Why Children Omit Function Morphemes: Metric vs. Syntactic Structure

Misha Becker
UCLA

1. Introduction

It is an interesting and noticeable fact that in the early stages of language production, children omit certain elements from their speech. Upon closer inspection, it becomes clear that children regularly omit certain function morphemes, as opposed to content words. Function morphemes (including determiners, pronouns, auxiliary verbs, and inflection morphemes) form a natural class from both a phonological and a syntactic perspective: phonologically speaking, in many languages (including English), function morphemes tend to be unstressed and monosyllabic; syntactically speaking, function morphemes are the heads of functional projections, which appear notoriously late in children's speech (Radford 1988). Thus, we can ask whether children omit functional elements for phonological or syntactic reasons.

If children omit function morphemes for reasons related to metrical stress, we should find similarities between unstressed function morphemes and unstressed lexical syllables, and, depending on our assumptions about metric structure, we might find asymmetries in the omission rates of subject vs. object determiners.[1] On the other hand, if children omit function morphemes for syntactic reasons, according to some accounts (e.g. Hoekstra, Hyams & Becker 1997, Clahsen et al. 1996) we should find correlations between omissions in the subject and the predicate, following the structural contingency known as spec-head agreement. In fact we find both of these patterns, which suggests that both metrical and syntactic processes constrain children's output forms, but in different ways.

1.1 Overview of predictions

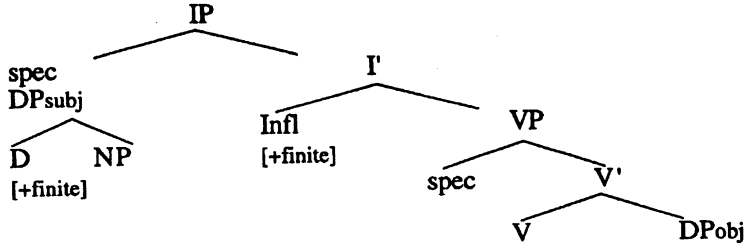
Many researchers have noticed the following pattern of syllable omission in young children's speech: while children show an overall preference for preserving strong (stressed) syllables, they also tend to preserve weak syllables in trochaic feet (S-w) more than weak syllables that are either in an iambic foot or unfooted. This fact has led a number of linguists to suggest that children's output forms are constrained by a so-called "trochaic template", such that weak syllables in trochaic feet are favored over all other weak syllables (Allen & Hawkins 1980, Wijnen, Krikhaar & den Os 1994, Gerken et al. 1990, Gerken 1991, 1994a, 1994b, 1996, i.a.).

To illustrate the predicted pattern, the determiner in (1a) should be omitted more frequently than the pronoun in the same sentence, because the subject determiner, unlike the object pronoun, fails to form a trochaic foot with anything.[2] Comparing (1b) and (1c), the object determiner in (1c) should be omitted more frequently than the one in (1b), because the syllabic inflectional morpheme on the verb in (1c) (namely *-es*) forms a trochaic foot with the verb root, forcing the determiner to form an iambic foot with the object noun (*duck*). (Stress is indicated by all caps.)

(1) a. [The BOY] [HUGGED him]
b. [BERT] [PUSHED the] [DUCK]
c. [BERT] [PUSHes] [the DUCK]

A syntactic account of children's speech also predicts a certain pattern of function morpheme omissions. Within recent generative theories, the subject is represented in the specifier position of the projection whose head is the Infl (inflection) node, as shown in (2).

(2)



Based on the idea that a specifier and its head should "agree" in features (known as spec-head agreement), we would expect the subject and the verb to agree in their specification (or unspecification) of, for example, finiteness features. A proposal by Hoekstra, Hyams & Becker (1997) suggests, specifically, that the omission of a subject pronoun (i.e. a null subject) or a missing determiner (a bare N subject) is indicative of a lack of "nominal finiteness" in the subject, and should thus correspond to a lack of finiteness in the verb (as evidenced through a lack of verbal inflectional morphemes, such as 3sg *-s*). In brief, the expected correlation hinges on finiteness features in the verb (shown in (3)): if the verb is finite, the subject should be fully specified; if the verb is nonfinite, we might expect either a specified or an unspecified subject. For a detailed discussion of "nominal finiteness" see Hoekstra & Hyams (1996).

(3) a. finite V -> full DP subject, *null subject/bare N
b. nonfinite V -> full DP subject, null subject/bare N[3]

2. Object determiner and pronoun omissions

Let us look at how the metrical and syntactic approaches square with children's omissions of object pronouns and determiners.

2.1 Gerken's data

In children's imitative reproductions of sentences of the sort in (1), Gerken (1991, 1994, 1996) found that children in fact omitted an object pronoun or determiner less often if it occurred in a trochaic foot. In example (4) I give some of the results from a study by Gerken (1991, average age 2;3), where the percentages below the unstressed words indicate the rate of omission:

(4) a. [she KISSED] [the DOG]
39% 28%
b. [the DOG] [KISSED her]
39% 0%
c. [PETE] [KISSED the] [DOG]
11%

As we see in (4), the object pronoun in (4b) is never omitted, and the object determiner in (4c) is omitted only 11% of the time; both of these words occur in trochaic feet. In contrast, the object determiner in (4a) is omitted much more frequently: at 28% of the time.

More specifically, Gerken (1996) shows that children (age 2;1-2;3) display a contrast in omission frequencies of the object determiner in the near minimal pair given in (5):

(5) a. he [KICKS the] PIG
16%
b. he [CATCHes] [the PIG]
29% 48%

As predicted by Gerken's metrical hypothesis, the children in her experiment omitted the object determiner in pattern (5b) more frequently (48% of the time) than the one in pattern (5a) (only 16%). Note that there is no syntactic difference between the two verbs, so no syntactic account would predict a difference in the omission rates between the determiners in the two conditions.

2.2 Progressive verbs

Note that monosyllabic verbs with progressive aspect (e.g. *running*) constitute a trochaic foot by themselves.

(6) [RUNning]
S w

Thus, given the facts outlined above for the omission of object determiners and pronouns following finite verbs, we should expect that object determiners & pronouns following progressives should be omitted as frequently as in (5b), as opposed to (5a). The expected omissions in this condition, according to the metrical hypothesis, are given in (7), where the morphemes with high expected rates of omission are underlined:

(7) a. [<u>The</u> BOY] <u>is</u> [HUGging] [<u>the</u> DUCK].
b. [<u>The</u> BOY] <u>is</u> [HUGging] <u>him</u>.

I searched 12 files of spontaneous (natural) speech from the Adam corpus (Brown 1973, from the CHILDES database, MacWhinney & Snow 1985), in which Adam's age was 2;3 to 3;7 (years;months). I found a total of only 18 examples of sentences in which Adam produced a progressive verb with an unambiguous singular count noun object (e.g. *eating (an) apple*). Of those 18 sentences, 13 (or 72%) contained an overt object determiner or pronoun, and only 5 (28%) lacked an object determiner or pronoun. I show the results in Table 1, comparing this result

with his omissions of object determiners or pronouns following plain, monosyllabic verbs. I give example sentences for each category.

Table 1. Adam's omission of object determiner/pronoun by verb type (age 2;3-3;7)

object type	progressives: n (%)	plain Vs: n (%)
overt det/prn	13 (72%) dey are holding it Mommy making a road	79 (85%) Adam do it dat fits my train
null det/prn	5 (28%) dis man is getting __ doggie getting __ meat	14 (15%) baby have __ tooth Adam see __ sun

$\chi^2<1$

Table 1 shows that Adam omits object determiners and pronouns after progressive verb forms at a higher rate (28%) than after plain, monosyllabic verbs (15%) (e.g. *eat(s) (an) apple*), as would be predicted by the metrical account.

2.3 A finiteness effect

Although the predictions of the metrical hypothesis for object omission are confirmed by these data, the metrical approach does not give us the whole story. In particular, there is a finiteness effect on Adam's omissions of object determiners and pronouns. That is, while 14 (25%) of Adam's sentences containing nonfinite plain verbs lacked an object determiner or pronoun, none of his finite utterances did. The breakdown by finiteness is given in Table 2.

Table 2: Adam's object det/prn omissions for finite vs. nonfinite verbs (age 2;3-3;7)[4]

object type	finite V	nonfinite V
overt det/prn	38 (100%) he takes it	41 (75%) kitty eat_ a apple all up
null det/prn	0 (0%)	14 (25%) have __ mouth

$\chi^2= 9.48$, $p \leq .005$

The same pattern of distribution with respect to finiteness occurs with Adam's progressive verbs, as shown in Table 3.[5]

Table 3: Adam's object det/prn omissions for finite vs. nonfinite progressives (age 2;3-3;7)[6]

object type	finite progressive (is)	nonfin. progressive (0-is)
overt det/prn	4 (100%) dey are holding it	9 (64%) Mommy _ making a road
null det/prn	0 (0%)	5 (36%) Adam _ taking _ diamond

$\chi^2 < 1$

Another child, Nina (Suppes 1973), shows a similar pattern of object determiner and pronoun omission to that of Adam. Searching 20 files, in which Nina ranged in age from 2;0 to 2;6, Nina omitted an object determiner or pronoun

after a plain, monosyllabic verb about 10% of the time, but about 20% of the time following progressive verbs.[7] I give the results in Table 4:

Table 4: Nina's omissions of object determiner & pronouns by verb type

	progressives	plain verbs
overt object det/prn	140 (80%)	1089 (90%)
null object det/prn	37 (20%)	115 (10%)

$\chi^2 = 19.16$, $p \leq .001$

Nina shows a slight finiteness effect, but not a significant one.

Nina's and Adam's data came from spontaneous, natural speech samples, while Gerken's results come from experiments involving elicited speech (in particular, imitative speech), making comparisons difficult. In order to better compare these sets of data, I conducted a small pilot study, adopting Gerken's methods. I measured omission rates for object determiners and pronouns following progressive verbs. The study involved 8 native English-speaking children between ages 2;0 and 2;6. Each child was tested on a total of 36 different sentences, broken down into two sets with two randomizations of each set. Each sentence was recorded onto a tape, spoken with normal intonation and stress, by a native speaker of English who was blind to the task.

I played the tape for each child individually, stopping the tape after each sentence. I then asked the child to repeat exactly what he or she had heard on the tape (a sample was given beforehand to ensure that the child understood the task). The sentence was replayed up to 3 times if the child gave no response, and if there was still no response after 3 repetitions, the next sentence was given. All sessions were recorded on another tape, and after each session, these tapes were checked against the results recorded during the actual experiment.

The predicates of the test sentences and the results are given in Table 5. I compared productions of utterances containing the verbs *hug* and *push*: These verbs differ in their metric structure in present tense (*hugs* vs. *pushes*), they are both monosyllabic in past tense (*hugged, pushed*), and they are both disyllabic in progressive aspect (*hugging, pushing*). Although the total number of omissions I recorded was quite small, the trend followed the pattern that Gerken reports. Object determiners and pronouns following progressives were omitted at roughly the same rate as the object determiners and pronouns following plain verbs with syllabic inflection, and more frequently than object determiners and pronouns after finite verbs with nonsyllabic inflection.

Table 5: object det/prn omissions following di- vs. monosyllabic verbs (age 2;0-2;6)

[V+es/ing] + the/him	n	%	[V+the/him]	n	%
[PUSHes] the	2	10%	[PUSHED the]	0	0%
[PUSHes] him	1	5%	[PUSHED him]	1	5%
[PUSHing] the	2	10%	[HUGGED the]	1	5%
[PUSHing] him	1	5%	[HUGGED him]	0	0%
[HUGging] the	1	5%	[HUGS the]	0	0%
[HUGging] him	1	5%	[HUGS him]	0	0%

As shown in Table 5, object determiners and pronouns following disyllabic verbs were omitted slightly more frequently than those following monosyllabic verbs (e.g. *pushes the* vs. *pushed the*). There was one exception involving an object pronoun: *pushes him* vs. *pushed him.*

As for a connection between verb finiteness and object determiner or pronoun omission, we find a similar correlation with finiteness as was found in Adam's natural speech data. That is, the children in my pilot study produced sentences with null object determiners and pronouns both with finite and nonfinite verbs (i.e., they produced a nonfinite verb if they omitted the verbal inflectional morpheme in the utterance they repeated). However, a much higher percentage of finite utterances contained an overt object determiner or pronoun (97%) than nonfinite utterances (86%). I give the relevant figures in Table 6.

Table 6: finiteness and object det/prn omission: data from pilot study (2;0-2;6)

	finite verbs	nonfinite verbs
overt object det/prn	145 (97%)	38 (86%)
null object det/prn	4 (3%)	6 (14%)

$\chi^2 = 6.21, p \leq .05$

To summarize this section, Gerken's experimental data suggest that children preserve weak syllables at a higher rate if they occur in trochaic feet, and the results of my small pilot study appear to support her claims. However both the spontaneous speech data I studied and my experimental data suggest a possible finiteness effect. Thus a syntactic correlation between verb finiteness and omission of object dets/prns cannot be ruled out.

3. Subject determiner and pronoun omissions

Let us turn now to the predictions regarding omission patterns of function morphemes in the subject. Both the metrical approach and the syntactic approach described above predict that subject determiners and pronouns should be omitted, but the precise predictions each theory makes are different. The metrical view makes the straightforward prediction that subject determiners and pronouns should be omitted relatively frequently, since they will not form a trochaic foot with a stressed syllable, as we saw earlier in (cf. 1a, 2a-b). Gerken (1991, 1994) provides experimental support for this hypothesis. Gerken 1991 showed that subject determiner and pronoun omissions ranged from 25-39%. In contrast, stressed syllables in the subject were omitted less requently: 3-25% of the time. In (8) below, I give example sentences from her 1991 experiment; percentages indicate rates of omission (average age 2;3).

(8) a. He kissed her.
35% 0%
b. He kissed John.
33% 0%
c. He kissed the dog.
39% 28% 0%
d. The boy kissed her.
39% 3% 0%

e. The boy kissed John.
25% 25% 0%
f. The boy kissed the dog.
28% 11% 14% 0%

The syntactic account proposed in Hoekstra, Hyams & Becker (1997) makes a more specific prediction. Namely, at a stage at which children have a tendency to omit subjects and verb inflection, we expect to find a contingency between omissions of the subject or subject determiner, and finiteness in the verb. That is, when the verb is finite, the subject should be fully specified (an overt pronoun or determiner); if the verb is nonfinite, the subject may be specified or unspecified (either null or lacking a determiner). Concretely, we should see examples of (9a, b & c) in children's speech, but not (9d).

(9) a. The boy dances
b. Ø boy dance/ Ø dance
c. The boy dance
d. *Ø boy dances

As table 7 shows with data from Nina (from the CHILDES database), this prediction is borne out. While Nina's finite verbs almost always take a fully specified subject (i.e. one with an overt determiner), her nonfinite verbs are more evenly distributed with respect to occurrence with overt and null subject determiners.

Table 7: Nina's production of subject determiners and verb finiteness (2;4-2;10)

	Finite verbs	Nonfinite verbs
overt determiner	34 (92%)	12 (57%)
null determiner	3 (8%)	9 (43%)

$\chi^2 = 7.85, p \leq .01$

As we can see in Table 7, there is a strong correlation between overt determiners and finite verbs. Adam's data show a slightly different pattern, but one that is still consistent with the syntactic hypothesis. As shown in Table 8, 93% of Adam's finite verbs take an overt subject determiner, but almost none of his nonfinite verbs do.

Table 8. Adam's production of subject determiners and verb finiteness[8] (2;3-3;7)

	Finite Verbs	Nonfinite Verbs
Overt determiner	53 (93%)	2 (5%)
Null determiner	4 (7%)	39 (95%)

$\chi^2 = 71.64, p \leq 0.001$

As for subject pronouns, 70% of Adam's utterances with a pronominal subject were finite (30% were nonfinite), and of Adam's nonfinite utterances, only 29% of them had an overt pronominal subject. (That is, Adam produced many utterances of the type *He is sleeping/he sleeps* and very few utterances of the type *He sleeping/he sleep.*) This fact about Adam's speech shows that the same nonrandom cooccurrence pattern between subject determiners and verb finiteness also holds

true for pronominal subjects. The metrical hypothesis does not predict such a correlation.

3.1 Subject-finiteness correlation from pilot data

The results of my pilot study also confirm the predictions of the syntactic account (i.e. there is a subject-verb correlation for finiteness). The data are given in Table 9.

Table 9: subject-verb finiteness correlation from pilot study (2;0-2;6)

	finite verb (pushed, pushes, hugged, hugs, is)	nonfinite verb (push, hug, null-is)
overt subject / subj det (Bert, the boy, he)	135 (85%)	20 (53%)
null subject / subj det (0, boy)	25 (15%)	18 (47%)

$\chi^2 = 17.17$, $p \leq 0.0001$

Within finite clauses, there is a strong tendency also to have an overt subject pronoun or determiner (85%). However when children failed to produce a finite verb, there was roughly an even number of productions containing as lacking a subject or subject det (53% and 47%, respectively). Note that this is the same pattern we found in Nina's speech (cf. Table 7), and is the pattern predicted by the syntactic account of Hoekstra, Hyams & Becker (1997).

4. Universality

A further potential advantage of the syntactic account is its universality. We find the same pattern of correlated finiteness between the subject and verb in other child languages such as German, Dutch and French. Some child data from German are given in Table 10.

Table 10: Subject determiners and finiteness in German children (avg. age 2;6)[9]:

	finite verb	nonfinite verb
overt subject det.	9 (90%)	2 (15%)
null subject det.	1 (10%)	11 (85%)

$\chi^2 = 9.798$, $p \leq 0.005$

Furthermore, evidence from child Italian (Schaeffer 1996) suggests that children inflect past participles to agree with the object if their object clitic is overt. When the object clitic is null, they do not produce agreement inflection on the participle. Whether the participle shows agreement or not, it has the same metric structure. The sentences in (10a-b) show the adult pattern of agreement. (11a) shows children's pattern with overt clitics, (11b) shows their pattern with null clitics, and (11c) shows that they do not inflect with null clitics. (The examples in (10-11) are schematic examples of the pattern with transitive verbs.)

(10) a. le ho viste
3pl.obj.cl. have-1sg. seen-3pl.
"I have seen them."
b. la ho vista
3sg.fem.obj.cl. have-1sg. seen-3sg.fem.
"I have seen her."

(11) a. le ho viste (same as 10a)
b. ho visto *pro* (meaning = (11a))
have-1sg. seen-1sg.
c. *ho viste/vista

Although these data do not run directly counter to Gerken's hypothesis, since she does not claim that all grammatical omissions are necessarily due to metric conditions (Gerken 1991), they show that there is a syntactic effect on object omissions that is not also a metric effect. That is, it is not the case that the omission of object clitics in child Italian is correlated with a metric effect on other words in the utterance, but rather with syntactic agreement inflection that does not bear on the metric structure of the sentence.

5. Data collection

Finally, let me point out some problems and advantages of using the two methods of data collection I used here. The main difficulty of using spontaneous speech, especially when looking for object omissions, is that in many cases it is difficult to tell whether an object was in fact omitted, i.e. whether it is obligatory. Particularly in the case of progressive verbs, many normally transitive verbs (e.g. *pushing*) can be used without an object in progressive aspect, taking a meaning like "push habitually". In this respect, using imitative speech is helpful since it is completely clear when a given morpheme is omitted.

On the other hand, it seems quite plausible to me that cadence and metrical structure play a much more important role in imitative than spontaneous speech. If this is true, then imitation tasks may well tell us about how metrical structure constrains a certain type of child speech, but not necessarily children's natural speech.

6. Conclusion

The main point of this paper was to explore the possibility that children's pattern of function morpheme omissions might be explained entirely by either a metrical or a syntactic story. There is strong support for both approaches, i.e. children preserve function morphemes that occur in trochaic feet, and they preserve determiners and pronouns in the subject and object when their verb is finite. Nevertheless, we have seen that there are cases like (5), which are explainable by a metrical story but not by a syntactic one, and the finiteness patterns in Tables 2 and 3, and 6-10, which are explainable by a syntactic story, but not by a metrical one. Thus, as the metrical and syntactic theories stand, neither one can fully cover the range of omission patterns we find in children's speech. In the future, we might find a way of refining and/or expanding one of the two approaches such that it captures the full range of phenomena, or we might discover a new approach that takes both metrical and syntactic phenomena into account.

Notes:

[1]This view is supported in Gerken 1991, 1994a, 1994b, 1996, Wijnen, Krikhaar & den Os, 1994, Fikkert 1994, among others.

[2]Brackets indicate foot boundaries; feet are formed according to the principles outlined in Gerken (1996).

[3]There is no direct prediction for omissions in the object, given finiteness features in the verb.

[4]Finite verbs that took a syllabic inflectional morpheme (/es/) were excluded from this data set, since the metrical hypothesis predicts that object determiners/pronouns following them should omitted more frequently. There were not enough productions of verbs taking or requiring syllabic inflection to do a comparison strictly between them and verbs not requiring syllabic inflection.

[5]A progressive verb is considered finite if the auxiliary verb *is* is overt; it is considered nonfinite if the auxiliary is omitted. This analysis is supported by considerations discussed in Hoekstra & Hyams, forthcoming.

[6]Please note that while the object omissions after progressive verbs are all predicted by the metrical approach, regardless of finiteness, the correlation between verb finiteness in plain verbs and object omissions cannot be a result of an accidental overlap of metrical & syntactic predictions. In fact, the metrical hypothes predicts the sequence [push the/him] to have a higher rate of preservation for the object determiner or pronoun than its finite counterpart; however as table 2 shows, the nonfinite verbs had a lower rate of preserved object determiners and pronouns.

[7]I am very thankful to John Grinstead (UCLA) for supplying me with the relevant utterances from the Nina corpus.

[8]Table 2 is taken from Hoekstra, Hyams & Becker (1997).

[9]Table 10 is reproduced from Hoekstra, Hyams & Becker (1997).

Bibliography

Allen, George & Sarah Hawkins. 1980. Phonological rhythm: Definition and development. Child Phonology, vol. 1: Production. ed. by Yeni-Komshian, Kavanagh & Ferguson. New York: Academic Press. 227-256.

Brown, Roger. 1973. A first language. Cambridge, MA: Harvard University Press.

Clahsen, Harald, Sonja Eisenbeiss & Martina Penke. 1996. Lexical Learning in Early Syntactic Development. Generative Perspectives on Language Acquisition. ed. by Harald Clahsen. Philadelphia: John Benjamins.

Fikkert, Paula. 1994. On the acquisition of prosodic structure. Ph.D. dissertation. Dordrecht: HIL.

Gerken, LouAnn, Barbara Landau and Robert Remez. 1990. Function morphemes in young children's speech perception and production. Developmental Psychology 27. 204-216.

Gerken, LouAnn. 1991. The metrical basis of children's subjectless sentences. Journal of Memory and Language, vol. 30. 431-451.

Gerken, LouAnn. 1994a. A metrical template account of children's weak syllable omissions. Journal of Child Language, vol. 21. 565-584.

Gerken, LouAnn. 1994b. Young children's representation of prosodic structure: Evidence from English-speakers' weak syllable omissions. Journal of Memory and Language, vol. 33. 19-38.

Gerken, LouAnn. 1996. Prosodic Structure in Young Children's Language Production. Language 72. 683-712.

Hoekstra, Teun & Nina Hyams. 1996. The syntax and interpretation of dropped categories in child language: a unified account. Proceedings of WCCFL XIV. Stanford: CSLI.

Hoekstra, Teun & Nina Hyams (forthcoming) Agreement and the finiteness of V2: Evidence from child language. Procedings of the 22nd annual BUCLD, Brookline: Cascadilla Press.

Hoekstra, Teun, Nina Hyams & Misha Becker. 1997. The Underspecification of Number and the Licensing of Root Infinitives. Proceedings of the 21st annual BUCLD. Brookline: Cascadilla Press.

MacWhinney, Brian & Catherine Snow. 1985. The Child Language Data Exchange System. Journal of Child Language, vol. 12. 271-296.

Radford, Andrew. 1988. Small children's small clauses. Transactions of the Philological Society, vol. 86. 1-43.

Schaeffer, Jeannette. 1996. Direct object scrambling in Dutch and Italian child language. UCLA Ph.D. dissertation.

Suppes, Patrick. 1973. The semantics of children's languages. American Psychologist, vol. 88. 103-114.

Wijnen, Frank, Evelien Krikhaar & Els den Os. 1994. The (non)realization of unstressed elements in children's utterances: A rhythmic constraint? Journal of Child Language, vol. 21. 59-83.