

주어 탈락 현상의 언어간 비교와 이론적 모색

김 영 주 (홍익대학교)

Null Subjects in Crosslinguistic Acquisition Data and Theoretical Implications

Young-joo Kim (Hong-Ik University)

한 문장의 주어가 우리말에서처럼 수의적으로 나타날 수 있는가 혹은 영어에서처럼 반드시 표면에 나타나야 하는가에 따라 공주어언어 (null-subject language)와 비공주어언어 (non-null-subject language)로 분류된다. 이러한 주어 탈락 현상에 대하여 이론적으로 다양한 가설이 제기되어 왔다. 본 논문에서는 한국어의 언어 습득 자료에 나타나는 주어 탈락의 양상을 살피고, 이를 비공주어언어인 영어자료와 공주어언어인 이탈리아어, 포르투갈어, 중국어자료와 비교함으로써 궁극적으로 이론적 가설을 비교 평가하는 데 이바지하고자 한다.

1. Introduction

Korean is a so-called "null subject" language. That is, the lexical realization of the subject in a sentence in Korean is optional. In this respect, Korean contrasts with "non-null subject" languages such as English and French -- which require the subject of a sentence to be lexically realized -- while it is grouped with other null subject languages such as Chinese, Japanese, Italian, and Spanish.

This variation in whether or not the subject need be phonologically realized has been the focus of much attention in current theoretical linguistics and in language acquisition research. In particular, a lot of different theories have been proposed to account for the phenomenon of null subjects in early grammars of non-null subject languages such as English. In this chapter, I will introduce and discuss certain major hypotheses about the null subject phenomenon in child language. My presentation of previously proposed theoretical models, however, is not exhaustive. My main interest in this paper is in general principles of acquisition, and how Korean data are relevant to those principles. Thus, I will concentrate on models which include crosslinguistic predictions, or have crosslinguistic implications.

We may roughly sort various theories of the null subject phenomenon in child language into two groups: competence-deficit

explanations and performance-deficit explanations. A competence-deficit model is that children produce null subjects because their grammars are syntactically different from adults'. A performance-deficit model is that children have basically the same grammar as adults, but they omit subjects due to processing limitations or other performance factors.

2. Competence-Deficit Hypotheses

2.1. Hyams (1986)

Hyams (1986) presents a competence-deficit hypothesis, in the framework of Government and Binding (GB) theory (Chomsky, 1981), about the phenomenon of null subjects in child language. Hyams proposes that all children start off with a "pro-drop" grammar, and that children acquiring languages such as English later switch the parameter setting from pro-drop to non-pro-drop. This "Pro-Drop Hypothesis" sees English-speaking children as having a competence deficit which erroneously allows small pro in their grammar.

Hyams' (1986) Pro-Drop Hypothesis is based on the observation that not only children acquiring null subject languages, but also those acquiring non-null subject languages produce subjectless sentences in the early period of language acquisition. For example, English-speaking children have been reported to produce "subjectless" sentences (Bloom, Lightbown, & Hood, 1975); and it has also been claimed that this same period of development is characterized by a notable lack of expletive pronouns - the pleonastic *it* and existential *there* (Hyams, 1986). Some time later these subjectless sentences become rare, and English-speaking children become consistent in their use of lexical subjects. At the same time, the expletive elements *it* and *there* are said to make their first appearance (Hyams, 1986).¹

Hyams (1986) proposes a parameter-setting account of the null subject phenomenon in early child language. She assumes, following ideas of Rizzi (1982), that the empty category which appears in subject position of tensed clauses in null subject languages is pro

¹. Valian (1991) reports, however, that expletives were rarely used by the 21 American children across the MLU range she observed (MLU 1.53-4.22). Even when children were producing subjects for 90% of their utterances containing verbs, they produced few expletives, probably for semantic reasons. On the other hand, if expletive use is taken to be diagnostic of knowledge that subjects are obligatory, even the lowest MLU group children understood that. Valian therefore concludes that there is nothing in her English data to suggest that expletive use is related to subject use.

Wang et al. (1992) also report that some English-speaking children in their data used both overt and null expletives at the time when they were producing some null subject utterances.

- a pure pronominal without phonetic content. Hyams claims that null subjects in early language can be explained much in the manner of adult null subject languages such as Italian and Spanish. Specifically, she argues that the *Pro-Drop Parameter*, which accounts for the difference between languages like Italian and English with respect to the possibility of unexpressed subjects, comes fixed at an initial setting, one which permits phonologically null subjects. The central claim of Hyams' (1986) proposal is that all children start out speaking a null subject language: children acquiring a non-null subject language, such as English, eventually change the initial parameter setting based on certain information in the input data.

2.2. Jaeggli and Hyams (1988); Hyams (1992)

Jaeggli and Hyams (1988) and Hyams (1992) note that this original account of Hyams (1986) has some inadequacies. One of its most serious problems is as follows. The inflectional systems of Italian or Spanish are "rich" enough to allow for the null subject phenomenon. Pronominal AGR features in these languages get realized in the form of rich overt inflection, and the latter serves to recover or identify the content of the null category. Children acquiring richly inflected languages such as Italian do acquire the inflectional system at a very early age. Thus we could say that in Italian children's grammar a null subject can be identified through rich inflection. However, we cannot say the same thing about the null subject phenomenon in early English. English-speaking children use null subjects despite the fact that verbal morphology has not yet been acquired at that point. Thus, if young English speakers are speaking "Italian," their production of null subject utterances in the absence of the overt manifestation of pronominal agreement features in the form of rich inflection violates the spirit of recoverability, i.e., the identification requirement on null elements, which is a standard assumption in GB theory.

Jaeggli and Hyams (1988), therefore, make a major modification of the null subject analysis proposed in Hyams (1986). Following Jaeggli and Safir's (1989) proposal, Jaeggli and Hyams (1988) argue that null subjects are permitted only in those languages which have morphologically uniform inflectional paradigms. "Morphological Uniformity" is defined as follows (Jaeggli and Safir, 1989, p.30):

An inflectional paradigm P in a language L is morphologically uniform if and only if P has either only underived inflectional forms or only derived inflectional forms.

The following matrix represents an example of a morphologically uniform inflectional paradigm: in Spanish, a tensed verb is inflected for number, person, tense, and mood:

Spanish

habl-o	'I speak'
habl-as	'you:SG speak'
habl-a	'he speaks'
habl-amos	'we speak'
habl-ais	'you:PL speak'
habl-an	'they speak'

Note that all the forms for 'speak' in Spanish have a derived morphology.

According to Jaeggli and Safir (1989), some of the inflections may be identical in form. The only thing that matters is whether or not all of them are derived (=morphologically complex). Compare the Spanish verbal paradigm with examples of morphologically "mixed" or "non-uniform" paradigms, such as English or French verbal forms:

English

to talk	infinitive
talk	imperative (=stem)
talk	present 1SG, 2SG, 1PL, 2PL, 3PL (=stem)
talk-s	present 3SG
talk-ed	past
talk-ing	gerund

In this paradigm, some forms are exactly identical to the verb stem. Thus, predicate inflectional paradigms in English are not morphologically uniform.

What about languages like Japanese or Chinese, which show no number-person agreement at all? Jaeggli and Safir (1989) classify Japanese as morphologically uniform, on the basis of the fact that its verbal paradigms show overt morphological markers of tense, mood, and aspect, as well as markers for negation, causation, volition, etc., as illustrated below: all of its verbal forms are analyzed as stem + affix. Chinese is also a language with morphological uniformity, since its verbal paradigms show no inflection at all: all are underived forms.

Japanese

yom-u	'read-present'
yom-ta	'read-past'
yom-eba	'read-conditional'
yom-oo	'read-imperative'
yom-ita-	'read-volitional'
yom-are-	'read-passive'
yom-ase-	'read-causative'

Chinese
xihuan 'like'

According to Jaeggli and Safir's definition, then, Japanese and Chinese are morphologically uniform, just like Italian-type languages, whereas languages like English and French are morphologically "mixed." And the licensing of null subjects is restricted to morphologically uniform languages.

The most important part of Jaeggli and Safir's (1989) and Jaeggli and Hyams' (1988) modification of Hyams' (1986) earlier account of the null subject phenomenon is that they now distinguish between 'licensing' and 'identification' of null subjects. The null subjects are licensed by morphological uniformity, but they are identified by some other means. This is where the role of rich agreement comes in: in Italian-type languages, which show rich person-gender agreement in verbal inflections, null subjects can be identified by AGR. In contrast, in languages like Japanese or Chinese, which are morphologically uniform (and therefore license null subjects) but do not have person-gender agreement, null subjects can be identified by a Discourse Topic. The crucial difference between Jaeggli and Hyams (1988) and Hyams (1992) on the one hand, and Hyams (1986) on the other, is that the new account now proposes that the licensing of null subjects is not dependent on pronominal AGR(eement). Hyams had to make this alteration to account for (i) the fact that young English-speakers sometimes use null subjects despite the lack of rich agreement in their language, and (ii) the fact that the so-called "discourse-oriented" languages also permit null subjects in the absence of rich AGR. Under the new analysis, null subjects in early English can be *identified* by a discourse topic just like null subjects in East Asian languages.

2.3. Hyams and Wexler (1993)

Hyams and Wexler (1993) continue to argue for the topic-drop (rather than *pro*-drop) hypothesis of the missing subjects in early English proposed by Jaeggli and Hyams (1988), although they do not particularly adopt Morphological Uniformity as the licensing condition for null subjects. Hyams and Wexler argue that English- as well as Chinese-speaking children have a topic-drop grammar, whereas Italian-speaking children have a *pro*-drop grammar, in which null subjects are identified by "rich" AGR(eement) as in the adult language. They predict that Chinese-speaking children should drop subjects at a lower rate than Italian-speaking children, and at a rate similar to that of English-speaking children (all else being equal). Such a prediction is made on the ground that since "every finite sentence contains AGR, null subjects will always be a grammatical option for the Italian-speaking children, whereas not every sentence has an appropriate (subject) topic, and the Chinese- and the English-speaking children will therefore have fewer opportunities to drop subjects." We will discuss this prediction later.

Hyams (1986), Jaeggli and Hyams (1988), Hyams (1992), and Hyams and Wexler (1993) all consider English-speaking children as having an incomplete grammar of English. Their hypotheses, therefore, are variants of a competence-deficit hypothesis for the null subject phenomenon in non-null subject languages.

3. *A Performance-Deficit Hypothesis*

Valian (1990, 1991) provides a type of processing-deficit model which aims at a crosslinguistic explanation. She proposes to re-examine English acquisition data, arguing that English-speaking children's use of subjects has been erroneously described in previous research. Her own examination of the English acquisition data shows that even below MLU 2, American children use subjects quite often (in about 69% of clauses containing a verb) and that there is a shift to much greater usage of subjects, around 89% of the time, just after MLU 2. In contrast, the total proportion of subject-containing +V clauses in her Italian acquisition data stays constant at about 30% for the age range of 1;6 to 2;6. On the basis of English and Italian data, she suggests that children learning both languages quickly establish the correct value of the null subject parameter for their languages -- much more quickly than has generally been suggested.² According to Valian, the data indicate no point at which American children's grammar clearly licenses null subjects. Valian also analyzed observational data from Greek children, ranging in age from 2;0 to 2;5, and the children produced subjects only 19% of the time (Valian, in press). Yet in another study, Valian and Eisenberg (1994) discovered that two-year-old Portuguese-speaking children increased their use of subjects from 28% in the lowest-MLUW (Mean Length of Utterance in Words) group to 72% in the highest-MLUW group, the latter almost perfectly matching the adult speakers. She concludes that the children as a whole cannot be described as having an incomplete grammar, and that all children, regardless of language, produce subjects less often than adults do, as a result of their performance limitations.

4. *Korean Data - Subject-Drop Rate*

Both Hyams' and Valian's discussion of the null subject issue depend mainly upon the acquisition data of English and Romance languages. Not much research finding is available on a different type of null subject language, i.e., so-called "discourse-oriented

². Hyams and Wexler (1993) calculated proportion of subject drop in the CHILDES transcripts of Brown's (1973) Adam and Eve. They used eight 2-hour samples taken from Adam 2;5-3;0 and eight samples from Eve 1;6-2;1. They excluded the copula and auxiliary *be*, and imperative sentences from those utterances containing a verb. They report that Adam dropped subjects 55% in Period I and 29% in Period II; Eve dropped subjects 39% in Period I and 15% in Period II.

languages" which (unlike Italian) do not exhibit subject-verb agreement but which still allow null subjects. This section aims to provide relevant information that can be used for crosslinguistic comparisons on some crucial points that have yet to be done.

The proportion of overtly expressed subjects have been calculated in the speech of the three Korean-speaking children in my samples in the earliest period of multi-word utterances. Detailed general guidelines for coding are available in Kim (in press). Table 1 presents the proportions of overt subject utterances out of clauses containing a verb in monthly samples of three children. Tables 2 and 3 respectively present the proportions of overt subject clauses containing a verb in adult input to children, and in adult-to-adult conversations.

Table 1. Proportion of Clauses Containing an Overt Subject Out of Non-Imperative/Non-Propositive Clauses Containing a Verb in Korean Children's Speech

P

age	sample length	overt subject +V clauses excluding quasi-set-phrases		overt subject +V clauses including quasi-set-phrases	
1;7	(90min)	0	(0/13)	0	(0/13)
1;8	(90min)	0	(0/18)	0	(0/18)
1;9	(90min)	.17	(26/157)	.16	(26/158)
1;10	(90min)	.39	(92/236)	.32	(92/292)
1;11	(90min)	.37	(106/284)	.34	(106/310)
2;00	(60min)	.22	(64/285)	.20	(62/315)
2;01	(90min)	.38	(133/353)	.36	(133/367)
2;02	(90min)	.34	(148/440)	.33	(148/453)
2;03	(90min)	.50	(178/358)	.43	(178/413)
2;04	(60min)	.47	(101/216)	.38	(101/265)
2;05	(90min)	.42	(163/392)	.35	(163/463)
2;06	(90min)	.39	(149/383)	.30	(149/498)

C

age	sample length	overt subject +V clauses excluding quasi-set-phrases		overt subject +V clauses including quasi-set-phrases	
1;11	(60min)	.15	(7/47)	.15	(7/47)
2;00:00	(60min)	.13	(13/100)	.12	(13/105)
2;00:17	(60min)	.25	(31/123)	.25	(31/126)
2;01	(60min)	.23	(43/188)	.21	(43/205)
2;02	(60min)	.53	(59/111)	.50	(59/119)
2;03	(60min)	.43	(60/140)	.40	(60/150)
2;04	(60min)	.37	(35/95)	.31	(35/112)
2;05	(60min)	.48	(94/194)	.45	(94/211)
2;06	(60min)	.36	(79/222)	.32	(79/247)

J

age	sample length	overt subject +V clauses excluding quasi-set-phrases		overt subject +V clauses including quasi-set-phrases	
2;00	(120min)	.42	(140/334)	.33	(140/419)
2;01	(120min)	.53	(232/437)	.46	(232/508)
2;02	(150min)	.42	(259/618)	.40	(259/649)
2;03	(60min)	.39	(107/274)	.38	(107/279)
2;04	(60min)	.39	(91/234)	.40	(91/229)

Table 2. Proportion of Clauses Containing an Overt Subject Out of Non-Imperative/Non-Propositive Clauses Containing a Verb in Korean Adult Input

child age	sample length	overt subject +V clauses excluding quasi-set-phrases		overt subject +V clauses including quasi-set-phrases	
P 1;07	(45min)	.41	(111/269)	.40	(111/281)
P 2;00	(30min)	.36	(103/285)	.35	(103/292)
P 2;04	(30min)	.37	(122/332)	.36	(122/337)
C 1;11	(30min)	.40	(80/202)	.40	(80/202)
C 2;06	(30min)	.43	(152/356)	.43	(152/356)
J 2;00	(30min)	.37	(75/203)	.34	(75/223)
J 2;04	(30min)	.43	(124/289)	.42	(124/298)

Table 3. Proportion of Clauses Containing an Overt Subject
Out of Clauses Containing a Verb in Korean
Adult-to-Adult Conversations and Writing
(from Hong, 1985)

	proportions of overt subject +V clauses	
Spoken ^a	.31	(462/1509)
Edited spoken ^b	.32	(196/613)
Written ^c	.51	(529/1031)

^a. Spoken: transcribed data of informal interviews of 5 native Korean males (ages: 17, 24, 45, 48, and 49) recorded in Seoul, Korea, in 1984. The interviewer elicited narratives from each interviewee, and the conversation included various topics about the interviewees' lives. The data included few, if any, imperatives, propositives, set-phrases and quasi-set-phrases of the type categorized above for the purpose of Tables 9 and 10 (Hong, personal communication).

^b. Edited spoken: radio comedy "Caychi puin Acha puin" and a TV drama (20 min.)

^c. Written: children's version of the Hong Kiltong Story, a college newspaper editorial, an academic article, and two light essays written by women

Table 1 shows that Korean children start with predicates alone, with no overt subjects, as indicated by 0 in P's samples collected at 1;7-1;8. The proportion of overt subjects increases as children get older, except for J, who is already exhibiting a stabilized adult pattern (.35-.45) in this respect at the start of the data collection. From Tables 2 and 3, the proportion of overt subjects in adults' input samples is about .40, which is somewhat higher than the proportion of subject use in adult-to-adult conversations (i.e., .31). The difference could be attributed, among others, to the frequent use of label-asking questions of the type *i-ke mue-ya?* 'this-one what-COP:INT' (= 'What is this?') and their declarative counterparts *i-ke X-ya* 'This is X' in adult-child interactions. This type of questions, along with non-existence statements produced by children when they look for something: *X eps-ta* 'X not:exist-DECL' (= 'X is gone. '), also has the effect of increasing the proportion of overt subjects in children's speech compared with adult-to-adult conversations. More frequent change of conversational topics in adult-child interactions than in elicited narratives from adults, and caregivers' consideration to make discourse topics explicit for children may also have contributed to a higher proportion of subject use in adult-child conversations than in adult-adult conversations. The results presented in these tables show that Korean children start with subjectless predicates,

but that they soon begin to speak much like adults -- as early as around their second birthday -- as far as subject use is concerned: the proportion of overt subjects increases with age until it gets stabilized at about 35-45%.

5. *Comparison with English, Italian, Portuguese, and Chinese*

Valian (1991) investigated the use of subjects in data collected from 21 American children and adults. The proportions of utterances containing a verb which also contain a subject noun phrase in her data are as follows:

Table 4. *American Children*
Mean Proportion of Non-Imitative, Non-Imperative
Clauses Containing a Verb and an Overt Subject
(from Valian, 1991)

Mean MLU	N	Mean Age(Range)	Proportion Subject Use
1.77 (Group I)	5	2;0 (1;10-2;2)	.69 (sd = .12)
2.49 (Group II)	5	2;5 (2;3-2;8)	.89 (sd = .04)
3.39 (Group III)	8	2;5 (2;3-2;6)	.93 (sd = .04)
4.22 (Group IV)	3	2;7 (2;6-2;8)	.95 (sd = .03)
Adults (in immediate replies to children's utterances)			.96-.98 (sd = .02-.04)

The five children in Group I showed an average use of subjects at almost 70%. The five children in Group II showed an average use of almost 90%, and the proportion of overt subjects thereafter was consistently higher, increasing slightly in the next two groups.

Table 5, from Valian (1991), shows the Italian children's use of subjects in utterances containing a verb. Five Italian-speaking children were observed 11 times beginning at 1;6 or 1;7 and ending at 2;4-2;5. Time I covers the first 5 sessions (1;6 - 1;10), and Time II, the last 6 sessions (2;0 - 2;5).

Table 5. *Italian Children*
 Mean Proportion of Non-Imitative, Non-Imperative
 Clauses containing a Verb and an Overt Subject
 (adopted from Valian 1991)

	N	Age Range	Proportion Subject Use
Time I	5	1;6-1;10	about .30
Time II	5	2;0-2;5	about .30
Adults	2		.46-.56 ³

The comparison of American and Italian children's data exhibits a clear difference in the pattern of subject use: the overall proportion of overt subjects in Italian stays constant at about 30%, less than half the average usage of Group I American children. On the basis of these data, Valian concludes that American children look very different from Italian children with respect to the use of subjects. According to Valian, Italian children look the way one would expect the learner of a prototypical Romance null subject language to look: during the period of 1;6-2;6 they include subjects in a minority of their sentences (about 30%). In contrast, American children in each of the four MLU groups included subjects in most of their sentences. Valian therefore argues that young American children actually do not think their language is a null-subject language like Italian.

Valian and Eisenberg (1994) investigated subject drop in Portuguese-speaking children's and adults' speech. They collected spontaneous cross-sectional data from 20 children, for 15 to 30 minutes for each subject. Speech of 8 interacting adults was also analyzed. Table 6 shows the result.

³. Bates (1967) supplies figures for the Italian parents in a total of three samples. According to Bates, the parents used subjects about 30-40%, but Valian's (1991, p.48) calculations from Bates' data yielded 46-56%.

Table 6. *Portuguese-Speaking Children*
 Mean Proportion of Non-Imitative, Non-Imperative
 Clauses containing a Verb and an Overt Subject
 (adopted from Valian and Eisenberg, 1994)

	N	Age Range	MLUW	Proportion Subject Use
Group I	7	2;2-2;9	1.71	.28 (range = .08-.59)
Group II	10	2;0-2;8	2.43	.38 (range = .27-.72)
Group III	3	2;3-2;10	3.83	.57 (range = .44-.82)
Adults	8		3.88	.56 (range = .49-.65)

Valian (in collaboration with Arsenidou) also analyzed observational speech from five Greek children, ranging in age from 2;0 to 2;5: the children produced subjects in about 20% of their clauses containing verbs (Valian, in press).

Wang, Lillo-Martin, Best, and Levitt (1992) collected elicited connected discourse data from nine Chinese children, ranging in age from 2;0 to 4;6, nine English-speaking children, ranging in age from 2;5 to 4;5, and nine Chinese adults. Since the method of data collection and the coding system are different, only an indirect comparison could be made between their Chinese data and my Korean data. Table 7 shows the results from Wang et al. (1992).

Table 7. *Chinese and American Children, and Chinese Adults*
Mean Proportion of Sentences with Overt Subjects
in Elicited Production Data
(adopted from Wang et al. (1992))

Lang/Group	N	Age Range	MLU	Proportion Subject Use
Chinese I	3	2;0-2;5	3.41	.44 (range = .38-.52)
English I	3	2;5-2;10	3.51	.74 (range = .41-.83)
Chinese II	3	3;1-3;5	4.41	.54 (range = .41-.66)
English II	6	3;6-4;5	4.47	.91 (range = .81-.96)
Chinese III	3	4;1-4;4	5.28	.62 (range = .59-.71)
Chinese Adults	9			.64 (range = .57-.75)
Chinese Adult-Adult Conversation	5	(total # of sentences = 1507)		.54 (range = .49-.61)

Wang et al. conclude on the basis of their results that the mean overt subject rate in Chinese children's speech is much lower (.53) than that of American children (.85) and also a bit lower than that of the Chinese adults (.64).

If we compare Chinese and Korean adult-to-adult conversation data (presented in Table 3), which are quite similar in data collection method and sample size, Korean adults appear to produce considerably less overt subjects (.31) than Chinese adults (.54). The comparison shows that even within the "discourse-oriented" language group, individual languages may quite differ in the rate of subject ellipsis. Such a difference with respect to subject drop between Chinese and Korean adult languages seems to be reflected in child language data, too. Although only an indirect comparison is possible due to different characteristics of the data, Korean children appear to drop subjects more than Chinese children.

Since similar coding procedures have been used for spontaneous production data, it would be meaningful to make crosslinguistic comparisons between the Korean data presented in Tables 1 and 2, and the English and the Portuguese data collected and analyzed by Valian and her colleague (Tables 4 and 6).

What first attracts our attention is the sharp contrast between the overt subject rates in Korean and English. While Korean children's production of overt subjects seems to get stabilized at about .35-.45 between the age range of 1;10 - 2;2, American children in the same age range produce subjects almost 70% of the time, and their overt subject use increases to almost 90% by 2;5.

Hyams and Wexler (1993) predicted that children speaking "discourse-oriented" languages should drop subjects at a lower rate than "rich-Agreement" languages, since every finite sentence contains Agr, always making null subjects a grammatical option, whereas not every sentence has an appropriate (subject) topic, providing fewer opportunities for (subject) topic drop. They also argued that the lower rate of subject drop by Wang et al.'s Chinese children than by Valian's (1991) Italian children confirmed their prediction. The overt subject rates both in children's and adults' data in Korean and Portuguese go exactly in the opposite direction. That is, Korean, a discourse-oriented language, exhibits a higher rate of subject drop than Portuguese, a rich Agreement language. The low overt subject rate in Korean also makes it hard to claim that early English patterns with discourse-oriented languages with respect to subject drop, as Hyams and Wexler (1993) actually do on the basis of Wang et al.'s data.

A common characteristic among the children in the three language groups is that regardless of their target language they all start producing subjects less often than adults do, as Valian (in press) also points out. But they reach the adult rate of overt subjects very early. Since even children acquiring null subject languages produce overt subjects less often than adults do at early stages of producing verb-containing utterances, it has been suggested that part of their nonproduction is due to performance factors (Valian, 1991). Children learning Korean, where the adult overt subject rate is probably the lowest among the languages where the relevant data have so far become available, also appear to go through a period during which their production of overt subjects remains only at half the adult rate. This supports the account that all children at the beginning of combinatorial speech experience certain performance difficulties with respect to the production of overt subjects.

Developmental changes within languages and crosslinguistic comparisons taken together, there seems to be no compelling evidence to posit that English-speaking children at the beginning stage of combinatorial speech have their parametric value(s) incorrectly set with respect to the production of subjects, to say nothing of theoretical advantages of the continuity hypothesis that it does not have to account for how children come to change their analysis or to reset the relevant parameter. Crosslinguistic data suggest that very young children not only know whether their language allows null subjects or not, but also show sensitivity to actual frequencies of subject drop in their target language.

6. Summary

The crosslinguistic comparison of English, Italian, Chinese, Portuguese, and Korean data reveals that children learning each language omit subjects at different rates. English-speaking children supply lexical subjects at the highest overall rate from very early stages of acquisition (.69 for Group I and .93 for Group III; .97 for adults). Portuguese-speaking children (Valian & Eisenberg, 1994) increased their use of overt subjects with development: from .28 (Group I) to .57 (Group III), while the adult rate was .56. Chinese children (Wang et al., 1992) increased their use of overt subjects from .44 (Group I) to .62 (Group III), while the adult rate was .64. The proportion of overt subjects in Italian children's speech (Valian, 1991) was the lowest: about .30 for Time I and II, while the adult rate was about .50. My Korean data come from a longitudinal study, and therefore provides a closer picture of developmental changes in individual children, while the finding has to be confirmed by a larger number of subjects. Korean children increased their use of overt subjects from 0 to .35-.45, and the adult rate was also .35-.45. Korean therefore seems to exhibit the lowest rate of overt subjects among languages investigated so far in this respect.

A crosslinguistic generalization is that the frequencies of overt subjects in children's speech closely reflect the patterns in their target languages. That is, in each language, the patterns of subject use in children's speech by the third year look very much like their adult input. The results from Korean acquisition data are consistent with Valian's (1991) earlier findings that both American and Italian children are very sensitive, early on, to how frequently subjects are used in the input.

American children produced a higher proportion of pronominal subjects over overt subjects than Portuguese-speaking children: .53 in Group I and .79 in Group III American children, vs. .18 in Group I and .43 in Group III Portuguese-speaking children. Italian children studied in Valian (1991) produced much less pronominal subjects: .22 at Time I; .35 at Time II. Statistics on Chinese children's use of pronominal subjects is not available. Korean children and adults alike used pronominal subjects infrequently. Calculation of the rate of pronominal subjects over the entire subject positions, filled and unfilled, also shows that both Korean children and adults use very small proportion of overt pronouns and especially personal pronouns, in subject positions.

According to Clancy (1994a), lexical subjects in Korean children's speech most frequently exhibit properties of discourse prominence, and null subjects least exhibit such properties, while pronominal subjects are placed in the middle ground in the scale of discourse prominence. Such finding confirms the widely held assumption that null subjects in discourse-oriented languages such as Korean represent "old" information recoverable from the context,

while lexical subjects more frequently represent "new" information. Crucially, future research will have to examine the nature of ellipited subjects in languages like English and Portuguese/Italian and to compare the results with findings about "discourse-oriented" languages like Korean.

The data presently available to us indicate that concrete distributional characteristics in the input play a greater role -- and at a much younger age -- than has previously been thought, thereby shaping children's grammar. The data also suggest that youngest children acquiring different null subject languages all seem to experience certain amount of processing difficulties at the beginning of combinatorial speech, because they produce less overt subjects than older children and adults. This, in turn, suggests that the subject ellipsis observed in early non-null subject languages does not necessarily have to be attributed to children's incomplete grammar. Although it has been argued that young English speakers have a topic-drop grammar (Jaeggli and Hyams, 1988; Hyams and Wexler, 1993), the actual rate of subject drop in English and Korean exhibits the two extremes in the crosslinguistic variation. For future research on this issue, detailed (longitudinal) data especially from children who are just beginning to produce their first verbs in various languages seem to be essential. Presentation of Korean data in this paper was an attempt in that direction.

REFERENCES

- Bloom, L. M., Lightbown, P., & Hood, L. Structure and variation in child language. *Monographs of the Society for Research in Child Development*, 1975, 40.2, Serial No. 160.
- Chomsky, N. *Lectures on government and binding*. Dordrecht: Foris Publications, 1981.
- Clancy, P. M. Reference and the co-construction of argument structure in the acquisition of Korean. Unpublished manuscript, University of California at Santa Barbara,
- Hong, Y. A quantitative analysis of *ka/i* and *(n)un* as subject markers in spoken and written Korean. In S. Kuno, J. Whitman, I.-H. Lee & Y.-S. Kang (Eds.), *Harvard studies in Korean linguistics* (Vol.1). Seoul: Hanshin Publishing Company, 1985, 145-153.
- Huang, J. On the distribution and reference of empty pronouns. *Linguistic Inquiry*, 1984, 15.4, 531-574.
- Hyams, N. *Language acquisition and linguistic theory*. Dordrecht: D. Reidel Publishing Company, 1986.

- Hyams, N. A reanalysis of null subjects in child language. In J. Weissenborn, H. Goodluck & T. Roeper (Eds.), *Theoretical issues in language acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1992, 249-268.
- Hyams, N., & Wexler, K. On the grammatical basis of null subject in child language. *Linguistic Inquiry*, 1993, 24.3, 421-459.
- Jaeggli, O., & Hyams, N. Morphological uniformity and the setting of the null subject parameter. In *Proceedings of NELS*, 1988, 18. GLSA, University of Massachusetts, Amherst. 238-253.
- Jaeggli, O., & Safir, K. The null subject parameter and parametric theory. In O. Jaeggli & K. Safir (Eds.), *The null subject parameter*. Dordrecht: Kluwer Academic Publishers, 1989.
- Kim, Y. The acquisition of Korean. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition* (Vol. 4). Hillsdale, NJ: Lawrence Erlbaum Associates, in press.
- Rizzi, L. *Issues in Italian syntax*, Dordrecht: Foris Publications, 1982.
- Valian, V. Logical and psychological constraints on the acquisition of syntax. In L. Frazier & J. de Villiers (Eds.), *Language processing and language acquisition*. Dordrecht: Kluwer Academic Publishers, 1990.
- Valian, V. Syntactic subjects in the early speech of American and Italian children, *Cognition*, 1991, 40, 21-81.
- Valian, V. (in press). Children's postulation of null subject: Parameter setting and language acquisition. In B. Lust, G. Hermon & J. Kornfilt (Eds.), *Syntactic Theory and First Language Acquisition: Crosslinguistic Perspectives: Binding, Dependencies, and Learnability* (Vol.2). Hillsdale, NJ: Erlbaum.
- Valian, V., & Eisenberg, Z. Syntactic subjects in the spontaneous speech of Portuguese-speaking children. Unpublished manuscript, Hunter College, 1994.
- Valian, V., Hoeffner, J., & Aubry, S. Young children's imitation of sentence subjects: Evidence of processing limitations. Unpublished manuscript, Hunter College, 1994.
- Wang, Q., Lillo-Martin, D., Best, C. T., & Levitt, A. Null subjects and objects in the acquisition of Chinese. *Language Acquisition*, 1992, 2, 221-254.