MAXIMIZING PROCESSING IN AN SOV LANGUAGE

Mieko Ueno
Department of Psychology and Beckman Institute, University of Illinois, Urbana-Champaign

Maria Polinsky
Department of Linguistics, University of California, San Diego

July, 2005

Address correspondence to:

Mieko Ueno
2217 Beckman Institute
University of Illinois, Urbana-Champaign
405 N. Matthews
Urbana, IL  61801

phone: (217)244-1952
fax:   (217)333-8206
email:  ueno@psych.uiuc.edu
MAXIMIZING PROCESSING IN AN SOV LANGUAGE
ABSTRACT

Head-driven parser models predict that SOV languages are harder to process than SVO languages, since the parser has to hold both S and O until it reaches V, instead of just S as in an SVO language. However, since no reading time differences have been attested between SOV and SVO languages, we hypothesize that either these models are wrong or SOV languages have strategies to compensate for the late appearance of the verb. The results presented in this paper indicate that SOV languages such as Japanese do use compensatory strategies to reduce the distance before the verbal head. First, a language may use a higher percentage of one-place predicates, thus minimizing the number of structural arguments; second, it may minimize the number of overtly expressed preverbal arguments (regardless of the adicity of the predicate). While the former strategy seems specific to OV languages, the latter strategy, manifested as a higher incidence of subject pro-drop with two-place predicates, is found in a number of languages, regardless of word order and verb agreement, which suggests more a universal processing constraint than a compensatory strategy for head-final languages.
1. **INTRODUCTION.** In both theoretical linguistics and psycholinguistics, a verb is considered to play a crucial role in a clause; it subcategorizes for arguments and determines the overall predicate structure. The well-known conceptual and grammatical asymmetry between nouns and verbs (Gentner 1978, 1981, Langacker 1987, Baker 2003) has to do with the inherent complexity of verbal meanings, which make reference not just to entities, as nouns do, but to the relationships between them as well. In psycholinguistics, a verb is considered to play a crucial role in a parsing decision; at the verb position the parser determines how other elements of the sentence, such as a dislocated filler, is interpreted (Gorrell 1993, Gibson & Hickok 1993, Pickering & Barry 1991, Pickering 1993), and the parser immediately utilizes the information regarding how likely a given verb would take an NP or sentential complement (Trueswell et al. 1993, Garnsey et al. 1997). This notion is representatively expressed by Pritchett’s (1992) head-driven parser model, which argues that syntactic attachment happens at a verbal head that gives the parser information about thematic roles and grammatical relations between argument NPs.

For example, the sentence fragment in (1a) *What did Calvin ...* does not give the parser enough information for determining syntactic attachment, since the sentence could continue as *What did Calvin think was good?*, with *what* as the theme and subject of the embedded clause, or *What did Calvin eat?*, with *what* as the patient and object of the matrix clause. In (1b), it is not until the verb *eat* that the parser receives enough information to determine syntactic attachment.

(1) a. What did Calvin ... not enough information for syntactic attachment
b. What did Calvin eat? enough information for syntactic attachment

Since the verb crucially determines the thematic roles and grammatical functions carried by the referents of the NPs, these roles and functions may remain unclear until the verb is reached. Gibson (1991) argues that each NP that is momentarily without a thematic role and each
thematic role that is momentarily unassigned both impose a burden on working memory (also see Gibson 1998, 2000, Babyonyshev & Gibson, 1999, among others, for more recent versions of his locality theory).

Following this line of logic, we might predict that SOV languages are harder to process than SVO languages. Since the parser would have to hold both S and O until it has reached V in an SOV language, as opposed to only S in SVO languages, it would be reasonable to think that this might create an extra processing cost. However, this is false—to our knowledge, there has been no attested difference in the reading times of SOV and SVO languages. In the absence of such a difference, there can be two possibilities, either SOV and SVO are processed in the same way, or SOV is actually harder to process than SVO but SOV languages have strategies to compensate for the late appearance of the verb. These two hypotheses are shown in (2).

(2) a. Hypothesis 1:

SOV and SVO are equally easy to process. There is no need for compensatory strategies for the late appearance of the verb in SOV languages.

b. Hypothesis 2:

SOV is harder to process because the verb comes later in the utterance following nominal arguments. SOV languages have a need for strategies to compensate for the late appearance of the verb.

One of the possible compensatory strategies is overt case-marking on argument nouns in SOV languages. Overt case-marking allows an early determination of the thematic role and grammatical function of each nominal referent before the parser reaches the clause-final verb position. This is consistent with Greenberg’s (1966) classic claim that ‘if in a language SOV is the dominant order, the language almost always has a case system’ (Universal 41). And indeed, of many SOV languages, only a small subset do not have case-marking—Abkhaz is the most
notable example (Hewitt 1979). Meanwhile SVO languages often have little or no case-marking; the mere percentage of such languages is much higher than that of their SOV counterparts.

Another possibility to consider is the reduction of overt arguments. If there were a smaller number of overt arguments, the parser would not have to hold as many NPs before reaching the verb position. This corollary to Hypothesis 2 is stated as follows in 3.

(3) Corollary to Hypothesis 2:

In addition to overt case-marking, SOV languages may minimize the number of overt arguments. The minimization can be achieved in two different ways—either by using a higher number of one-place predicates (intransitive proper, passive, antipassive, middle, etc.), or by omitting overt argument/s via pro-drop. It is unknown which of these strategies if any is found in a head-final language.

In order to test our hypotheses, we conducted a comparative corpus study of English (an SVO language) and Japanese (an SOV language) sentences from various genres. We show that Japanese uses constructions with a smaller number of overt argument NPs than English does, and argue that this is a strategy for facilitating processing which compensates for the late appearance of V in a head-final language. Section 2 explains how the corpus was constructed and analyzed. Section 3 presents the main results obtained in the study. Section 4 discusses these findings in relation to our hypotheses, followed by some remaining questions in Section 5 and a conclusion in Section 6.

2. METHODS.

2.1. MATERIALS. A total of 2,400 sentences (300 sentences/genre x 4 genres x 2 languages) were surveyed for predicate types. The four genres included home decoration
magazines, mystery novels, books about Japanese politics, and children’s utterances from CHILDES (age around 3;8) in both English and Japanese (see the data source references for full citations).

2.2. Procedure. Sentences were manually coded by a Japanese graduate student in syntax at UCSD\(^1\), who was unaware of the purpose of the study, and then checked by the first author. Sentences were coded for matrix clause constructions only; adjoined/conjoined and embedded clauses were not considered, in order to avoid possible differences between matrix and subordinate structures. Relative clauses modifying arguments in the matrix clause were considered part of the relevant NP; their internal structure was not considered (cf. 4c and 5a). Complement clauses in argument positions were counted on a par with non-clausal arguments (cf. 5b).

Matrix clauses were classified as ‘one-place predicates’, ‘two-place predicates’, and ‘sentence fragments’. ‘One-place predicates’ include ‘intransitives’, ‘non-verbal predicates (adjectival and nominal)’, and ‘passives (with and without the by-phrase)’. ‘Two-place predicates’, or more accurately 2+-place predicates, include ‘transitives’ and ‘ditransitives’. For the purposes of this study, the main contrast was between one-place and 2+-place predicates, which is why we made a decision to collapse ‘transitives’ and ‘ditransitives’. ‘Sentence fragments’ include bare NPs and interjections.

4 shows examples of one-place predicates in English and Japanese.
(4) One-place predicates

a. Intransitives

\[
\begin{array}{c|c}
S & V \\
\end{array}
\]
Sometimes they died.

(English—mystery novel # 7)

\[
\begin{array}{c|c}
S & V \\
\end{array}
\]
思い出が、水がこみあげてきた。
omoigakezu namida-ga komiagete-kita.
unexpetedly tear-N came.out

‘Tears came out unexpectedly.’

(Japanese—mystery novel # 16)

b. Non-verbal predicates

\[
\begin{array}{c|c|c}
S & Adj & \\
\end{array}
\]
He’s bigger.

(English—children’s utterance #88)

\[
\begin{array}{c|c|c|c}
S & NP \\
\end{array}
\]
ほら、こっちが火事やて。
hora, kocchi-ga kaji-yate
look here-N fire-be

‘Look, here is the fire.’

(Japanese—children’s utterance #184)
c. Passives

\[ S \quad pV \]

Few wood cabinets are constructed of solid wood

(English—home decoration magazine #61)

分厚い 木の ドアには 職人の 手作りになる 鉄の 金具が 付けられています。

\[ S \quad pV \]

thick wood-G door-D-T craftsman-G handmade iron-G metal-N attached be

‘An iron metal handmade by a craftsman is attached to a thick wooden door.’

(Japanese—home decoration magazine #202)

5 shows examples of two-place and higher predicates in English and in Japanese. In assembling corpus statistics, we did not make a distinction between NP and sentential complements; they were all counted as ‘objects’.

(5) Two-place predicates

a. Transitives with NP complements

\[ S \quad V \quad O \]

These dimensions represent cultural continua, not dichotomies, and the differences are in degree, not in kind.

(English—book about Japanese politics #8)

たとえば、鈴木内閣と 同じように、高い 支持率のもとで 船出した

\[ S \quad O \quad V \]

Suzuki.ministry-with same.as high support.rate-G under sailed

池田内閣の 初の 状況が このことをよく示している。

\[ S \quad O \quad V \]

Ikeda.ministry first.time-G situation-N this-A well indicating is

‘For example, the initial situation of the Ikeda ministry which sailed out under a high support rate just like the Suzuki ministry is indicating this well.’

(Japanese—book about Japanese politics #20)
b. Transitives with sentential complements

\[
\begin{array}{ccc}
S & V & O \\
\end{array}
\]

You said I could go.

(English—children’s utterance #25)

\[
\begin{array}{ccc}
S & O & V \\
犬も & pro & pro \\
pro & 乗りたいなあて & ゆうとる。
\end{array}
\]

‘The dog is also saying that (he) wants to ride (it).’

(Japanese—children’s utterance #227)

c. Ditransitives

\[
\begin{array}{ccc}
S & V & O \\
\end{array}
\]

Cool colors—such as mint green or sky blue—give a room a fresh, airy ambiance.

(English—home decoration magazine #215)

\[
\begin{array}{ccc}
S & O & O & V \\
バリ島の & 高級 & バンブー家具が、涼しげな & ナチュラル感を部屋に & 運んでくれます。
barito-no & kokyu & banbu & kagu-ga & suzushigena nachurarukan-o & heya-ni & hakondekuremasu.
\end{array}
\]

‘The high class bamboo furniture from Bali Island brings the room a cool natural feeling.’

(Japanese—home decoration magazine #278)

When coding Japanese, nominative-marked NPs were coded as subjects, but we also kept separate track of constructions that, arguably, contained nominative-marked objects, some examples of which are shown in 6. We distinguish three main types of such constructions (see also Takezawa 1987):

(a) ‘existential-possessive constructions’ with the possessor in the dative case and possession in the nominative, as in 6a
(b) ‘potential constructions’ with the subject in the dative case and the theme in the nominative (Dubinsky 1993), as in 6b

c. ‘need constructions’ with the agent in the dative case and the theme in the nominative, as in 6c.

(6) Constructions with nominative-marked objects

a. Existential-possessive construction

木の皮や　枯れ枝　などの　植物には　乾いた　ぬくもり感が　あります。

ki-no kawa-ya kareeda nado-no shokubutsu-ni-wa kawaita nukumorikan-ga arimasu

‘Some plants, such as wooden bark and withered brunches, have dry warmth.’ (lit.: ‘For some plants such as wooden bark and withered branches, dry warmth exists.’)

(Japanese—home decoration magazine #231)

b. Potential construction

夫が　いわんと　している　ことが、　にわかには　信じられなかった。

otto-ga iwanto shiteiru koto-ga niwakani-wa shinji-rare-na-katta.

‘(I) could not instantly believe what (my) husband was going to say.’ (lit.: ‘What (my) husband was going to say was not instantly believable.’)

(Japanese—mystery novel #80)

c. ‘Need’ construction

選挙に　出馬し　当選を　期する　ためには　、　巨額の　資金が　必要と　なった。

senkyo-ni shutsubashi tousen-o kisuru tame-ni-wa kyogaku-no shikin-ga hitsuyoto natta.

‘In order to run for an election and expect to win, (they) came to need a large sum of money.’ (lit.: ‘In order to run for an election and expect to win, a large sum of money became necessary.’)

(Japanese—book about Japanese politics #99)
Additionally, we kept record of middle constructions such as 7 in both languages; this was particularly relevant in keeping track of the alternation between transitive and intransitive predicates.

(7) Middle constructions

Both semi-gloss and high-gloss channels wipe clean easily and resist stains and scuffing.

(English—home decoration magazine #207)

ゆみ（うみ）が
Yumi (umi)-ga
ocean-N

見えた。
mieta.
showed

‘I saw the ocean.’ (lit.: ‘The ocean was visible.’)

(Japanese—children’s utterance #116)

The occurrences of pro-drop were also recorded, for both null subjects and objects in both languages. Imperatives were not included as cases of pro-drop. 8 shows some examples.

(8) Pro-drop examples

a. S(subject)-drop only

pro Got it!

(English—mystery novel # 147)

pro
boku-ni
me-D

向かって、
mukatte
toward

一目散に
ichimokusanni
for.one’s.life

走って 来たんです。
hashitte kitandesu.
running came

‘(She) came running toward me for her life.’

(Japanese—mystery novel #11)
b. SO(subject and object)-drop

no examples attested in the English corpus

\[ pro \text{ ？} \text{mou ？ dashitan?} \]

‘Have (you) already got (it) out?’

(Japanese—children’s utterances #50)

c. O(object)-drop only

Her’s gonna make \textit{pro}

(English—children’s utterances #37)

\[ \text{otto-ni mukatte watashi-ga jitsuwa anata-ni kakurete chiwawa-o katte irunoyo} \]

\[ \text{丈夫-D facing I-N actually you-D hiding chihuahua-A keeping am.you.know} \]

\[ \text{と 告白して いた としても、 丈夫 pro 信じなかった に違いない。} \]

\[ \text{that confessing was even.if husband-T believed.not-that must.not} \]

‘Even if I was confessing to (my) husband, “Actually (I) am keeping a chihuahua behind your back.”, (my) husband must not have believed (it).’

(Japanese—mystery novel #253)

For each language, the number of sentence tokens of each classification in each genre, such as ‘two-place predicates in the Japanese mystery genre’, was added up. The occurrences of one-place and two-place predicates for the two languages were placed into a two-by-two table (English vs. Japanese and one-place vs. two-place) for each genre and examined by a Pearson chi-square test to determine whether there was any significantly different distribution of one-place and two-place predicates between the two languages. This was to test whether Japanese would have a higher percentage of one-place predicates than English.
The constructions with arguably nominative-marked objects (see 6 above) were counted separately, and separate chi-square tests were run treating those constructions as two-place, instead of one-place predicates, in order to see if these predicates would make any statistically significant difference in the one-place vs. two-place distribution between English and Japanese tested above. Further, different types of one-place predicates, such as ‘passives’ and ‘intransitives’ were counted separately to see whether there was any difference in the proportions of these predicate types between English and Japanese.

The number of sentence tokens with null arguments was also tallied up for each predicate type\(^3\), genre, and language. Across languages, the occurrences of sentences with and without pro-drop were placed into a two-by-two table (English vs. Japanese and pro-drop vs. no-pro-drop) for each genre and tested for any significantly different distribution of pro-drop between the two languages (i.e. whether Japanese has more instances of pro-drop than English) by a Pearson chi-square test. Within a given language, the sentence tokens with or without pro-drop for either one-place or two-place predicates were placed into a two-by-two table (one-place vs. two-place and pro-drop vs. no-pro-drop) for each genre to determine whether there were any statistically different distributions of pro-drop across predicate types by a Pearson chi-square test. This was to test whether either language would show a higher rate of pro-drop with two-place than one-place predicates. An alpha level of .05 was used for all statistical tests.

3. RESULTS.

3.1 ONE-PLACE VS. TWO-PLACE PREDICATES. Figure 1 shows the distribution of sentence tokens for one-place vs. two-place predicates for English and Japanese. When collapsed across
genres (Figure 1a), Japanese exhibits a significantly higher percentage of one-place predicates in comparison to English (English vs. Japanese: 51% vs. 73%, $\chi^2(1) = 107.13$, $p < .001$).

When each genre is examined separately, this is also the case for all the genres except for books about Japanese politics (home decoration: E 42% vs. J 81%, $\chi^2(1) = 89.59$, $p < .001$; mystery: E 49% vs. J 65%, $\chi^2(1) = 13.73$, $p < .001$; children: E 41% vs. J 76%, $\chi^2(1) = 44.90$, $p < .001$; politics: E 68% vs. J 71%, $\chi^2(1) = .58$, $p = .447$). Notice that the English discussion of Japanese politics suddenly shows a much higher percentage of one-place predicates (68%) than other genres (home decoration: 42%; mystery: 49%; children: 41%). Meanwhile, Japanese has a smaller inter-genre difference (range of one-place predicates in Japanese: 65% - 81%).

Table 1 shows the number of constructions that include arguably nominative-marked objects in Japanese. We can see that such sentence tokens are fairly small in number, 76 tokens or 6% of the total number of the Japanese sentences examined.

Even when these constructions are counted as two-place predicates instead of one-place predicates, Japanese still exhibits a significantly higher percentage of one-place predicates than English collapsed across genres (E 51% vs. J 65%, $\chi^2(1) = 46.22$, $p < .001$). When each genre is examined separately, this is also the case for home decoration magazines (E 42% vs. J 72%, $\chi^2(1) = 52.60$, $p < .001$) and children’s utterances (E 41% vs. J 74%, $\chi^2(1) = 41.58$, $p < .001$), with mystery novels following the same trend (E 49% vs. J 56%, $\chi^2(1) = 2.88$, $p = .090$). The politics genre again shows no significant difference between the two languages (E 68% vs. J 63%, $\chi^2(1) = 1.98$, $p = .159$).
Table 2 shows the occurrences of middle constructions in both languages. There are only one tokens of middle constructions in English, which is about 0.2% of the total number of sentence tokens. There are 16 tokens of middle constructions in Japanese, which is about 1.3% of the total sentence tokens in Japanese.

Table 3 shows the details of one-place predicate types.

Collapsed across genres, the higher ratio of one-place predicates over two-place predicates in Japanese appears to be due to higher percentages of non-verbal predicates (E 299 vs. J 429) and intransitives (E 151 vs. 295) in Japanese than in English. When individual genres are examined, Japanese exhibits higher percentages of non-verbal predicates than English in home decoration, politics, and children’s utterances, while intransitives show this pattern in home decoration, mystery, and politics. Interestingly, more passive sentences are used in English (105 in total) than in Japanese (37 in total) in all the genres. The political science genre includes many more passive tokens (82) in English than in Japanese (27), and this seems to have made the total number of one-place predicates in both languages equivalent within that genre, as discussed above with respect to Figure 1.

3.2 Pro-drop. Figure 2 shows the proportion of sentence tokens with and without pro-drop for English and Japanese.

When collapsed across genres Japanese exhibits a significantly higher percentage of pro-drop than English (E 3% vs. J 26%, $\chi^2(1) = 215.37$, $p < .001$). This is also the case when individual
genres are examined separately (home decoration: E 1% vs. J 27%, $\chi^2(1) = 79.38$, $p < .001$; mystery: E 2% vs. J 23%, $\chi^2(1) = 54.21$, $p < .001$; politics: E 0% vs. J 11%, $\chi^2(1) = 35.46$, $p < .001$; children: E 12% vs. J 56%, $\chi^2(1) = 86.70$, $p < .001$).

Table 4 shows the details of pro-drop types. For both languages, a majority of instances are S-drop (E 63%, J 90%). Other cases include O-drop only for English (37%) and SO-drop (8%) and O-drop only (3%) for Japanese.

Table 5 shows the types of null subject arguments. For all the genres (except for politics in English which has no pro-drop), the first and second person references are dominant in English (home decoration: 100%; mystery: 60%; children: 56%), while that is not the case in Japanese (home decoration: 26%; mystery: 52%; politics: 29%; children: 16%). All of the third person references are animate in English, while more of the third person references are inanimate than animate in Japanese. All of the referents appear to be (at least to the coder) previously-mentioned information (if not the speaker or the addressee) in the discourse that are more likely to be pronouns/demonstratives than lexical NPs, if overtly expressed.

Table 6 shows the types of null object arguments. Unlike null subjects, all the referents were third person and inanimate (except for only one instance in the mystery genre in Japanese) for all the genres in both languages. All of the referents again appear to be (at least to the coder) previously-mentioned information in the discourse that are more likely to be pronouns/demonstratives than lexical NPs, if overtly expressed.
Figure 3 shows the distribution of pro-drop for one-place vs. two-place predicates in English. When collapsed across genres, English shows a significantly higher percentage of pro-drop with two-place predicates than with one-place predicates (one-place vs. two place: 1% vs. 6%, $\chi^2(1) = 25.76$, $p < .001$). This seems for the most part due to children’s utterances. When individual genres are compared, children’s utterances is the only genre that shows a significant difference (one-place vs. two place: 3% vs. 19%, $\chi^2(1) = 13.20$, $p < .001$). The home decoration (one-place vs. two place: 0% vs. 2%, $\chi^2(1) = 2.24$, $p = .135$) and mystery (one-place vs. two-place: 1% vs. 4%, $\chi^2(1) = 2.61$, $p = .106$) genres also have more instances of pro-drop with two-place than with one-place predicates, but they are too few to show any statistical significance. Note again the politics genre has no instances of pro-drop.

--Insert Figure 3 about here--

Figure 4 shows the distribution of pro-drop with one-place vs. two-place predicates in Japanese. When collapsed across genres, Japanese also shows a significantly higher percentage of pro-drop with two-place predicates than with one-place predicates (one-place vs. two place: 21% vs. 38%, $\chi^2(1)= 31.62$, $p < .001$). This is true for all the genres (one-place vs. two place, home decoration: 21% vs. 52%, $\chi^2(1)= 22.70$, $p < .001$; politics: 6% vs. 26%, $\chi^2(1)= 24.19$, $p < .001$; children: 49% vs. 79%, $\chi^2(1)= 11.32$, $p < .001$), except for mystery novels, which, although it follows the trend somewhat numerically, does not show a statistically significant difference (one-place vs. two place: 21% vs. 25%, $\chi^2(1)= .63$, $p = .428$). As shown in Figures 3-4, for both English and Japanese, two-place predicates tend to involve more instances of pro-drop than one-place predicates. We term this ‘pro-drop bias’.

--Insert Figure 4 about here--
3.3 INTER-GENRE DIFFERENCES. As can be seen in Figure 1, Japanese exhibits more one-place predicates than two-place predicates for all the genres. The proportion of one-place vs. two-place predicates in English is about half and half for all the genres except for the political science genre, which has more one-place predicates due to a high number of passive constructions.

The order of likelihood of pro-drop by genre is similar for each language, namely (see Figures 3-4):

(9)  
a. politics < home decoration < mystery < children (English)  
b. politics < mystery < home decoration < children (Japanese)

The omitted arguments are mostly subjects in all the genres for both languages. Other cases are either object-drop only in English or both subject and object-drop or object-drop only in Japanese (see Table 4). We can see the pro-drop bias for children’s utterances in English (with other genres having too few instances of pro-drop to judge) and for all the genres in Japanese (except for the mystery genre which somewhat follows the trend but not statistically significantly).

4. DISCUSSION. The main results that emerge from the corpus study are as follows. First, one-place predicates are much more predominant in Japanese than in English. Second, Japanese shows an overall preference for pro-drop in two-place predicate clauses, a phenomenon we have termed as the ‘pro-drop bias’ in the Results Section. However, children’s utterances in English also show the pro-drop bias. Finally, Japanese and English show differences with respect to the featural characteristics of null arguments. In what follows, we will discuss each of these findings in relation to our hypotheses in turn.
4.1 Adicity of Predicates. Our results show that Japanese exhibits a higher ratio of one-place predicates to two-place predicates than English (Figure 1). The only exception to this pattern is writings about Japanese politics, which have approximately equal instances of one-place predicates across languages. We hypothesize that this is due to two different factors. First, the result is due in large part to the high number of passive constructions in the English sentences, which is probably because the use of passive is more common in academic writings than other genres in English (cf. Svartvik 1966). Second, this is the most academic, written (as opposed to spoken) of all the genres considered here, and the one that seems most immune to the constraints found in the other genres.

Recall the corollary to our hypothesis 2 repeated here:

(10) SOV languages may minimize the number of overt arguments.

This was motivated by the idea that if there were a smaller number of overt arguments, the parser would not have to hold as many NPs before reaching the verb position. We further hypothesized that the minimization may be achieved by the use of one-place predicates or pro-drop. The higher percentage of one-place predicates in Japanese supports the first possibility in that the strategy of choosing a one-place predicate over a 2+-place predicate is present in Japanese.

In addition to our frequency data, there is evidence from language acquisition that suggests that Japanese has a preference for one-place over two-place predicates. Fukuda (2005) reports that Japanese-speaking children start using the intransitive versions of the verbs that have transitive counterparts (e.g. *aku* ‘open(int.)’ vs. *ak-e-ru* ‘open(tr.)’) two to three months earlier than the transitive versions, while an English-speaking child starts producing transitive versions approximately six months earlier than their intransitive versions.
Further, the findings concerning the adicity of predicates in Japanese vs. English are also consistent with the observation in Nichols et al. (2004) that there may be cross-linguistic variation in the preference for one- or two-place predicates. To describe the difference in adicity, Nichols et al. (2004) propose the notion of lexical valence orientation. This notion captures the descriptive generalization that some languages are predominantly intransitive, and derive transitive verbs from intransitive, while other languages are predominantly transitive, deriving one-place predicates via detransitivization.\(^4\) In examining correlations between various language properties and lexical valence orientation, Nichols et al. note that languages of the SOV type (OV order in their terminology) favor the intransitive orientation (2004: 170); they suggest that this correlation is grammatically driven (as opposed to being an areal feature) but do not provide an explanation for it. If the results obtained for Japanese reliably replicate in other head-final languages, the correlation noted by Nichols and colleagues receives a processing explanation. The predominance of one-place verbs is consistent with the hypothesis that the number of overt arguments before the verb needs to be kept small.

Recall that we have used sentences in compatible genres in each language. Although our sentences for each language are not mutual translations of each other\(^5\), the sentence tokens within each genre are presumably similar to each other in terms of the concepts expressed. Then one may wonder how presumably similar concepts can be expressed equally well by one- or two-place predicates.

One possible answer to the question is that speakers of different languages describe the same event differently. For instance, Slobin (1991) proposes the notion of ‘thinking for speaking’, in which a child adopts a particular viewpoint for describing events with the architecture of a given language. For example, even when instructed to describe the same story
made of the same set of pictures, English and Spanish speakers tend to describe it differently as
in (11) (Berman & Slobin 1994: 11).

(11) a. English

And he starts running. And he tips him off over a cliff into the water. And he lands.

b. Spanish

El ciervo le llevó hasta un sitio, donde debajo había un río. Entonces el ciervo tiró al
perro y al niño al río. Y después, cayeron.

‘The deer took him until a place, where below there was a river. Then the deer threw the
dog and the boy to the river. And then they fell.’

Following the distinction between ‘satellite-framed’ and ‘verb-framed’ languages proposed by
Talmy (1985), Berman and Slobin (1994) argues that English is characterized by compact verbal
expressions which trace out the trajectory of the fall by a series of particles attached (or
’sattelited’) to a verb (satellite-framed language), such as off over a cliff into the water, while
Spanish analyzes the event into phrases, as in ‘the deer does something, and the boy and dog fall’
with the verbs functioning as bare descriptions of changes of state (verb-framed language).

Similarly, Tatsumi (1997) compares the descriptions of the same ‘frog story’ by English
monolinguals, Japanese monolinguals, and English-Japanese bilinguals and finds that English
monolinguals tend to provide elaborate descriptions of the path of the motion, Japanese
monolinguals tend to provide isolated descriptions of the location of the downward motion, and
English-Japanese bilinguals fall somewhere in between the two monolingual groups.

Berman and Slobin (1994) and Tatsumi (1997) focus on the temporal and spatial
properties of the same story narrated by speakers of different languages. However, in addition,
we might expect different sentence structures to be produced in different languages. Speakers
may focus on and express different aspects of the same events and encode them into one-place or
two-place predicates. Thus it is conceivable to think that ‘thinking for speaking’ also contributes
to the one-place vs. two-place predicate difference between English and Japanese.

Recall that the higher rate of one-place predicates in Japanese is due to more tokens of
intransitive and non-verbal (adjectives and NPs) predicates than in English (while there are more
passive constructions in English than in Japanese). As for intransitives, Japanese occasionally
treats an argument that would be expressed as an object in English as a subject. These are
realized as constructions with arguably nominative-marked objects, as discussed with respect to
6 and Table 1. Recall also from Table 2 that there are more instances of middle constructions in
Japanese (16) than in English (1) across genres. Although these facts alone do not necessarily
account for the difference in the ratio of one- vs. two-place predicates between English and
Japanese, they clearly show a preference for one-place predicates in Japanese. Also note that
there are seven middle constructions in the Japanese-speaking children’s utterances while there is
no middle construction in the English-speaking children’s utterances, being consistent with
Fukuda’s (2005) findings on the acquisition of intransitive/transitive verb forms in both
languages. We surmise that this preference for one-place predicates in Japanese can be
extended further. In addition to nominative-marked object and middle constructions, the
preference might result in the use of other intransitive verbs and non-verbal predicates. The
driving force for this might be ‘thinking for speaking’, in that Japanese speakers are taking a
grammaticized point of view to choose one-place predicates.

4.2 PRO-DROP. As discussed above, a higher rate of one-place predicates provides the
means to minimize the structural domain of the predicate. However, at this point it is unclear if
this is the only way to minimize such domain. To determine what other strategies of argument
minimization are used, we now turn to the role of pro-drop, first in Japanese, then in other languages.

Our data show that two-place predicates are more prone to pro-drop than one-place predicates, and we have termed this phenomenon the ‘pro-drop bias’. For Japanese, this is demonstrated by the results in Figure 4. These results suggest that pro-drop is certainly working to reduce the number of overt arguments before the parser hits the verb in Japanese. Given that Japanese has no verb agreement system, however, the question is whether processing a covert pronoun would really be easier than processing an overt NP. Nevertheless, null pronouns in Japanese (Kameyama 1985, 1988, Walker et al. 1994) and other languages (Turkish: Turan 1998, Yiddish: Prince 1999) are attested to be highly recoverable in a given discourse, in that the referents tend to have been previously-mentioned and highly-ranked (in terms of syntactic and pragmatic features such as being the topic or subject or being salient, thus at the center of the speaker’s attention) in the current and previous utterances. In our data, the omitted arguments, in both English and Japanese, typically map onto previously mentioned information that are easily identifiable to the coder (see Section 3.2). Thus we assume that recovering null pronouns exacts hardly any processing cost, and that the reduction of overt argument NPs by pro-drop facilitates the processing of SOV.

The question, however, is whether this pro-drop bias is specific to SOV languages, or if it would also apply to other word order types. In fact, the corpus data presented here show that English also allows for some pro-drop as well; the pro-drop rate is within 0-2% for adults but 12% for children. Although there are too few instances of pro-drop in adult English to determine whether it shows any pro-drop bias, children’s utterances certainly show it, being consistent with

If we look at other SVO pro-drop languages (see Table 7), a spoken Spanish (with rich verb agreement) corpus by Bentivoglio (1992) also shows a significantly higher percentage of pro-drop with two-place predicates than with one-place predicates \( (one\text{-}place \ vs. \ two\text{-}place: \ 458 \ (67\%) \ vs. \ 331 \ (76\%), \ \chi^2(1)= 11.73, \ p < .001) \). Similarly, a spoken Mandarin (with no verb agreement) corpus by Tao (1996) also shows a significantly higher percentage of pro-drop with two-place predicates than with one-place predicates\(^8\) \( (one\text{-}place \ vs. \ two\text{-}place: \ 130 \ (40\%) \ vs. \ 232 \ (81\%), \ \chi^2(1)= 106.53, \ p < .001) \). In addition, Sacapultec Maya, an ergative VOA language with rich verb agreement, also shows a pro-drop bias; according to a spoken corpus by Du Bois (1987), there is a significantly higher percentage of pro-drop\(^9\) with two-place predicates than with one-place predicates \( (one\text{-}place \ vs. \ two\text{-}place: \ 124 \ (47\%) \ vs. \ 156 \ (87\%), \ \chi^2(1)= 71.11, \ p < .001) \). Therefore, there seems to be a general trend towards shortening two-place predicate utterances by pro-drop for languages of various basic word orders and agreement systems, and we seem unable to say that the pro-drop bias is SOV specific.

If not a compensation strategy specifically for an SOV language, what motivates the pro-drop bias? It might be a general processing economy principle that applies to the languages of all word orders, namely, it is easier to comprehend or produce shorter utterances than longer ones. In terms of comprehension, pro-drop with a two-place predicate reduces the number of overt NPs to process, whether before or after the verb. Assuming again that recovering null pronouns is not costly, it would be more economical to comprehend shorter constructions with pro-drop than those without. In terms of production, shorter utterances would also be easier to
Bloom (1990) reports that VP length in English speaking children’s utterances increases as a function of the subject type, as in full NP < pronominal < null, and attributes it to a performance limitation of children, in that they are not capable of producing long utterances. Our data are consistent with his claim for all the genres in Japanese and children’s utterances in English, in addition to the adults’ spoken corpora in Spanish, Mandarin, and Sacapultec Maya discussed above; even adult speakers are more likely to drop argument nouns in two-place than one-place predicates. Although adults do not have as heavy a performance restriction as children do, they still seem to prefer shorter utterances over longer utterances. This is consistent with a view that pro-drop serves to satisfy performance constraints, at least to a certain degree, as opposed to a view that pro-drop is fundamentally a competence-based phenomenon (e.g. Hyams & Wexler 1993)\textsuperscript{11}.

Further, in addition to just the principle of economy for shortening utterances, there may be an interaction with discourse factors. Recall from Table 4 that subjects are much more likely to be dropped than objects, both in Japanese and English\textsuperscript{12}. This is consistent with the previous data reported for children’s utterances in both languages (Mazuka et al. 1986 for Japanese, Bloom 1990, Hyams and Wexler 1993 for English), as well as the adult spoken data in Mandarin (Tao 1996) and Sacapultec (Du Bois 1987)\textsuperscript{13} discussed above. In addition to these token frequency data, a cross-linguistic observation on the grammaticality of pro-drop (Gilligan 1987) also shows object pro-drop is more limited; among the 100 languages that semi-proportionally represent the language families of the world, 89% allow subject-drop, while 73% direct object-drop and 26% indirect object-drop. Bloom (1990, 1993) argues that this may be due to pragmatic factors surrounding subjects, which tend to be more ‘given’ than objects and thus prone to be omitted at a processing bottleneck.
Similarly, Du Bois (1987) argues that in Sacapultec Maya there tends to be only one lexical argument (which comprises new information) in a clause, with other arguments expressed as covert/overt pronouns, and that the lexical argument appears preferentially in the S or O roles, but rarely in the A role. Du Bois argues that introducing a new lexical argument referent demands enough attention that a simultaneous introduction of a second new referent within the same clause would be too costly. The driving force here again seems to be the notion that easily recoverable discourse entities do not have to be fully expressed (especially within a performance constraint), and as a result, tend to be expressed as covert/overt pronouns, making more room for new information. If subjects typically convey ‘given’ (previously mentioned and already activated (e.g. Chafe (1976)) information while objects convey ‘new’ (newly introduced) information, it is plausible to think that a transitive construction tends to have the object as a full lexical argument and the subjects as a reduced form.

As a side note, it is interesting to see a consistent difference between English and Japanese in the features of null subject pronouns (although there is no such distinction in null objects, see Table 6). As discussed in relation to Table 5, English can only take the first or second person null pronouns, with the exception of a very few animate third person null pronouns, for both children’s and adults’ utterances. On the other hand, Japanese null subject pronouns index all three persons, and in fact the majority of the null pronouns in this study have animate or inanimate third person referents. This is consistent with the general characteristic of Japanese as a pro-drop language with pro-drop not limited to a subset of person/number features, while that is not the case for English.
4.3 GENERAL DISCUSSION. Recall that our original hypothesis was that SOV languages may have strategies to compensate for the late appearance of the verb by minimizing the number of overt arguments. We further hypothesized that the minimization may be achieved by two possible strategies, namely, either by using a higher rate of one-place predicates or by pro-drop.

We have found in our data that Japanese exhibits a higher percentage of one-place predicates than English, thus the first of our hypotheses has been supported. We have also seen that this coincides with the earlier acquisition of intransitive rather than transitive verb forms in Japanese (Fukuda, 2005), as well as a morphological preference for intransitive rather than transitive verb forms in OV languages (Nichols et al. 2004).

As for the second hypothesis, however, we have found that it is true not only for Japanese but also for other languages. As a result, we are unable to say that the pro-drop bias is an SOV specific strategy. Although pro-drop is certainly working to reduce the number of overt arguments with two-place predicates in Japanese, we have also seen the same type of pro-drop bias in children’s English (SVO), Spanish (SVO), Mandarin (SVO), and Sacapultec Maya (ergative VOA). We have argued that this may be due to a general principle of economy for shortening utterances, which possibly involves an interaction with discourse factors. Given the above, we will focus on the first strategy, higher use of one-place predicates, and continue with our discussion.

The preference for one-place predicates can be understood quite naturally as a processing constraint designed to reduce the number of arguments that need to be held in working memory until the verb is encountered in a verb-final language, and this points to the role of verbal heads as discussed in the Introduction. The head-driven parser model (Pritchett 1992) has recently been called into question by evidence of the more incremental nature of Japanese sentence
processing (e.g. Kamide & Mitchell 1999, Kamide et al. 2003, Aoshima et al. 2004; see Miyamoto 2003 for an overview). However, it seems we cannot completely dismiss the importance of verbal heads in parsing operations. The parser may utilize other information that becomes available before hitting the verbal head, yet having a verb appear later rather than sooner would seem significant enough a challenge to warrant a compensatory strategy. We may be able to argue further that this compensatory strategy exists to make languages with different basic word orders to be equally easy to process.

5. REMAINING ISSUES

5.1 PRODUCTION VS. COMPREHENSION. We have argued that the Japanese language tries to reduce the number of argument NPs before the verb by heavy use of one-place predicates. This was motivated by the prediction based on the head-driven parser model, namely, holding S and O until V in an SOV language would be harder than holding only S in an SVO language. In other words, we used a comprehender’s point of view in making this prediction. However, the question is what motivates this pattern of results in terms of language production. It is possible to think that speakers and writers are sensitive to the needs of comprehenders and produce something that is easier to comprehend. But what if they do what they do because it is also easier for themselves?

One way to think about this is that since it is hard to produce a longer SOV structure, producers avoid it and instead produce a shorter SV structure. As discussed in Section 4.2 with respect to pro-drop, longer sentences may be harder to produce as well as to comprehend, regardless of the basic word order. However, can there be anything SOV specific that favors one-place predicates?
Yngve (1960) argues that in his sentence production model a phrase structure is generated from top to bottom and left to right, and the processing load at each node is proportional to the number of yet-to-be-expanded nodes (referred as ‘depth’) that must be kept in working memory. Yngve (1960) argues that this explains heavy NP shifts in English, such as *He gave the candy to the girl that he met in New York while visiting his parents for ten days around Christmas and New Year’s* (preferred) as opposed to *He gave the girl that he met in New York while visiting his parents for ten days around Christmas and New Year’s the candy* (ungrammatical?). In order to minimize the memory load, a potentially deep expression has to start at the minimum depth, the right-most position. If we follow this logic, having an object NP (deeper than V) before the verb involves more production cost in an SOV language, compared to an SVO language that has an object NP after the verb. But if a one-place SV predicate is used, there would be no such extra production cost.

Besides the phrase structural account, we may be able to say that it is generally easier to produce a smaller number of overt NPs before the verb. Lindsley (1975) reports that English-speaking subjects (after looking at a picture) took the same amount of time to initiate saying a transitive or intransitive sentence, but a shorter amount of time to say the subject only. He concludes that speakers start their utterances before they have syntactically encoded the object of a transitive action but not before they know the verb. Then a verb might have a special status in production as well. If a predicate obligatorily needs a verb that gives it some commitment (in terms of grammatical functions and thematic roles, as in comprehension), then producers might benefit from minimizing the obligation and getting to the verb quickly.\textsuperscript{15} Then the above suggests a benefit for producers to use SV instead of SOV.
It is also possible to think that the strategies for ease of comprehension have become the internalized properties of the language, and producers just follow them without any conscious thought. Such strategies might include using constructions utilizing one-place predicates. Recall from Section 4.1 that Slobin (1991) argues that when we present events or experiences in a given language, we have to take a grammaticized point of view in order to fit them into the structure of the language. Let us assume here a standard model of language production which consists of three basic levels, namely, conceptualization (deciding what to say), formulation (deciding how to say it), and articulation (saying it) (e.g. Garrett 1976, Bock & Levelt 1994). Like Slobin’s (1991) ‘thinking for speaking’, Levelt (1989) argues that there is a language-specific difference at the conceptualization level. An adult speaker of a given language has learned what to encode when preparing a message for expression, as the language specific requirements on semantic structure have become represented in the conceptualizer’s procedural knowledge base. Then it might be the case that Japanese speakers are just automatically conceptualizing elements that are likely to be expressed in one-place predicates rather than in two-place predicates as their default procedure, as in ‘The ocean was visible’ instead of ‘I saw the ocean’.

To summarize, it seems possible to come up with an production-based explanation for the dominance of one-place predicates in Japanese found in our study. The true remaining question would be the exact relationship between comprehension-based and production-based motivations and if and how they interact with each other.

5.2. What is predicted for other SVO and SOV languages and languages with other basic word orders? The second remaining question is whether the pattern of results found in this study would generalize for other SVO and SOV languages. From our data and the
additional data discussed in Section 4.2, we assume that the pro-drop bias is universal. However, how about the use of one-place predicates? Let us make some preliminary comparisons based on some of the available corpora here. Recall that when all the genres are combined in our data, the proportion of one-place and two-place predicates is 51% (555) vs. 49% (534) in English and 73% (761) vs. 27% (285) in Japanese. Regarding SVO languages, Spoken Spanish according to Bentivoglio (1992: 16) shows the proportion of 66% (920) vs. 34% (480), which still leaves Japanese with a significantly higher percentage of one-place predicates ($\chi^2(1) = 13.80, p < .001$). Similarly, Spoken Mandarin according to Tao (1996: 115) shows the proportion of 60% (513) vs. 40% (344), and this also leaves Japanese with a significantly higher percentage of one-place predicates ($\chi^2(1) = 35.39, p < .001$). Therefore, we maintain that Japanese is more one-place dominant than such SVO languages as English, Spanish, and Mandarin. We will then have to investigate other SOV languages in order to further confirm our hypothesis.

Further, if we assume that one-place predicates are truly more dominant in SOV languages than in SVO languages, what would we predict for languages with other basic word orders? First, what do we predict for verb-initial, VSO and VOS languages such as Welsh and Malagasy? Since the verb comes first in these languages, they may be more similar to an SVO than an SOV language, in that there may not be a need for the reliance on one-place predicates. Spoken Sacapultec Maya (VOS) according to Du Bois (1987: 822) shows the proportion of one-vs. two-place predicates as 59% (262) vs. 41% (180), and this again leaves Japanese with a significantly higher percentage of one-place predicates ($\chi^2(1) = 26.27, p < .001$), consistent with our prediction. Second, one may also wonder what would happen in languages that have a mixture of SVO and SOV order depending on clause type. If we examine matrix vs. subordinate clauses in Dutch or German, would we see more one-place predicates in subordinate
clauses than in matrix clauses? The question is whether the use of one-place predicates is a clause-type effect or a typological effect. There remains more to study to answer these questions.

6. CONCLUSION. In summary, a corpus consisting of 2,400 sentences in English and Japanese in four genres was examined in terms of the proportion of one-place and two-place predicates and the occurrence of pro-drop. This was to test our hypotheses as repeated here in 12.

(12) a. Hypothesis 1:

SOV and SVO are equally easy to process. There is no need for compensatory strategies for the late appearance of the verb in SOV languages.

b. Hypothesis 2:

SOV is harder to process because the verb comes later in the utterance following nominal arguments. SOV languages have a need for strategies to compensate for the late appearance of the verb.

Our data present evidence in favor of Hypothesis 2. There was a higher ratio of one-place to two-place predicates in Japanese than in English. This suggests that SOV is potentially harder to process, and the dominance of one-place predicate works as a compensatory strategy needed to make the number of overt arguments before the verb smaller and thus easier to process, possibly both in terms of comprehension and production. Although pro-drop occurred more frequently in two-place predicates than in one-place predicates in Japanese, this was also the case for other pro-drop languages with other basic word orders. The pro-drop bias thus seems to be a more universal processing constraint, such as a general economy principle of making utterances shorter, possibly interacting with discourse factors.
Perhaps there are several properties in the Japanese language that in combination contribute to ease of processing. These properties include case-marking on argument nouns and the use of more one-place predicates than two-place predicates, which is potentially due to a grammaticized way of thinking by Japanese speakers to prefer one-place predicates. Further research on other SVO and SOV languages is needed to confirm whether this difference between English and Japanese is systematic, however, our data are promising in suggesting that by utilizing various strategies, all human languages are designed to be equally easy to process despite their typological differences.
REFERENCES


Prince, Ellen. 1999. ‘Subject-pro-drop in Yiddish.’ Focus: linguistic, cognitive, and computational perspectives, eds. by Peter Bosch and Rob van der Sandt, 82-104. Cambridge, MA: Cambridge Univ. Press.


Yamashita, Hiroko, Franklin Chang, and Yuki Hirose. Producers build structures only with overt arguments. Poster presented at the Eighteenth Annual CUNY Conference on Human Sentence Processing, The University of Arizona, Tucson, AZ.

DATA SOURCE REFERENCES

1. Home decoration magazines:
   House Beautiful May 1998.80-81.
   Home Remodeling & Decoration Winter 1997-98.25-34.

2. Mystery Novels:

3. Books about Japanese Politics:

4. Children's Utterances: CHILDES
NOTES

1 Except that the types of null subject and object arguments (Tables 5-6) were coded by the first author.

2 As an operational procedure, if an NP cannot appear as a passive subject, it is not an object of a respective verb.

3 As Bloom (1990) notes, it is not clear whether children’s intuition about which verbs take obligatory objects would be identical to the adults’, but we assumed it is.

4 Nichols et al. (2004) establish this generalization on the basis of overt morphological marking. They also recognize a third type, languages that seem to derive both types from a single stem using different morphology. This type is irrelevant for the points made in this paper.

5 We chose to use sentences originally written in a given language rather than translation, in order to avoid any effect of translation and to keep the sentence tokens as natural as possible.

6 In adult genres, the omission of subjects, which is numerically quite small, consists mostly of diary-drop (Haegeman 1990, Rizzi 1994). Diary-drop has been shown to be limited to root clauses, and given that our data are based exclusively on those it is not surprising that English shows missing subjects in the written genres examined here. According to Rizzi, the null element is not a true pro but rather a trace-like null constant that picks its identification from discourse. Note however that outside children’s utterances, the majority of the null subjects are identified with first or second person (except for two third person tokens in the mystery genre). It is unclear why a null constant would be limited to (or preferentially identified with) first or second person; nothing in its features ([−anaphor, −pronoun]) points to such an identification. The pattern observed in English diary-drop actually resembles patterns of limited pro-drop
observed in several natural languages, where pro-drop is limited to first and second person (e.g. in Hebrew). This suggests that diary-drop may actually be close to true pro-drop but under very stringent conditions such as root clauses and deictic person features which allow for easier recoverability.

7 This is based on a conservative calculation using only the overt and covert pronouns on ‘Table 2a: Distribution of S & A subjects according to from (Ø, P, N)’ in Bentivoglio (1992: 16); we did not include lexical NPs into our calculation at all as they can be mere presentational NPs, such as ‘Here comes the bride’, instead of real subjects. As there are more (overt) NPs in intransitive than transitive constructions, adding non-presentational NPs is likely to further magnify the pro-drop bias in the data.

8 Note that ‘one-place predicates’ here only include ‘intransitives’ and ‘statives’ and do not include copular constructions (Tao 1996: 116-117).

9 The calculation here again is conservative, in that it is only based on S-drop and SO-drop. There may be cases of O-drop only, but we have no way of knowing the number of such cases from Du Bois’ (1987: 822) data.

10 Similar to the case of comprehension, one may wonder if producing (?) a null pronoun would be really easier than producing an overt pronoun. But it seems intuitive to think that overtly articulating something would take more resources than not doing so, and there is syntactic priming evidence suggesting that producers might actually build structures only with overt arguments (Yamashita et al. 2005).

11 Hyams & Wexler (1993: 439-441) themselves show that the VP length and subject type is correlated in adults’ Italian. They attribute it to a grammatical property of a null-subject
language, possibly resulting from an interaction of pragmatic and grammatical factors. However, it is also possible to argue that this involves a processing constraint.

12 English-speaking children’s utterances show the highest percentage (38%) of object-drop (ignoring the English home decoration genre that only has three instances of pro-drop). This, however, is still lower than subject-drop (62%), and may have been due to counting seven instances of I don’t know, one instance of She wants to, and one instance of I want to. as object-drop. If we do not count them as object-drop, the object-drop rate will be just 10%, but the pro-drop bias still holds (one-place vs. two place: 3% vs. 12%, $\chi^2(1) = 6.31, p = .012$)

13 Presumably, rich agreement langauges should allow object pro-drop if the object is indexed by verb agreement. However, even in Sacapultec that has an object agreement system, object-drop occurs less frequently than subject-drop. Among the languages without agreement that exhibit object pro-drop, Korean is the one cited most. Even in Korean, subject pro-drop occurs more often than object pro-drop (Cho 1994; Kim 2000). For instance, Kim (2000) shows that subjects are dropped in 61.9% of the cases (329: Table I) while objects are dropped in 46.9% of the cases (339: Table IV) in the data collected from six Korean children aged 1;4 through 2;5. This suggests that the mapping of an argument to an information structural status is still the deciding factor in pro-drop.


15 We thank Kay Bock for pointing this out.

16 This time the tokens with lexical NPs are included.

17 Here ‘one-place predicates’ include ‘intransitives’, ‘statives’, and ‘copulars’.

18 We thank Kay Bock again for pointing this out.
Collapsed across Genres
\( \chi^2(1) = 107.13, p < .001^{**} \)

**FIGURE 1**
One-place vs. two-place predicates for English and Japanese
FIGURE 2
Pro-drop in English vs. Japanese
FIGURE 3
Pro-drop for one-place vs. two-place predicates in English
Figure 4
Pro-drop for one-place vs. two-place predicates in Japanese
<table>
<thead>
<tr>
<th></th>
<th>Home Decoration</th>
<th>Mystery</th>
<th>Japanese Politics</th>
<th>Children</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Existential</td>
<td>9</td>
<td>12</td>
<td>12</td>
<td>2</td>
<td>35</td>
</tr>
<tr>
<td>Potential</td>
<td>16</td>
<td>12</td>
<td>9</td>
<td>0</td>
<td>37</td>
</tr>
<tr>
<td>‘Need’</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>25</td>
<td>24</td>
<td>25</td>
<td>2</td>
<td>76</td>
</tr>
</tbody>
</table>

**Table 1**

Number of constructions arguably including nominative-marked objects in Japanese

<table>
<thead>
<tr>
<th></th>
<th>Home Decoration</th>
<th>Mystery</th>
<th>Japanese Politics</th>
<th>Children</th>
<th>Total</th>
</tr>
</thead>
</table>
### Table 5
Types of null subject arguments

<table>
<thead>
<tr>
<th></th>
<th>Home Decoration</th>
<th>Mystery</th>
<th>Japanese Politics</th>
<th>Children</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>English</td>
<td>Japanese</td>
<td>English</td>
<td>Japanese</td>
<td>English</td>
</tr>
<tr>
<td>1st</td>
<td>0(0%)</td>
<td>10(13%)</td>
<td>3(60%)</td>
<td>26(41%)</td>
<td>4(22%)</td>
</tr>
<tr>
<td>2nd</td>
<td>1(100%)</td>
<td>0(0%)</td>
<td>0(0%)</td>
<td>0(0%)</td>
<td>0(0%)</td>
</tr>
<tr>
<td>3rd (animate)</td>
<td>0(0%)</td>
<td>2(3%)</td>
<td>2(40%)</td>
<td>19(30%)</td>
<td>8(24%)</td>
</tr>
<tr>
<td>3rd (inanimate)</td>
<td>0(0%)</td>
<td>54(71%)</td>
<td>0(0%)</td>
<td>12(19%)</td>
<td>16(47%)</td>
</tr>
<tr>
<td>Total</td>
<td>1(100%)</td>
<td>75(100%)</td>
<td>5(100%)</td>
<td>64(100%)</td>
<td>34(100%)</td>
</tr>
</tbody>
</table>

### Table 6
Types of null object arguments

<table>
<thead>
<tr>
<th></th>
<th>Spoken Spanish</th>
<th>Spoken Mandarin</th>
<th>Spoken Sacapultec Maya</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>one-place</td>
<td>two-place</td>
<td>one-place</td>
</tr>
<tr>
<td>pro-drop</td>
<td>458(67%)</td>
<td>331(76%)</td>
<td>130(40%)</td>
</tr>
<tr>
<td>no-pro-drop</td>
<td>227(33%)</td>
<td>102(24%)</td>
<td>195(60%)</td>
</tr>
<tr>
<td>Total</td>
<td>685(100%)</td>
<td>433(100%)</td>
<td>325(100%)</td>
</tr>
</tbody>
</table>

### Table 7
Pro-drop for one-place vs. two-place predicates in spoken Spanish (adopted from Bentivoglio, 1992: 16), Mandarin (adopted from Tao, 1996: 116-117), and Sacapultec Maya (adopted from Du Bois, 1987: 822)