

**Proceedings of the
GALA '97 Conference on
Language Acquisition**

Edited by

Antonella Sorace
Caroline Heycock
Richard Shillcock

Edinburgh
4-6 April 1997

A Discourse-Pragmatic Explanation for the Subject-Object Asymmetry in Early Null Arguments: The Principle of Informativeness Revisited

Shanley E. M. Allen

Max Planck Institute for Psycholinguistics
Postbus 310, 6500 AH Nijmegen, Netherlands
allen@mpi.nl

Abstract

The present paper assesses discourse-pragmatic factors as a potential explanation for the subject-object asymmetry in early child language. It identifies a set of factors which characterize typical situations of informativeness (Greenfield & Smith, 1976), and uses these factors to identify informative arguments in data from four children aged 2;0 through 3;6 learning Inuktitut as a first language. In addition, it assesses the extent of the links between features of informativeness on one hand and lexical vs. null and subject vs. object arguments on the other. Results suggest that a pragmatics account of the subject-object asymmetry can be upheld to a greater extent than previous research indicates, and that several of the factors characterizing informativeness are good indicators of those arguments which tend to be omitted in early child language.

1. Introduction

Children at the early stages of language learning (two-word stage to 4;0+, peak before 2;6 in English) tend to omit arguments in their utterances, regardless of whether the target language permits this or not. Examples below from English, which does not normally permit null arguments, show utterances with null subject (1a), null object (1b), and both null subject and null object (1c) (data from Bloom, 1970, 1973).

- (1) a. No like celery. (Kathryn 1;10)
(telling her mother she doesn't like celery)
- b. Mommy, you wiping. (Allison 1;8)
(wanting her mother to wipe a doll)
- c. Spill. (Allison 1;10)
(noting that her mother has spilled some juice)

It has frequently been noted in the literature that a clear asymmetry exists between omission of subjects and objects, such that subjects are omitted more frequently than objects in child speech across several languages. This phenomenon occurs regardless of whether the target language uniformly requires overt arguments as in English (L. Bloom, 1970; Hyams, 1986; P. Bloom, 1990; Valian, 1991; Wang *et al.*, 1992), or permits null arguments in topic-drop contexts as in German (Hamann, 1995), or uniformly permits null arguments as in Chinese (Wang *et al.*, 1992), Japanese (Mazuka *et al.*, 1986; Hirakawa, 1992), Korean (Clancy, 1993), and Mauritian Creole (Adone,

1994) (though note that Dutch, a topic-drop language, does not seem to show a subject-object asymmetry (de Haan & Tuijnman, 1988; Krämer, 1995)).

Three major types of explanations have been put forth to explain the subject-object asymmetry: grammatical, performance and pragmatics accounts. Grammatical accounts assume something about the child's early grammatical structure is non-adult-like and thus causes the child to omit subjects more frequently than objects. Hyams's (1986) prodrop parameter account claims that all children are born with the prodrop parameter in a default setting permitting the omission of subjects, as in Italian and Spanish, but not affecting objects. Radford's (1990) functional category account claims that children at the early stages of language development only have lexical categories (e.g. NP, VP), and not functional categories (e.g. IP, CP). Thus, children cannot produce the subject which appears in the (functional) SPEC of IP position, but can produce the object which appears in (lexical) daughter of VP. Topic drop accounts (de Haan & Tuijnman, 1988; Hamann, 1995; Jaeggli & Hyams, 1988; Wang *et al.*, 1992; Hyams & Wexler, 1993) assume that early null arguments are a result of topic drop, a mistaken parameter setting for English, which tends to affect subjects more than objects.

Performance accounts typically assume that the child has adult-like grammatical structure from the earliest stages of language learning, but has difficulty in producing all that he or she is capable of due to performance limitations. The most common of these involves processing (L. Bloom, 1970; Mazuka *et al.*, 1986; P. Bloom, 1990; Valian, 1991). Since the processing load of a sentence is assumed to be greater at the beginning of the sentence where the subject tends to be, than at the end of the sentence where the object tends to be, the subject is more often dropped than the object. A second type of performance account involves prosody and metrical structure (Gerken, 1990). By this account, the subject tends to appear in a prosodically and metrically more weak position than the object, and thus is omitted more often by children.

Finally, pragmatics accounts assume that the child's non-adult-like production of null arguments stems from a hypersensitivity to pragmatic features of discourse. Greenfield & Smith's (1976) Principle of Informativeness suggests that children omit those arguments which are less "informative", and that the subject tends to be less informative than the object. Similarly, various

authors suggest that the subject more often tends to represent given information while the object tends to represent new information, and therefore the subject is omitted more often than the object (Hyams, 1986; Bloom, 1990; Valian, 1991; Hirakawa, 1993).

While much has been written about the various grammatical and performance accounts explaining the subject-object asymmetry in early null arguments, relatively little attention has been paid to pragmatics accounts. This paper, then, focuses on the viability of a pragmatics account to explain this phenomenon.

2. History of the Pragmatics Account

A pragmatics account of null arguments assumes that arguments are omitted from early child speech for various pragmatic reasons. While several researchers mention the distinction between given and new referents as a reason for null arguments, Greenfield & Smith's (1976) Principle of Informativeness (different from the Gricean Principle of Informativeness), is rather more explicit. Working on the premise that "child language is sensitive to the informational structure of an event" (p. 222), Greenfield & Smith note that children tend to encode those aspects of the event that are most informative: "where [they see] alternatives, where there is uncertainty in terms of the situational structure" (p. 64). On the other hand, children will tend not to encode those aspects that are presupposed, such as the subject. They state that the subject is "often taken for granted" and typically "less informative than other constituents of the sentence that resist 'deletion'" (p. 223), while the agent "is only spoken under unusual conditions, such as conflict about agency or an actual change of agent" (p. 184).

Few detailed studies of the effects of pragmatic factors in explaining the subject-object asymmetry have appeared to date in the literature. Hyams (1986) quickly dismisses the potential effect of pragmatic factors, reasoning that in most child speech the subject and object are equally uninformative and recoverable since "early language centers largely on objects and events in the immediate environment" (p. 97), and thus the Principle of Informativeness does not sufficiently distinguish between subject and object in determining which arguments are omitted. Hyams & Wexler (1993) also dismiss pragmatic factors since they find that non-lexical arguments (which they assume represent given referents) are most often pronominalized when they occur in object position, but dropped or pronominalized in about equal proportions when they occur in subject position. While the Principle of Informativeness would predict an equal proportion of pronominal to null arguments in both subject and object position, a grammatical model would predict the data found since the option to drop non-lexical arguments is available only for subjects. Thus, Hyams (1986) and Hyams & Wexler (1993) both conclude that while pragmatic factors may well be involved in explaining the null argument phenomenon, the need for a principled

grammatical account is not vitiated.

While Hyams's and Wexler's conclusion that a pragmatic account cannot provide the full explanation for the subject-object asymmetry may certainly be true, it is not clear that they have given this account a fair trial. The reasoning in Hyams (1986) is based on assumptions which are not empirically proven, since she does not examine any concrete data set to actually determine which arguments are informative or not, and she does not explicate her criteria for what constitutes informativeness. Though Hyams & Wexler (1993) do work with a concrete data set, they do not assess individual arguments on the basis of any criterion other than form of argument to determine whether they are informative or not. Thus, while their argumentation sounds reasonable on the surface, it would certainly benefit from more detailed investigation addressing these shortcomings.

The present paper seeks to give a more thorough treatment to a pragmatics account of the subject-object asymmetry. It identifies a set of factors which characterize typical situations of informativeness, and uses these factors to identify informative arguments in a data set from four children aged 2;0 through 3;6 learning Inuktitut as a first language. In addition, it assesses the extent of the links between features of informativeness on one hand and lexical vs. null and subject vs. object arguments on the other. Results indicate that a pragmatics account of the subject-object asymmetry can indeed be upheld to a much greater extent than previous research indicates, and that the factors characterizing "informativeness" are good indicators of those arguments which tend to be omitted in early child language.

3. Criteria for Informativeness

While Hyams (1986) is right that most referents involved in early child speech are present in the discourse context, this does not mean the child never has reason to say anything informative (which she also notes). In fact, several types of situations arise in everyday discourse in which the child might feel the need to reduce the potential uncertainty of the listener regarding the referents that he or she is talking about. Eight such situations are identified in the following paragraphs (including ideas taken from Greenfield & Smith, 1976; Clancy, 1980, 1993; Givón, 1983; Hyams, 1986; Chafe, 1987; Du Bois, 1987; Hirakawa, 1993).

Contrast. Children often tend to emphasize contrast in their speech, typically between candidate agents when they want to prohibit others from doing something they are doing, or when they want to do something someone else is doing. Contrast represents a situation of informativeness since information about a change in the current situation is being conveyed, which reduces uncertainty about the referent in question.

Query. Asking or responding to a question to which a referent in subject or object position is the answer is a situation characterizing informativeness since the

uncertainty regarding the identity of a given referent is either identified or resolved.

Absence. Identifying referents which are not present in the physical context represents informativeness since language is used to establish the identity of a referent whose identity cannot be gleaned from context.

Newness. Occasion often arises to introduce a new referent into discourse which has not been previously talked about. This represents informativeness since the referent being introduced cannot yet be presupposed.

Differentiation in discourse. Differentiating one referent from other potential referents already established in discourse represents informativeness since it decreases potential uncertainty concerning which of the recently-mentioned potential referents is the one now being talked about.

Differentiation in context. Differentiating one referent from other potential referents in the physical context which could fit the verb semantics and identifying features of the argument in question is informative since it selects among the alternative candidates.

Third person. Talking about third person rather than about a speech act participant greatly increases the search space for finding a referent for the argument. Rather than searching among two participants, the set of potential referents is almost infinite. Thus, defining the third person being talked about represents a situation of informativeness.

Inanimacy. The same search space logic can be applied to talking about inanimate rather than animate entities. In the typical child discourse, there are a limited number of animate entities (child, mother, father, sibling, dog, etc.) compared to the vast number of inanimate entities (table, cup, toy, juice, television, plant, clothes, etc.). Thus, the search space is vastly increased once an inanimate entity is being discussed.

These eight features, then, will be used in the remainder of the paper as characterizations of informativeness. An argument which represents one or more of these features is informative, while an argument which represents none of these features is not informative.

4. Methodology

In the remainder of the paper, this issue of informativeness as an explanation for the subject-object asymmetry is investigated using data from Inuktitut child language. Inuktitut is a language of the Eskimo-Aleut family. It is polysynthetic and morphologically ergative, has basic SOV word order, 4 persons, 3 numbers, and a rich system of nominal and verbal inflections. Particularly relevant for this paper is the morphological form in which arguments appear. Most importantly, verbal inflection is obligatory in Inuktitut. Each verb stem takes one verbal inflection, a portmanteau morpheme which gives information about verbal modality and person and number of both subject and object. Third person arguments may optionally be represented by a lexical or demonstrative noun phrase. However, representation of first and second person

arguments other than in the inflection (i.e. by an independent pronoun) is ungrammatical in the adult language. The utterances in (2) show each of the three possible options of argument representation with the intransitive verb *sinik-* 'sleep': (2a) shows a lexical argument, (2b) a demonstrative, and (2c) a null argument. Each of the three options in (2) is possible for a third person argument, while only (2c) is possible for a first or second person argument.

- (2) a. *Piaraıt sinisijuq.*
 piaraq-it sinik-si-juq
 baby-ABS.2Ssg sleep-PRES-PAR.3Ss
 'Your baby is sleeping.' (Paul 3;3)
- b. *Una sinisimmat.*
 u-na sinik-si-mmat
 this.one-ABS.SG sleep-PRES-CSV.3Ss
 'This one is sleeping.' (Lizzie 2;10)
- c. *Sinilimat.*
 sinik-liq-mmat
 sleep-PRES-CSV.3Ss
 'He/she is sleeping.' (Elijah 2;9)

Note that the term "argument" here is meant to indicate mention of a referent in either subject or object position, including both independent and inflectional forms. "Overt" means that there is an independent lexical or demonstrative representation of the argument in addition to verbal inflection, while "null" means that the argument is represented only by verbal inflection (or occasionally nothing if the inflection is omitted). Use of this terminology does not intend to indicate a particular position on whether the verbal inflection in Inuktitut is actually a pronominal argument or not.

Data used in this study is taken from four Inuit children aged 2;0 through 3;6, living in a functionally monolingual Inuit community of some 200 inhabitants in arctic Quebec (Allen, 1996). Data were collected by videotape, four hours per month for nine months, in naturalistic communication situations with friends and family. The tapes were transcribed and checked by native speakers, and coded for morphology and syntax relevant to this research. All utterances included in the analysis were fully intelligible and complete from the point of view of the child's intonation. Repetitions and imitations were included though routines (e.g. songs) were not. In addition, all utterances analyzed contained a verb root and/or a verbal inflection.

This paper analyzes the first, middle and last sessions for each child, including a total of 2019 intransitive clauses and 641 transitive clauses. Intransitive clauses include those which are simple active, passive, antipassive, noun incorporation, and complex active (e.g. want to V); transitive clauses include those which are simple active, causative, noun incorporation, and complex active (e.g. want to V). Example (3a) shows a simple active intransitive utterance with an overt lexical subject, while (3b) shows a transitive causative utterance with both arguments null.

- (3) a. *Qupanuarulu qailangamat.*
 qupanuaq-guluk-Ø qai-langa-mmat
 bird-DIM-ABS.SG come-FUT-CSV.3sS
 'The little bird is going to come.' (Lizzie 2;11)
- b. *Ijukkatilauruk.*
 ijuikkaq-tit-lauq-gup
 fall-CAUS-POL-IMP.2sS.3sO
 'Make it fall.' (Louisa 3;2)

All subject and object arguments are coded for the factors described in section 3 above.

5. Results

The question to be answered, then, is whether informativeness could be a reasonable explanation for the subject-object asymmetry found in argument ellipsis in early child language. Since this paper uses data from Inuktitut, it must first be determined that a subject-object asymmetry of this sort in fact exists in Inuktitut. As noted above, Inuktitut has verbal inflection for both subjects and objects, leading to the expectation that subjects and objects will be treated equally in terms of grammar. Nevertheless, Inuktitut child data show a subject-object asymmetry for null arguments: subjects are omitted in 85% of cases, while objects are omitted in only 69% of cases.

It is possible, however, that the subject-object asymmetry shows up in Inuktitut just because first and second person arguments can never be overt, and first and second person arguments appear more often in subject position than in object position. Two types of data show that this does not appear to be the full explanation. First, if the subject-object asymmetry were due to the ungrammaticality of overt first and second person arguments, then no asymmetry should be observed for third person arguments ($n=1542$) in these positions. However, the asymmetry still occurs for third person arguments alone, though it is weaker: third person subjects are omitted in about 70% of cases, while third person objects are omitted in only about 63% of cases.

Second, some independent evidence suggests that first and second person arguments would be omitted more often than third person arguments if the opportunity were given in Inuktitut. Research in discourse indicates that first and second person arguments are omitted more often than third person arguments in languages in which overt first and second person arguments are possible. In Inuktitut child language, a similar effect can be seen for the omission of verbal inflection, which is not grammatical in adult language, but is equally possible for all three persons. Children do occasionally omit the verbal inflection ($n=163$). Thus, if first and second person inflections are omitted more than third person inflections, this would suggest that first and second person overt arguments might well be omitted more than third person arguments if the former were grammatical in Inuktitut. In fact, ellipsis of verbal inflection does occur more frequently in first

and second person contexts (7.5%) than in third person contexts (2%), suggesting that first and second person arguments are preferentially dropped regardless of the typology of Inuktitut in this respect.

Thus, it seems that the whole explanation for the subject-object asymmetry in Inuktitut does not lie in the fact that first and second person arguments can never be overt in Inuktitut. Nonetheless, the remainder of the results will be presented for both all arguments taken together, and for third person arguments only.

Returning to the main question, then, we have established that there is a subject-object asymmetry in Inuktitut. In order to show that informativeness could provide an explanation for this asymmetry, it must first be shown that it plays a part in determining whether an argument is overt or not. Figures 1 and 2 show that arguments which are overt tend to have a given informativeness feature more frequently than arguments which are null. For example, some 47% of all the overt arguments in the data set introduce new referents, while only about 7% of the null arguments in the data set introduce new referents. This pattern is the same for each of the informativeness features observed, apart from query for which there is too little data to make any clear statements. Figure 2 indicates that the same pattern holds if only third person arguments are considered, though the differences are smaller.

A final piece of evidence needed to argue that informativeness can explain the subject-object asymmetry is whether informative arguments appear more frequently in object position than in subject position. Figures 3 and 4 show that the percentage of object positions containing arguments with a given informativeness feature is higher than the percentage of subject positions containing the same feature for three of the features, and for informativeness overall, when all persons are included. Thus, some 25% of objects serve to introduce new referents, while only some 10% of subjects have this feature. The effect is not seen for query since so few arguments have this feature. The reverse effect is seen for contrast, likely since so much of the contrast in this data set involves contrasting the agents who are permitted to do certain activities.

The difference between subject and object in terms of percentage of arguments having a given informativeness feature is somewhat smaller when only third person arguments are considered, as illustrated in Figure 4. Note that contrast patterns more like the other features when only third person arguments are considered; this is likely because much of the contrast involves disputes between first and second person agents over involvement in activities, and once these are removed from the analysis, the remaining contrasts fall into the expected pattern. It is not clear why more subjects more often represent absent referents than do objects.

To summarize so far, then, four significant results have arisen from examination of the Inuktitut data. First, Inuktitut child language does display a subject-object asymmetry with regard to early null arguments.

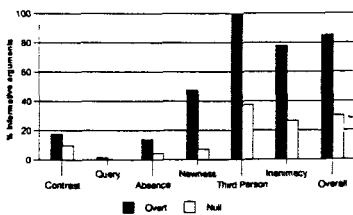


Figure 1: Percentage of arguments with informativeness features for overt vs. null arguments, all persons

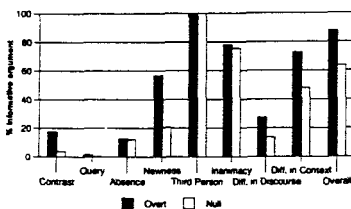


Figure 2: Percentage of arguments with informativeness features for overt vs. null arguments, third person only

Second, this asymmetry does not seem to just be due to the ungrammaticality of first and second person overt arguments in Inuktitut. Third, informativeness does seem to have something to do with overtness of arguments. Fourth, informativeness features also appear more frequently in association with arguments in object position than with arguments in subject position. Thus, we are led to conclude that informativeness could well be an adequate explanation for the subject-object asymmetry.

A stronger position would be that informativeness is a better explanation for the subject-object asymmetry than is a grammatical account. This idea was investigated by comparing two types of subjects - subjects of intransitives (Si) and subjects of transitives (St). Under a grammatical account of the subject-object asymmetry, subjects of transitives should be null equally as much as subjects of intransitives since they are licensed in the same way and positioned in the same place. However, DuBois's (1987) Preferred Argument Structure, a set of statistical tendencies observed across a number of typologically diverse languages, indicates that St and Si behave differently in discourse: both lexical and new arguments rarely appear in St position, but often appear in Si position. Thus, a pragmatic but not a grammatical account would predict an asymmetry between St and Si in occurrence of null arguments.

The first question to ask, then, is whether there is an asymmetry between St and Si for null arguments in the

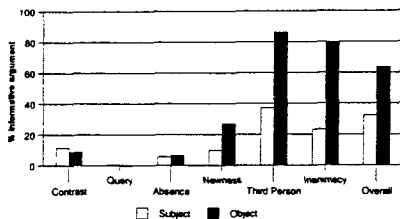


Figure 3: Percent of arguments with informativeness features in subject vs. object position, all persons

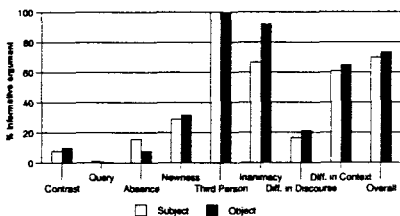


Figure 4: Percent of arguments with informativeness feature in subject vs. object position, third person only

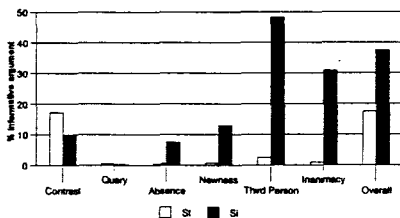


Figure 5: Percentage of arguments with informativeness features in St vs. Si vs. O positions, all persons

Inuktitut data. In fact, such an asymmetry does occur: 98% of transitive subjects are null, while only 85% of intransitive subjects are null (compared to 70% of objects).

Not surprisingly, arguments with given informativeness features also appear more frequently in Si position than in St position. For example, Figure 5 indicates that some 15% of arguments in Si position introduce a new referent, whereas only 1% of those in St position introduce a new referent. This is consistent with adult discourse research such as DuBois's (1985, 1987) Preferred Argument Structure.

Since a grammatical account would not predict a difference between Si and St either in terms of the

morphological form of arguments (null vs. overt) or in the distribution of informativeness features between Si and St, the results here suggest that a pragmatics account may well be preferable to a grammatical account in explaining the subject-object asymmetry.

6. Conclusion

This paper has shown that a subject-object asymmetry occurs in Inuktitut even though both subject and object ellipsis are licensed through verbal inflection. In the Inuktitut data examined here, overt arguments are more frequently informative than null arguments (Figures 1-2), and informative arguments appear more frequently in object position than in subject position, especially arguments that are new, third person and inanimate (Figures 3-4). While the ungrammaticality of overt first and second person arguments in Inuktitut clearly has an effect, it does not appear to be the whole effect (Figures 2 and 4). The asymmetry between St and Si for both overtness and informativeness of arguments indicates that a pragmatics account of the subject-object asymmetry may well be preferable to a grammatical account (Figure 5). Finally, discourse-pragmatic factors seem to play a much larger role than previously attributed to them in the distribution of null arguments in general, and in the subject-object asymmetry in particular.

Acknowledgements

I wish to thank Melissa Bowerman, Penny Brown, Pat Clancy, Cornelia Hamann, Heike Schröder and David Wilkins for discussions relevant to this paper, and the Inuit children and their families for participating in the research. Funding for data collection was provided by the Kativik School Board and the Social Sciences and Humanities Research Council of Canada.

Abbreviations

ABS = absolutive	POL = politeness
CAUS = causative	PRES = present
CSV = causative	SG = singular
DIM = diminutive	2sS = 2nd sg subject
IMP = imperative	3sS = 3rd sg subject
PAR = participative	3sO = 3rd sg object
2Ssg = 2nd sg possession	

References

Adone, D. (1994). *The acquisition of Mauritian Creole*. Amsterdam: Benjamins.
 Allen, S.E.M. (1996). *Aspects of argument structure acquisition in Inuktitut*. Amsterdam: Benjamins.
 Bloom, L. (1970). *Language development: Form and function in emerging grammars*. Cambridge, MA: MIT Press.
 Bloom, L. (1973). *One word at a time*. The Hague: Mouton.
 Bloom, P. (1990). Subjectless sentences in child language. *Linguistic Inquiry*, 21(4), 491-504.

Chafe, W. (1987). Cognitive constraints on information flow. In R. Tomlin (Ed.), *Coherence and grounding in discourse*. Amsterdam: Benjamins.
 Clancy, P. (1993). Preferred argument structure in Korean acquisition. In E. V. Clark (Ed.), *Proceedings of the 25th Annual Child Language Research Forum* (pp. 307-314). Stanford, CA: CSLI.
 DuBois, J. W. (1987). The discourse basis of ergativity. *Language*, 63, 805-855.
 Gerken, L. (1990). A metrical account of subjectless sentences. In J. Carter, R. Dechaine, B. Philip & T. Sherer (Eds.), *Proceedings of NELS 20* (Vol. 1). Amherst, MA: GLSA.
 Givón, T. (Ed.). (1983). *Topic continuity in discourse: A quantitative cross-language study*. Amsterdam: Benjamins.
 Greenfield, P. & Smith, J.H. (1976). *The structure of communication in early language development*. New York: Academic Press.
 de Haan, G.J. & Tuijman, K. (1988). Missing subjects and objects in child grammar. In P. Jordens & J. Lalleman (Eds.), *Language development*. Dordrecht: Foris.
 Hamann, C. (1995). Null arguments in German child language. In D. MacLaughlin & S. McEwen (Eds.), *Proceedings of the 19th Annual BUCLD* (pp. 240-254). Somerville, MA: Cascadia Press.
 Hirakawa, M. (1993). Null subjects versus null objects in an early grammar of Japanese. *McGill Working Papers in Linguistics*, 9, 30-45.
 Hyams, N. (1986). *Language acquisition and the theory of parameters*. Dordrecht: Reidel.
 Hyams, N. & Wexler, K. (1993). On the grammatical basis of null subjects in child language. *Linguistic Inquiry*, 24, 421-459.
 Jaeggli, O. & Hyams, N. (1988). Morphological uniformity and the setting of the null subject parameter. In J. Blevins & J. Carter (Eds.), *Proceedings of NELS 18* (Vol. 1). Amherst, MA: GLSA.
 Krämer, I. (1995). The occurrence of subjects and objects in child Dutch: A longitudinal study of the spontaneous speech of seven children aged 1;8 to 3;1. Master's thesis, University of Amsterdam, Netherlands.
 Mazuka, R., Lust, B., Wakayama, T. & Snyder, W. (1986). Distinguishing effects of parameters in early syntax acquisition: a cross-linguistic study of Japanese and English. *Papers and Reports on Child Language Development*, 25, 73-82.
 Radford, A. (1990). *Syntactic theory and the acquisition of English syntax*. Cambridge, MA: Basil Blackwell.
 Valian, V. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition*, 40, 21-81.
 Wang, Q., Lillo-Martin, D., Best, C. T. & Levitt, A. (1992). Null subject versus null object: Some evidence from the acquisition of Chinese and English. *Language Acquisition*, 2(3), 221-254.