expressive grammar ability according to their score on a standard clinical task and in terms of expressive vocabulary. The study aimed to compare children with CIs to children with TH with comparable expressive language ability according to assessment measures. Therefore, it was important to match the groups both on their score on a standard expressive grammar and their score on a standard expressive vocabulary task. This final sample of children included 12 children, matched with the children with CIs on both their score on a standard expressive grammar task (p = .7; see Table 1) and their score on a standard expressive vocabulary task (p = .7; see Table 1). The descriptive data for age, hearing age, duration of CIs use, performance on the Formulated Sentences Test (Semel et al., 2004), and the expressive vocabulary task are shown in Table 1. The children with CIs had a median age of 6 years and 5 months (Range = [5 years 4 months–8 years 1 month]). Their median age of implantation was 8.5 months (Range = [5 months–36 months]) and their median duration of using CIs was 5 years and 5.5 months (Range = [3 years and 3 months–6 years and 3 months]). The median age of the children with TH was 6 years and 6.5 months (Range = [6 years 1 month–7 years 5 month]). All children with TH attended mainstream education.

### Formulated Sentences

The standard test for evaluating children’s expressive grammar skills used in research and clinical work in Sweden is the Formulated Sentences task, part of the CELF-IV (Semel et al., 2004). The test consists of 12 pictures and the child is instructed to describe each picture using a given target word. The degree of difficulty increases, with later target words requiring the child to produce more complex sentences. Correct sentences render 2 points, sentences with a maximum of two grammatical errors 1 point, and sentences with more than two errors 0 points. After four consecutive 0 points, the test is terminated. In this study, the raw score on the Formulated sentences task was used to match the children with CIs and the children with TH. This was done to make sure the children were comparable in terms of their raw scores. The sentences produced by the children for the Formulated sentences task were subsequently analyzed in terms of sentence complexity and length as well as in terms of numbers of global and local errors. This made it possible to investigate if children with CIs and children with TH matched on their raw scores on both a spoken expressive grammar and a spoken expressive vocabulary task show the same pattern in terms of error types and sentence complexity.

### Materials

#### Expressive Vocabulary

The standard test for evaluating children’s expressive vocabulary skills used in research and clinical work in Sweden is the Expressive Vocabulary task, part of the CELF-IV (Semel et al., 2004). In this task, children are asked to name pictures or parts of pictures. For the tested age group, the children had to name 18 pictures in total. The children received one point for every correct answer. If the child was not able to answer four consecutive trials, the task was terminated. In the current study, the raw score on the expressive vocabulary task was used to match the children with CIs and the children with TH on their expressive vocabulary skills.

#### Testing Procedure

The children were tested individually by a speech and language pathologist or by a speech and language pathology student in their last term. All the testing was done orally. The children were tested either in a quiet room in their school or a quiet room at home. The children were tested on more tasks as reported here and the data used in the current study were collected for a larger research project (Socher, 2020). The test sessions were recorded, and the order of tests was randomized. The test leader used the practice trails of each of the tasks to make sure the child understood the instruction.
Data Preparation

Sentence Complexity and Sentence Length

The length and complexity of the sentences produced by the children for the Formulated sentences task were analyzed. Different measures can be used to investigate the complexity of spoken language. The present study used a combination of sentence complexity and sentence length. Mean Length of Utterance measure (MLU; Arltman-Rupp et al., 1976; Brown, 2013; Hickey, 1991), calculating the mean number of words, syllables, and morphemes produced in sentences, is one common measure for grammar development. Calculating MLU on the morpheme level is standard in English literature, but studies suggest it to be less reliable in highly inflected languages (Hickey, 1991), such as Swedish. Hence, MLU on morpheme level has not been included in the present study. Producing long sentences and achieving a high MLU require certain grammatical skills. MLU is therefore considered a reliable indirect measure for overall grammatical development (Arlman-Rupp et al., 1976). However, it does not take specific indicators of sentence complexity into consideration.

Another estimate of grammar development is analyzing the level of complexity in sentences produced. In Swedish, sentences can be simple (consisting of one main clause), conjoint (consisting of more than 1 main clause), or complex (consisting of 1 or more main clauses and 1 or more subclauses). A clause is a full sentence with a subject, verb, and object. A subclause is a subordinate clause that cannot stand alone and can only be comprehended with the information from the main clause. All grammatically correct sentences were assessed for complexity and length by a native Swedish speaker (El). Each sentence was scored as being simple, conjoint, or complex. The nature of the Formulated Sentence Test results in varying numbers of sentences produced for each child. The ratio of simple, conjoint, and complex sentences was therefore used as an outcome measure. The overall sum of words and syllables for each child was calculated and divided by the number of produced sentences to estimate the MLU on word and syllable level.

Local and Global Grammatical Errors

The grammatical errors in the sentences produced by the children for the Formulated sentences task were classified into local and global errors. There are several ways to classify grammatical errors. The division into local and global errors (introduced by Burt & Kiparsky, 1972) is focused on how grammatical errors disrupt communication. Errors that hinder successful communication (i.e., syntax, and sentence connections) are considered global, whereas minor mistakes affecting only part of the sentence are considered local (Shaffer, 2005).

All grammatically incorrect sentences were scored as incorrect due to either local or global errors by a native Swedish speaker (EI). Errors considered local, and not hampering communication, included wrong/inconsistent tense, article, gender, number category (singular/plural), and preposition. All other errors, including free-standing sub-clauses, were considered global errors.

Statistical Analysis

We used R (Version 3.6.0; R Core Team, 2019) for all our analysis. To tidy the data, the R-packages car (Version 3.10; Fox & Weisberg, 2019), carData (Version 3.4; Fox et al., 2020), dplyr (Version 1.3; Wickham et al., 2021), purrr (Version 3.4; Henry & Wickham, 2020), readr (Version 1.4.0; Wickham & Hester, 2020), tibble (Version 3.5; Müller & Wickham, 2021), tidyr (Version 1.1.2; Wickham, 2020), and tidyverse (Version 1.3.0; Wickham et al., 2019) were used. To plot the data and to create the tables, the packages flextable (Version 6.3; Gohel, 2021), forcats (Version 0.5; Wickham, 2021), ggplot2 (Version 3.3.3; Wickham, 2016), ggpurrr (Version 4.0; Kassambara, 2020a), and gsubsummary (Version 1.3.7; Sjoberg et al., 2021) were used. For the statistical analysis, the packages effsize (Version 0.8.1; Wickham, 2016) and rstatix (Version 0.7.0; Kassambara, 2020b) were used. For writing the report, the package papaja (Version 1.9997; Aust & Barth, 2017) was used.

First, the ratio of grammatically correct sentences and the ratio of grammatically incorrect sentences were calculated for each child. Subsequently, the ratio of grammatically correct and grammatically incorrect sentences was compared between the groups. In addition, for the grammatically correct sentences, the ratio of simple, conjoint, and complex sentences as well as the MLU based on words, and the MLU based on syllables were compared between the groups. For the grammatically incorrect sentences, it was evaluated if the sentences were rated incorrectly due to local or due to a global error. Subsequently, the ratio of sentences with a global and sentences with a local error was compared between the groups. As all data were found to be non-normally distributed, Wilcoxon rank sum tests were used for the comparison. The datasets generated for this study will not be made publicly available. It was ensured to the parents in the information letter that no data will be send to anyone not part of the research team. This was also included in the ethics application.

Results

The median number of included sentences (completed sentences) was 10 (IQR = [3, 14], Range = [0, 20]) for the children with CIs. The median number of included sentences (completed sentences) was 12 (IQR = [4, 17], Range = [0, 21]) for the children with TH. The median number of correct sentences was 4 (IQR = [2, 12], Range = [0, 16]) for the children with CIs. The median number of correct sentences (completed sentences) was 6 (IQR = [2, 10], Range = [0, 18]) for the children with TH. No significant group difference between children with CI and children with TH was found for the ratio of grammatically correct, $W = 85.00, p = .468$, $\chi^2 = .02$, or ratio of grammatically incorrect sentences, $W = 59.00, p = .468, \chi^2 = .02$. A graphical representation of the results is provided in Figure 1.
in addition, no significant group difference was found for the ratio of simple sentences, $W = 91.00$, $p = .267$, $\epsilon^2 = .06$, and ratio of complex sentences, $W = 74.00$, $p = .928$, $\epsilon^2 < .001$. For a graphical representation of the results see Figure 2.

No significant difference between children with CI and children with TH was found in terms of MLU for words, $W = 88.50$, $p = .355$, $\epsilon^2 = .04$, or MLU for syllables, $W = 341$, $p = .831$, $\epsilon^2 < .001$. For a graphical representation of the results see Figure 3.

Most grammatical errors in both groups were global ones. There was no significant difference in the ratio of local, $W = 71.50$, $p > .999$, $\epsilon^2 < .001$, or global errors, $W = 68.00$, $p = .831$, $\epsilon^2 < .001$, between children with CI and children with TH. For a graphical representation of the results see Figure 4.

**Discussion**

The results of the current study suggested that children with CIs (early and bilateral implants; median = 8.5 months) and children with TH matched on the quantitative scores on clinical measures of spoken expressive grammar and spoken expressive vocabulary had comparable spoken language profiles in terms of sentence complexity, sentence length, and type of grammatical errors. The Formulated Sentences Test (Semel et al., 2004) from the CELF-JV battery did seem to accurately represent differences in these areas for children with CIs and children with TH using spoken Swedish age 5–8 years.

**Sentence Complexity and Sentence Length**

No difference in terms of the complexity of the sentences used in the Formulated Sentences Test (Semel et al., 2004) was found between children with CIs and children with TH. Most children (CIs and TH) produced mainly simple sentences, which was most likely due to the age of the tested children. Furthermore, children with CIs and children with TH produced sentences of similar length, and no significant difference in terms of MLU based on words or syllables was found. This result was in accordance with Werfel et al. (2021) who found that DHH children and children with TH aged 4–5 years who produced sentences with similar MLU in morphemes also produced similar rates of complex sentences. However, in the current study, the children with TH and children with CIs were matched on chronological age, while the groups differed by on average 5 months in the study by Werfel et al. (2021). One explanation for this might be that the children tested in the current study were older than the children tested in the study by Werfel et al. (2021). A study by Faes et al. (2015) indicated that children with CIs catch up with their peers with TH in terms of MLU in words by the age of 7.

**Local and Global Grammatical Errors**

Most children (CIs and TH) produced mainly global errors, which was most likely due to the age of the children. It was expected that they were not yet able to use all grammatical structures tested in the Formulated Sentences Test (Semel et al., 2004). The results from the current study suggested that children with CIs produced the same ratio of global and local grammatical errors as children with TH matched on expressive spoken grammar and expressive spoken vocabulary. This was in accordance with results by Guo and Spencer (2017) who found that the error patterns of children with CIs aged 4.5–6.5 years were similar to those of children with TH. However, Boons et al. (2013a) found that children with CIs aged 5–13 years produced more severe errors than children with TH even if they had expressive grammar ability within their age range. In addition, the current results differed from studies who found differences between children with TH and children with CIs in terms of morphological skills (Davies et al., 2020; Hansson et al., 2017; Wie et al., 2020). Specific problems with morphological skills should lead to a higher proportion of local
errors for children with CIs. The difference between our results and the results from previous studies was most likely due to the differences in the matching procedure. The studies referred to here had an applied approach with the aim to investigate possible overall delays in grammar development. In comparison, we aimed to investigate if the language profiles in terms of error type differed for children with CIs and children with TH matched on their standard expressive grammar and expressive vocabulary score.

Limitations
Results from the current study cannot be used to evaluate the real-life communication skills of children with CIs. The language access profiles of the included children were diverse and only four relied solely on oral communication. Six used oral language and sign as support and two were bilingual (Swedish sign - and oral language). Therefore, the real-life expressive language skills of the children might have differed from the skills indicated by the Formulated Sentence task used in the current study. In addition, it is important to note that the current study included a small sample. The results should be interpreted with caution. Children with CIs is a small group in Sweden and aiming for a homogeneous sample will result in a small study sample. Also, the group of children tested here had bilateral CIs and had been implanted at a young age. The results do not necessarily apply to children with CIs who received their implants later or who are unilaterally implanted.

Future Research
Replications of the study, as well as comparisons of studies on children with different language backgrounds, are needed to further investigate the development of expressive spoken grammar of children with CIs. In addition, these studies are needed to investigate if the Formulated Sentences Test (Semel et al., 2004) may not accurately represent differences in the language profiles of children from different age groups or who use different languages than spoken Swedish. In the current study, oral language has been tested. In future research, it would be especially important to look at the real-life multimodal conversation skills of children with CIs. Standard language tests often measure very specific language abilities, and it is seldom investigated if problems in, for example, specific grammar areas translate to any communication problems. A replication of the study design presented in the current study using real-life conversation data and a larger study sample would be of importance.

Conclusion and Implications
When children with CIs and children with TH are matched on their scores on clinical tests of spoken expressive grammar and spoken expressive vocabulary, they show comparable language profiles in terms of sentence complexity, sentence length, and error types. This is an encouraging result as it indicates that these standard clinical measures can be used to meaningfully compare the spoken expressive language of children with CIs and children with TH. The scores the children receive do not seem to be related to different language profiles, at least not at an age at which most children use short and simple sentences (age 5–8). Further studies are needed to investigate the real-life expressive grammar profiles of children with CIs. In such studies, it would be important to take language ability as a whole and not just spoken language into consideration. In addition, more research is needed to investigate how real-life communication is influenced by different subskills of expressive language.

Funding
This research was funded by the European Union Seventh Framework Program [FP7/2007–2013], the Swedish Research Council for Health, Working Life and Welfare [2013–01363], and by the Swedish Association of Hard of Hearing People [HFF19–0032].

Conflicts of Interest
The authors have no conflicts of interest to disclose.

Acknowledgments
A special thanks to all the children and caregivers participating in this study. The authors would like to acknowledge Linn Helgrien, Malin Gärskog, Ingrid Hedström, and Elias Ingebrand for helping with the data collection, and Björn Lyxell for support.

References

