Speech Prosody 2024
2-5 July 2024, Leiden, The Netherlands

Prosodic Clitics in English-speaking Children's Speech Production – An Acoustic Study

Rui Cai[1], *Paul Boersma1*[1], *Ivan Yuen*[2], *Katherine Demuth*[3], *Titia Benders*[1]

[1]University of Amsterdam
[2]Universität des Saarlandes
[3]Macquarie University

r.cai@uva.nl, p.p.g.boersma@uva.nl, ivan.yuen@mq.edu.au,
katherine.demuth@mq.edu.au, a.t.benders@uva.nl

Abstract

English-speaking children have been suggested to cliticize function words as early as 2 years of age. However, several limitations to previous research, notably the absence of identical acoustic measures of cliticization across all ages, pose challenges in elucidating apparent differences between 2-year-olds and school-aged children. Thus, this study aims to apply established methods from adult research to children's speech and provide insight into how children acoustically realize cliticization in their productions. The study explored the production of cliticization by comparing two prosodic structures: No-Clitic (e.g., *Boys often cut cards*) vs. Potential-Clitic (e.g., *Boys often cut the cards*). A total of 32 children, 3-year-olds (N=12) and 12-year-olds (N=20), were drawn from AusKidTalk, an audio corpus of Australian-English children's speech. This selection of the corpus' youngest and oldest age groups enables an initial exploration of age-related change in cliticization. The results show that, on average across the ages, children shorten verb durations in the Potential-Clitic condition compared to the No-Clitic condition, indicating that children cliticize articles in the leftward direction. Our findings then further suggest that a trading relationship between verbs and articles exists in the process of children's acquisition.

Index Terms: prosodic structure, cliticization, prosodic development.

1. Introduction

The acquisition of grammatical function words occurs gradually over an extended period to attain adult-like proficiency. Most researchers have attributed the omission of function words in early childhood to children's immature prosodic structure (see overview by Kehoe [1]). The current study considers the next step in the development of function words, once children have stopped omitting function words and thus appear to have acquired them: how do children then integrate function words into the overall prosodic structure?

English articles, functioning as prosodic clitics, can attach to an adjacent (lexically stressed) content word. However, the direction of cliticization of monosyllabic articles is not always clear. Articles are used to identify referents in a noun phrase. Some suggest that this close semantic–syntactic relationship should induce articles to undergo rightward cliticization with the following noun to form an iambic foot or (in English) attach at the higher level of the prosodic word [2]–[4]. However, other studies argue that if there is a monosyllabic content word preceding the article, it will show leftward cliticization [5], [6]. Given that most English function words are produced as weak syllables, they are highly constrained to prosodify as part of a trochaic foot, combining a strong and weak syllable. Several studies, using acoustic analysis from psycholinguistic experiments, support leftward cliticization [7], though others suggest that adult speakers may cliticize to the right [8].

Several studies have shown that English-speaking toddlers often omit the article when it cannot be prosodified as part of a trochaic foot [3], [9]–[13]. Gerken introduced the metrical foot constraint to investigate children's omission of function words using an elicited imitation task [13]. The results show that two-year-old children are more likely to produce articles when they occur in a footed context (1a) compared to an unfooted context (1b), tending to produce articles as part of a trochaic foot. This research, based on omission patterns, supports the leftward cliticization in early development.

1a. Tom [pushed the]$_{FT}$ zebra.

1b. Tom [pushes]$_{FT}$ the zebra.

Demuth & McCullough [12] later conducted a longitudinal study of children's spontaneous speech, replicating Gerken's findings. However, acoustic analysis of one child who exemplified a slightly different pattern showed that she first produced articles as independent prosodic words at 1;9 years, then integrated articles with the preceding word to form a trochaic foot at 2;0 years, providing further support for leftward cliticization.

Contrary to previous studies with 2-year-olds, Redford [14] suggests that school-aged children (5- and 8-year-olds) cliticize articles more tightly with the following noun when they are in footed position (e.g., *hits the bat*) than in unfooted position (e.g., *pushes the bat*), by measuring anticipatory coarticulation [15].

Thus, previous studies have suggested diverse cliticization patterns for two age groups: toddlers and school-aged children. A potential factor contributing to these different findings is the absence of identical tasks and measures across age groups, posing challenges for understanding apparent differences between 2-year-olds [12] and school-aged children [14]. One acoustic measure used in most investigations of prosodic cliticization is polysyllabic shortening, showcasing instances of verb shortening under leftward cliticization [7] and object noun shortening under rightward cliticization in adults [8]. We therefore use this well-established paradigm below.

Thus, the present study aims to apply established methods from adult research to answer the following research questions:

1. Do young age children produce function words with leftward cliticization but school-aged children with

10.21437/SpeechProsody.2024-12

rightward cliticization? The leftward hypothesis predicts that the article cliticizes with the preceding word, while the rightward hypothesis predicts cliticization with the following word.

Even if younger and older children were to cliticize articles in the same direction, that would not mean that they cliticize articles in exactly the same way. Many studies have indicated that it takes time for children to control speech duration, possibly leading to a limited reduction of articles [10], [14], [16], especially for younger children. Therefore, this study also aims to shed light on the acoustic implementation of cliticization, asking the question:

2. If young and older children cliticize articles in the same direction, is there any age-related change in the acoustic realization of cliticization? It is expected that younger children may reduce articles less than older children.

2. Method

2.1. Corpus

The data utilized for this study were sourced from AusKidTalk, a corpus comprising audio recordings of speech in Australian English [17]. Our study focuses exclusively on investigating articles within the Sentence Repetition task in this corpus. In the Sentence Repetition task, each child was instructed to listen to and repeat each of the 36 pre-recorded sentences. Speech in each trial produced by each child was saved as a single WAV file.

2.2. Participants

A total of 32 Australian English-speaking children – 3-year-olds (N=12, 9 boys, 3 girls) and 12-year-olds (N=20, 10 boys, 10 girls) – were selected from the AusKidTalk corpus. This selection of the corpus' youngest and oldest age groups provides an initial exploration of age-related change in cliticization. The mean age of the 3-year-old group was 3;7 years (Range: 3;2 years to 3;11 years). The mean age of 12-year-old children was 12;7 years (Range: 12;1 years to 12;11 years). They were all typically developing children without any language disorder according to their caregivers' report.

2.3. Materials

Two experimental conditions were used during the corpus collection (No-Clitic and Potential-Clitic) (see Table 1). The difference is that there is an article in Potential-Clitic condition. Figure 1 illustrates the hypothesized prosodic structures of two experimental conditions with two potential directions for the Potential-Clitic. Within each condition, three test sentences were employed, all starting with the phrase *boys often* and concluding with distinct combinations of verbs and nouns.

The total number of sentences for the 3-year-old group amounted to 72, while the 12-year-old group encompassed 120. Since young children produce repetition, substitution, disfluency, or omission in their speech, 29 sentences with such errors were excluded from the acoustic analysis. A total of 163 sentences were ultimately included for examination.

Table 1: *List of stimulus sentences.*

Prosodic structure	Utterance types
No-Clitic	Boys often hide gold Boys often cut cards Boys often cook beans
Potential-Clitic	Boys often hide the gold Boys often cut the card Boys often cook the beans

2.4. Data analysis

2.4.1. Measurements

The current study adopted the polysyllabic shortening measure, showcasing instances of verb shortening in adults under leftward cliticization [7], [8] and noun shortening under rightward cliticization [8].

Three primary durations were measured, namely the verb duration, the object noun duration, and the article duration. The subject noun duration was used in the analysis to account for differences in speech rate among participants and age groups. The acoustic annotation process was carried out using Praat [18], with a single annotator primarily responsible for annotating all acoustic landmarks within the utterances.

For the target verbs, object and subject nouns in each utterance, a boundary was marked between the onset and offset of sonorant segments. Boundaries were inserted at the onset of voicing after the release in a stop consonant to the offset of voicing before the release in a stop consonant (e.g., for /k/ to /t/ in "cut" and /k/ to /d/ in "cards") or before the closure in a fricative consonant (e.g., for /b/ to /s/ in "beans"). The article durations were marked between the onset and offset of schwa.

2.4.2. Statistical analysis

Three linear mixed-effects models with the verb, the object noun, and the article as dependent variables were separately constructed by using the lme4 package [19] in the R environment. All absolute durations (in milliseconds) were converted to natural logarithms, and subject noun durations were incorporated as a non-interacting continuous predictor. This normalization technique was done to absorb variance in verb, object and article durations due to age-related speech-rate differences caused by different abilities to control articulatory movement.

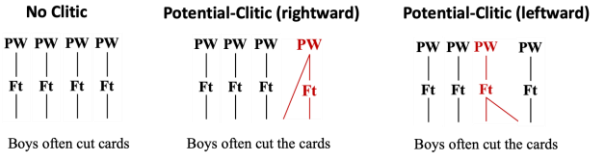

Figure 1: *Hypothesized prosodic structures for the two experimental conditions, with two potential directions for the Potential-Clitic.*

The Akaike information criterion (AIC) was used to evaluate the best-fit random structure for each model. The random-effects structure of the verb model incorporates by-participant random intercepts and by-participant random slopes for the following within-participant predictors: prosodic structure and verb type (with three levels, "cut", "cook" and "hide"). Verb type is a control predictor in our model. Similarly, the object model includes by-participant random intercepts and

by-participant random slopes for the following within-participant predictors: prosodic structure and subject duration. By-participant random slopes for the interaction were not included, because the model with these slopes failed to converge.

3. Results

3.1. The direction of cliticization

The results of verb durations, which could indicate leftward cliticization of articles, can be seen in Figure 2. Both the 3-year-olds and 12-year-olds exhibit longer durations in the No-Clitic condition compared to the Potential-Clitic condition across all verbs. The linear mixed effect model indicates that verb durations are between 1.17 and 1.36 times shorter in the Potential-Clitic condition than in the No-Clitic condition (point estimate 1.27 times; p from $1 = 4.7 \cdot 10^{-7}$) (with 95% CI). Thus, on average across both ages, children cliticize articles leftward, with the preceding verb.

There is also an age effect on verb duration: three-year-olds exhibit segment durations that are 1.05 to 1.38 times longer than 12-year-olds (point estimate 1.2 times; $p = 0.014$). This accords with earlier studies showing that younger children exhibit longer segment durations [20]. We found no significant interaction effect between age and prosodic structure on verb duration ($p = 0.24$).

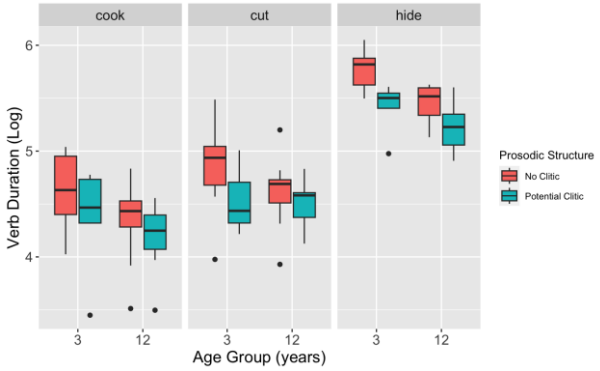

Figure 2: *The logarithm of verb duration across No-Clitic and Potential-Clitic conditions for 3-year-old and 12-year-old children, with separate plots for the verbs "cook", "cut", and "hide".*

The results of the object noun durations, which could indicate rightward cliticization, are presented in Figure 3. The object noun seems to be longer in the Potential-Clitic condition than in the No-Clitic condition in five of our six simple comparisons. However, the statistical analysis does not reveal a significant effect of prosodic structure: the object noun is between 0.94 and 1.12 times longer in the Potential-Clitic condition than in the No-Clitic condition (point estimate 1.03 times; $p = 0.52$). Therefore, we are unable to draw any definitive conclusions regarding cliticization within article–object phrases. The age and the interaction of age and prosodic structure on object duration show similar patterns to the effects on the verb, showing longer object durations for 3- than 12-year-olds by between 1.13 and 1.49 times (point estimate 1.3 times; $p = 0.00055$), and no evidence for the interaction ($p = 0.45$).

These results are consistent with leftward cliticization of articles, with no evidence of rightward cliticization nor of direction changing with age. Does this imply that they similarly implement all aspects of this linguistic phenomenon? In the next section, we further explore whether there are any age-related changes in the acoustic realization of cliticization, examining article duration under cliticization.

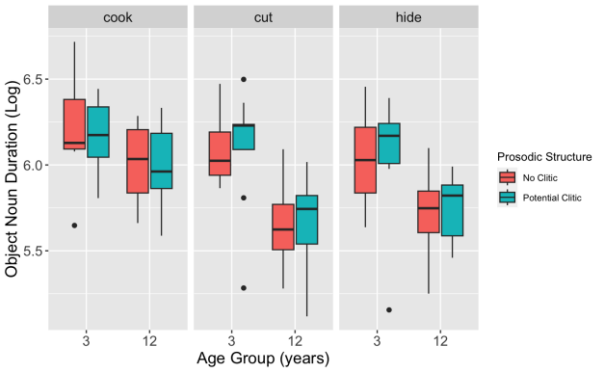

Figure 3: *The logarithm of object duration, with separate plots for the verbs "cook", "cut", and "hide".*

3.2. Age-related difference of acoustic realization

All sentences in the Potential-Clitic condition were further analyzed to answer our second research question regarding the relative duration of the article. Two linear mixed-effects models were employed in this analysis. The first model contained article durations as a dependent variable to test the possibility of article reduction. The second model employed the logarithm of the verb-to-article ratio as the dependent variable to examine the influence of age on the ratio between the durations of verbs and the durations of articles. The estimates from this model were converted into ratios (verb durations divided by article durations) by the exponential function. Note that a previous study [14] directly analyzed ratios (dividing content word by function word durations). Instead, we took the logarithm of verb-to-article ratios, which mitigates the impact of non-normality caused by using untransformed ratios as a direct dependent variable. Both models included age and verb type as fixed effects, with the random-effects structure limited to a by-participant random intercept only; no random slopes (for verb type) were included, because the model with these slopes failed to converge.

The "Article" column in Figure 4 illustrates that the 12-year-olds exhibited greater reduction of article duration than the 3-year-olds. The statistical analysis shows that, for 12-year-olds, the article is 1.16 to 1.99 times shorter than that of the 3-year-olds (point estimate 1.52 times; $p = 0.0033$). Therefore, we can conclude that 12-year-old children exhibit a greater reduction in articles compared to 3-year-olds.

The finding that 3-year-olds produce longer verbs and longer articles than 12-year-olds (over and above the speech-rate differences present in subject noun durations) (used as a control), could simply reflect uniformly longer durations of both, or a change in the temporal patterns within the verb-article unit. The graph in Figure 4 may also suggest a larger difference between verb and article durations in the 12-year-old than the 3-year-old children. Conversely, the more horizontal slopes for the 3-year-olds suggest more equal-syllable timing in the 3-year-olds than the 12-year-olds. The statistical analysis of the logarithm of the verb-to-article ratio indeed reveals a significant effect of age: the duration ratio between the verb and article is

1.04 to 1.76 times larger in 12-year-olds than in 3-year-olds (point estimate 1.35; p from 1 = 0.024).

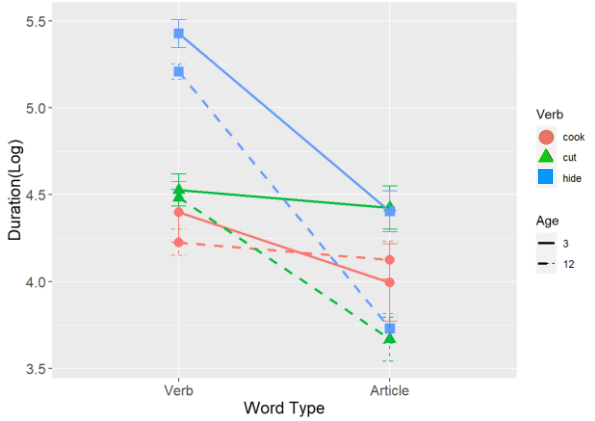

Figure 4: *Mean Durations and standard errors of verbs and articles across three verb conditions for 3-year-old and 12-year-old children.*

4. Discussion

The initial goal of this project was to identify the direction of cliticization in children's speech. The results have identified leftward cliticization for both 3-year-olds and 12-year-olds. This is consistent with the toddler studies by Gerken [5], [13] and Demuth & McCullough [12], as well as the adult study by Yuen et al. [7], but conflicts with Redford's claims of rightward cliticization in school-age children [14].

A possible explanation for the present lack of evidence for rightward cliticization is that the object-shortening effect, which was used to measure rightward cliticization, may have been confounded by the impact of phrase-final lengthening [21]. As phrase-final lengthening exists across both No-Clitic and Potential-Clitic conditions, one could perhaps argue that the comparison of these conditions should capture rightward cliticization beyond phrase-final lengthening effects. Another possibility is that the interplay between utterance-final lengthening and cliticization-induced shortening might differ from one child to another because they might assign different weights to these effects. Alternatively, it could be that task effects in [14] yielded these different results. Further work needs to be done with more reliable measurements to assess the rightward cliticization in children's speech.

These findings may help us to understand children's cliticization from the perspective of their phonological representations. The two directions of cliticization can be treated under two types of constraints: the metrical foot constraint governs leftward cliticization and the syntactic alignment constraint governs rightward cliticization. According to Turk & Shattuck-Hufnagel [8], the prosodic structure is influenced by the syntactic representation through different ranking of constraints. The finding of leftward cliticization in the present study suggests that the metrical foot constraint is ranked higher than the syntactic alignment constraint in 3- and 12-year-old children's representation. It is an interesting future direction to investigate the development of the ranking of cliticization-related constraints and identify when it stabilizes. For example, Demuth & McCullough [12] reported that different individual strategies were used in the production of articles in children's early language. Similar individual strategies could also be found in our data and account for variation in the rankings of children's constraints in the future.

Turning now to the second question on age-related change in the acoustic realization of cliticization, our findings align with most acquisition studies, finding that 3-year-old children produce fuller (i.e., less reduced) articles than 12-year-old children. The second major finding was that the ratio between the duration of the verb and the duration of the article was smaller in 3-year-olds than in 12-year-olds, which supports the idea that age-related differences in reduction may lead to age-related differences in how children trade off the verb and article within one functional unit. The result suggests that young children have relatively small timing differences between components within a functional unit. As they grow older, children's verb durations thus become relatively extended within the functional strong–weak or verb–article unit, while their article durations become increasingly reduced. This results in more durational prominence for content words than clitics. We propose to call this "a developing trading relationship" to characterize the dynamic interaction between verbs and articles in children's speech.

Taken together, the observed correlation between verbs and articles is consistent with other research that finds that younger children produce less reduction on cliticization than older children, and that this reduction particularly affects weak syllables [10], [14], [16].

The generalisability of these results is subject to a potential limitation of the experiment task, namely elicited imitation. Elicited imitation has been used in a well-established paradigm for studying cliticization acoustically [7], [8], [14], although Demuth & McCullough [12] conducted their acoustic case study on spontaneous speech in a corpus. As reviewed by Benders et al. [22], both elicited imitation and spontaneous speech tasks tap into similar aspects of children's phonological and phonetic processes. Most strongly related to the phenomenon of polysyllabic shortening, previous studies have demonstrated moraic structure effects on vowel production for 2-year-old children across imitated and spontaneous speech [23], [24]. This suggests that the present conclusion on children's acoustic realization of cliticization may similarly not be limited to elicited imitation tasks but generalize to spontaneously realized utterances.

5. Conclusion

This study aimed to investigate children's prosodic cliticization of grammatical function words. Drawing on the multisyllabic shortening paradigm employed by Yuen et al. [7], the results show that, on average, 3-year-old and 12-year-old children cliticize articles with content words (in this case: verbs) to their left. However, 3-year-olds did not reduce articles during cliticization to the same extent as their 12-year-old counterparts. Our findings suggest a trade relationship between verbs and articles in the process of children's acquisition. Future research will need to learn about the development of prosodic cliticization in children's speech production.

6. Acknowledgements

This work is supported by the China Scholarship Council.

7. References

[1] M. Kehoe, "The development of prosody and prosodic structure," in *Languages and Linguistics*. New York: Nova Publishers, 2013, pp.119–162.

[2] M. Nespor and I. Vogel, *Prosodic Phonology: With a New Foreword*, De Gruyter Mouton, 1986.

[3] J. J. McCarthy and A. Prince, "Generalized alignment," in *Yearbook of Morphology 1993*, G. Booij and J. V. Marle, Eds., Dordrecht: Springer Netherlands, 1993, pp. 79–153.

[4] E. Selkirk, "The prosodic structure of function words," in *Signal to Syntax*, Psychology Press, 2014, pp. 199–226.

[5] L. Gerken, "Young children′s representation of prosodic phonology: evidence from English-speakers′ weak syllable productions," *Journal of Memory and Language*, vol. 33, no. 1, pp. 19–38, Feb. 1994.

[6] B. Hayes, "Extrametricality and English stress," *Linguistic Inquiry*, vol. 13, no. 2, pp. 227–276, 1982.

[7] I. Yuen, K. Demuth, and S. Shattuck-Hufnagel, "Planning of prosodic clitics in Australian English," *Language, Cognition and Neuroscience*, vol. 37, no. 10, pp. 1271–1276, Apr. 2022.

[8] A. E. Turk and S. Shattuck-Hufnagel, "Word-boundary-related duration patterns in English," *Journal of Phonetics*, vol. 28, no. 4, pp. 397–440, Oct. 2000.

[9] A. Alkhudidi, *An Acoustic Investigation of Traces to Children's Early Omitted English Articles: A Case Study*, Dissertation, Macquarie University, 2022.

[10] G. D. Allen and S. Hawkins, "Phonological rhythm: definition and development," in *Child Phonology*, G. H. Yeni-Komshian, J. F. Kavanagh, and C. A. Ferguson, Eds., Academic Press, 1980, pp. 227–256.

[11] K. Demuth, "Markedness and the development of prosodic structure," in *Proceedings of the North East Linguistics Society*, Amherst, America, 1995, pp.13–25.

[12] K. Demuth and E. McCullough, "The prosodic (re)organization of children's early English articles," *Journal of Child Language*, vol. 36, no. 1, pp. 173–200, Jan. 2009.

[13] L. Gerken, "Prosodic structure in young children's language production," *Language*, vol. 72, no. 4, pp. 683–712, 1996.

[14] M. A. Redford, "Grammatical word production across metrical contexts in school-aged children's and adults' speech," *Journal of Speech, Language, and Hearing Research*, vol. 61, no. 6, pp. 1339–1354, Jun. 2018.

[15] D. H. Whalen, "Coarticulation is largely planned," *Journal of Phonetics*, vol. 18, no. 1, pp. 3–35, Jan. 1990.

[16] M. A. Redford and P. J. Howson, "Acoustic correlates and listener ratings of function word reduction in child versus adult speech," *The Journal of the Acoustical Society of America*, vol. 152, no. 3, pp. 1463–1475, Sep. 2022.

[17] B. Ahmed, K. Ballard, D. Burnham, T. Sirojan, H. Mehmood, D. Estival, E. Baker, F. Cox, J. Arciuli, T. Benders, K. Demuth, B. Kelly, C. Diskin-Holdaway, M. Shahin, V. Sethu, J. Epps, C. B. Lee, E. Ambikairajah, "AusKidTalk: an auditory-visual corpus of 3- to 12-year-old Australian children's speech," in *Interspeech 2021*, ISCA, pp. 3680–3684, Aug. 2021.

[18] P. Boersma and D. Weenink, "Praat: doing phonetics by computer," Version 6.1.09, retrieved 26 January 2020 from http://www.praat.org/.

[19] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, 67 (1), 48, https://doi. org/10, 2015.

[20] I. Yuen, F. Cox, and K. Demuth, "Three-year-olds' production of Australian English phonemic vowel length as a function of prosodic context," *The Journal of the Acoustical Society of America*, vol. 135, no. 3, pp. 1469–1479, Mar. 2014.

[21] D. Snow, "Phrase-final syllable lengthening and intonation in early child speech," *Journal of Speech, Language, and Hearing Research*, vol. 37, no. 4, pp. 831–840, Aug. 1994.

[22] T. Benders, N. X. Rattanasone, R. Thornton, I. Mulders, and L. Koring, "Experimental methods to study child language," in *The Routledge Handbook of Experimental Linguistics*, Routledge, 2023, pp. 375–389.

[23] K. Demuth, J. Culbertson, J. Alter, "Word-minimality, epenthesis and coda licensing in the early acquisition of English," *Language and Speech*, vol. 49, no. 2, pp. 137–173, Jun. 2006.

[24] K. Miles, I. Yuen, F. Cox, and K. Demuth, "The prosodic licensing of coda consonants in early speech: interaction with vowel length," *Journal of Child Language*, vol. 43, no. 2, pp. 265–283, May. 2016.