SHORT REPORTS

On the (non)universality of the preference for subject-object word order in sentence comprehension: A sentence-processing study in Kaqchikel Maya

Masatoshi Koizumi  Yoshiho Yasugi  Katsuo Tamaoka
Tohoku University  National Museum of Ethnology  Nagoya University

Sachiko Kiyama  Jungho Kim
National Center for Geriatrics and Gerontology  Tohoku University

Juan Esteban Ajsivinac Sian  Lolmay Pedro Oscar García Máztar
Comunidad Lingüística Kaqchikel

The processing load of sentences with three different word orders (VOS, VSO, and SVO) in Kaqchikel Maya was investigated using a sentence-plausibility judgment task. The results showed that VOS sentences were processed faster than VSO and SVO sentences. This supports the traditional analysis in Mayan linguistics that the syntactically determined basic word order is VOS in Kaqchikel, as in many other Mayan languages. More importantly, the result revealed that the preference for subject-object word order in sentence comprehension observed in previous studies may not be universal; rather, the processing load in sentence comprehension is greatly affected by the syntactic nature of individual languages.*

Keywords: basic word order, field-based psycholinguistics, Guatemala, processing load, syntactic complexity

1. Introduction. In many flexible word-order languages, including Basque, Finnish, German, Japanese, Korean, Russian, and Sinhalese, sentences in which the subject (S) precedes the object (O) (SO word order = SOV, SVO, VSO) induce a lower processing load in comprehension than those in which the opposite occurs (OS word order = OSV, OVS, VOS), and thus they are preferred by speakers (Sekerina 1997, Bader & Meng 1999, Mazuka et al. 2002, Kaiser & Trueswell 2004, Tamaoka et al. 2005, among many others). However, previous studies on sentence processing have all targeted languages, such as Finnish, in which the subject precedes the object in the syntactically basic word order (= SO languages). Hence, it remains unclear whether the preference for SO is a reflection of word order in individual languages or of human cognitive features that are more universal. What we refer to as individual grammar theory in this short report posits that a language’s syntactically determined basic word order has a low processing load in comparison to other possible word orders, whereas

* The present study, at its various developmental stages, has been presented at the 24th annual CUNY Conference on Human Sentence Processing, the 10th meeting of the Kansai Circle of Psycholinguistics, the 145th meeting of the Linguistic Society of Japan, the 8th Workshop on Altaic Formal Linguistics, the 2nd Formal Approaches to Mayan Linguistics, and the 6th Formal Approaches to Japanese Linguistics, as well as at the ZAS (Centre for General Linguistics, Berlin), MIT, Mie University, Hokkaido University, Kwansei Gakuin University, National Tsing Hua University, and Keio University. We would like to thank the audiences for their helpful comments and encouragement, especially Judith Aissen, Noam Chomsky, Jessica Coon, Nora England, John Haviland, Robert Henderson, Klaus von Heusinger, James Huang, Sabine Iatridou, Hisatsugu Kitahara, Jaklin Kornfilt, Alec Marantz, Shigeru Miyagawa, Wayne O’Neil, Umut Özge, David Pesetsky, Colin Phillips, Uli Sauerland, Mamoru Saito, Peter Sells, Shigeo Tonoike, Hubert Truckenbrodt, Wei-Tien Dylan Tsai, Tasaku Tsunoda, Kenneth Wexler, John Whitman, and Kazuko Yatsushiro. Part of this work was supported by a Grant-in-Aid for Scientific Research (S) entitled A Field-Based Cognitive Neuroscientific Study of the Processing of OS-Type Languages (No. 22222001, PI: Masatoshi Koizumi) from the Japan Society for the Promotion of Science.
what may be referred to as universal cognition theory hypothesizes that SO word order has a low processing load regardless of the basic word order of any individual language. To verify which of these two theories is correct, it is necessary to examine languages in which the object precedes the subject in the syntactically basic word order (= OS languages), for which the two theories develop different predictions. Therefore, as described below, we conducted a sentence-processing experiment in Kaqchikel, a Mayan language spoken in Guatemala. The syntactically determined basic word order of Kaqchikel is VOS, although in general, word order is relatively flexible (García Matzar & Rodríguez Guaján 1997:333). The results of the experiment revealed that for Kaqchikel speakers, the processing load of VOS is lower than that of the two other commonly used word orders, that is, VSO and SVO. This suggests that the preference for SO in sentence comprehension may not be universal; rather, syntactic features of individual languages significantly influence sentence-processing load.

2. SO word-order preference. Evidence that SO word orders are easier to process than OS word orders in flexible word-order languages is abundant in the psycholinguistic and neurolinguistic literature. In terms of behavioral indices, Japanese readers take less time to judge whether a sentence makes sense when it has SOV word order than when it has OSV word order (Tamaoka et al. 2005). Longer reading times for OSV sentences in Japanese were also reported using self-paced reading and eye-tracking methodologies (Mazuka et al. 2002, Imamura & Koizumi 2008b). Similarly, in Finnish, the SVO order is processed faster than OVS order even when an appropriate context is provided for the latter (Kaiser & Trueswell 2004). Parallel results from the processing of orthographically and phonologically presented sentences have been reported for many other languages (see Sekerina 1997 for Russian, Tamaoka et al. 2011 for Sinhalese, Kim 2012 for Korean, among many others). In terms of neurophysiological indices, studies with functional magnetic resonance imaging (fMRI) have found that the left inferior frontal gyrus is activated more during the processing of OS word orders compared to SO word orders (Grewe et al. 2007 for German, Kinno et al. 2008 and Kim et al. 2009 for Japanese). Research on event-related brain potentials (ERPs) also supports the claim that SO word orders are easier to process. Compared with SO orders, OS orders elicit P600 and/or (sustained) anterior negativity components, suggesting that processing OS word orders places a larger load on the working memory (Roesler et al. 1998 for German, Ueno & Klüender 2003 and Hagiwara et al. 2007 for Japanese, Erdöcsia et al. 2009 for Basque).

The SO word-order preference has been observed not only for flexible word-order languages but also for languages with less flexible word order, such as English and Hebrew. For example, in English, subject relative clauses such as The reporter [who sent the photographer to the editor] hoped for a good story are easier to process than object relative clauses such as The reporter [who the photographer sent to the editor] hoped for a good story, as evidenced by both behavioral and neurophysiological indices (Caplan & Waters 1999, Just & Carpenter 2001, Grodner & Gibson 2005, Santi & Grodzinsky 2010).

Thus, we seem to have solid evidence that SO word orders are preferred to OS word orders in many languages of the world. The question then arises as to the sources of this preference in sentence comprehension. A possibility that immediately comes to mind is that it is primarily due to syntactic canonicity (i.e. individual grammar theory). According to many sentence-processing theories, including Pritchett and Whitman’s (1995) representational theory of complexity, Gibson’s (2000) dependency locality
theory, and Hawkins’s (1990, 1994, 2004) early immediate constituents, other things being equal, a language’s syntactically determined basic word order is easier to process than other grammatically possible but noncanonical derived word orders in the language. Thus, from the perspective of individual grammar theory, SO word orders were preferred in previous studies because they are the syntactically basic word orders in the target languages.

Alternatively, the SO word-order preference in sentence comprehension may be largely attributable to human cognitive features that are more universal (i.e. universal cognition theory). That there may be such features is strongly suggested by the fact that a vast majority of the world’s languages have one of the SO word orders as the basic word order (SOV: 48%, SVO: 41%, VSO: 8%, VOS: 2%, OVS: 1%, and OSV: 0.5%, according to Dryer 2005). In particular, a number of studies have shown that entities that are prominent as a result of properties such as agency, animacy, concreteness, prototypicality, and prior mention in the discourse tend to appear as sentence-initial subjects (cf. Slobin & Bever 1982, Bock & Warren 1985, Hirsh-Pasek & Golinkoff 1996, Primus 1999, Branigan et al. 2008, Bornkessel-Schlesewsky & Schlesewsky 2009).

Universal cognition theory, therefore, suggests that SO word order has a low processing load regardless of the basic word order of any individual language, again consistent with what has been reported in the literature so far.

Both individual grammar theory and universal cognition theory correctly predict the SO word-order preference in sentence comprehension in SO languages. However, their predictions diverge when it comes to OS languages. According to individual grammar theory, OS word orders should be processed faster than SO orders in these languages; universal cognition theory predicts the opposite to be the case. It is therefore necessary to study OS languages to determine which theory is on the right track. We thus turn to an OS language, Kaqchikel.

3. KAQCHIKEL. Kaqchikel is one of the twenty-one Mayan languages spoken in Guatemala. It is mainly used in the highlands west of Guatemala City, the capital. With over 450,000 speakers, it is one of the principal Mayan languages along with Kʼiche’, Qʼeqchi’, and Mam (Tay Coyoy 1996:55, Brown et al. 2006:2, Lewis 2009).

Like other Mayan languages, Kaqchikel is head-marking: subjects and objects are unmarked, and person and number agreement for both subjects and objects are obligatorily expressed on the verb. Kaqchikel is ergative, like other Mayan languages. In Mayan linguistics, ergative agreement markers (i.e. those that indicate the subject of a transitive verb) are called set A, and absolutive agreement markers (which indicate either the subject of an intransitive verb or object of a transitive verb) are known as set B. The order of morphemes in the verb is [Aspect-B-A-Verb stem]. An example is given in 1.

1 There is additional evidence for the universal nature of SO word-order preference. First, Al-Sayyid Bedouin Sign Language arose within the last seventy years in an isolated community with a high incidence of profound prelingual deafness. In the space of one generation, it assumed a grammatical structure characterized by SOV order (Sandler et al. 2005). Given that none of the neighboring languages are SOV, the SOV order seems to have emerged spontaneously in the language without any apparent external influence. Second, Gell-Manna and Ruhlen (2011) argued, given the distribution of word-order types in world languages, that the original word order in the ancestral language was SOV. Finally, Goldin-Meadow and colleagues (2008) showed that speakers of languages that differ in their predominant word orders used the actor-patient-act order, analogous to the SOV pattern, when asked to describe or reconstruct events without speaking. They took this to suggest that actor-patient-act is the natural order we impose on events when describing and reconstructing them nonverbally and exploit when constructing language anew.

Since Kaqchikel is a pro-drop language, 1 functions as an independent sentence.

Although, like many other Mayan languages, Kaqchikel allows different grammatical word orders, its standard order is 'verb-initial'. If the sentence is irreversible as in example 2 (where the meaning of the sentence collapses with the reversal of the object and subject), it can be interpreted in either VOS or VSO order. However, VOS is preferred.

(2) a. X-∅-u-chöy ri chäj ri ajanel. (VOS)
   compl-B.3sg-A.3sg-cut DET pine.tree DET carpenter
   ‘The carpenter cut the pine tree.’

   b. X-∅-u-chöy ri ajanel ri chäj. (VOS)
      compl-B.3sg-A.3sg-cut DET carpenter DET pine.tree
      ‘The carpenter cut the pine tree.’

In cases like 3a,b, where the sentence is semantically reversible (it makes sense when the object and subject are reversed), a VOS interpretation is overwhelmingly favored (even though a VSO interpretation is still possible).

(3) a. X-∅-r-oqotaj ri me’s ri tz’i’. (VOS)
    compl-B.3sg-A.3sg-run.after DET cat DET dog
    ‘The dog ran after the cat.’

   b. X-∅-r-oqotaj ri tz’i’ ri me’s. (VOS)
      compl-B.3sg-A.3sg-run.after DET dog DET cat
      ‘The cat ran after the dog.’

In cases like 4, where the subject is preposed before the verb, the subject is naturally interpreted as topical or focused.

(4) Ri ajanel x-∅-u-chöy ri chäj. (SVO)
    DET carpenter compl-B.3sg-A.3sg-cut DET pine.tree
    ‘The carpenter cut the pine tree.’

In this sense, SVO order is pragmatically marked. Furthermore, SVO order has traditionally required transformation of the predicate, such as by adding an agent-focus morpheme and deleting the ergative agreement marker. Thus, it can be said that SVO is also marked from a morphological perspective. (In modern Kaqchikel, however, it is possible to attain the SVO word order without transforming the morphological form of the verbal complex (retaining the same morphological form as in VOS or VSO), as is evident in example 4. For these and other reasons, the majority of Mayan language researchers consider the syntactically determined basic word order of modern Kaqchikel to be VOS (Rodríguez Guaján 1994:200, García Matzar & Rodríguez Guaján 1997: 333, Tichoc Cumes et al. 2000:195, Ajsivinac Sian et al. 2004:162).5

3 All six word orders that are logically possible are indeed allowed in many of the Mayan languages, including Kaqchikel (England 1991, García Matzar & Rodríguez Guaján 1997:333).

4 Among the six grammatically allowed word orders in Kaqchikel, SVO is most frequently used. It has been suggested that this is due to the influence of Spanish (Maxwell & Little 2006), but the fact that all six word orders, including SVO, appear in sixteenth-century Kaqchikel texts shows that SVO was used before the language had contact with Spanish (Rodríguez Guaján 1989, quoted in England 1991, García Matzar & Rodríguez Guaján 1997:334).

5 Note also that results of a word-order acquisition study in Kaqchikel (Sugisaki et al. 2012) suggest that Kaqchikel-speaking three-year-old children know that VOS is the unmarked order in their language. Also, Pye (1992) showed that in K’iche’, a Mayan language closely related to Kaqchikel, children acquire the VOS order early.
According to England (1991:480), sentences with the above-mentioned three types of word orders have the syntactic structures shown in 5 (see also Aissen 1992, Tada 1993, Coon 2010, and Preminger 2011).

(5) ORDER DERIVATION

<table>
<thead>
<tr>
<th>Word Order</th>
<th>Derivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>VOS</td>
<td>[VOS]</td>
</tr>
<tr>
<td>VSO</td>
<td>[[V __ S] REORDERED O]</td>
</tr>
<tr>
<td>SVO</td>
<td>[REORDERED S [VO __ ]]</td>
</tr>
</tbody>
</table>

Aissen (1992) has proposed more elaborate syntactic structures for Mayan sentences with these word orders, but her analysis agrees with England’s in that VSO and SVO word orders are associated with more complex syntactic structures than VOS word order (see also Coon 2010 and Preminger 2011).⁶

Given this feature, the following predictions can be made about processing load in the comprehension of Kaqchikel sentences: if the preference for SO word order shown by speakers of SO languages is mainly caused by the syntactic structure of the individual language, as suggested by individual grammar theory, VOS sentences should have a lower processing load than VSO or SVO sentences in Kaqchikel. If, by contrast, SO triggers a lower processing load than OS regardless of the basic word order of the individual grammar, as suggested by universal cognition theory, then Kaqchikel VOS sentences should create a greater processing load than the other two word orders. The field-based psycholinguistic study described in §4 tested these predictions.

It should be noted at this point that, as mentioned in n. 4, even though VOS is Kaqchikel’s syntactically basic word order, it is the SVO order that is most frequently used in this language (England 1991:472, Rodriguez Guajan 1994:201, Maxwell & Little 2006:102, Kubo et al. 2012). According to Kubo and colleagues (2012), for example, of all the sentences with a transitive verb and nominal subject and object produced in their sentence-production experiment with a picture-description task, sentences with the SVO, VOS, and VSO orders constitute 74.4%, 24.2%, and 1.4%, respectively. In fact, not just in Kaqchikel but in many other Mayan languages also, word orders in which subjects are preposed appear more frequently than the syntactically determined basic word order. Therefore, it has been pointed out that the ‘syntactically determined word order’ from the standpoint of syntactic complexity needs to be distinguished from the ‘pragmatically determined word order’, commonly used for pragmatic purposes, when examining the ‘basic word order’ of Mayan languages (Broady 1984, England 1991). In the psycholinguistic literature, it has been reported that there are cases where the frequency with which words and sentence structures appear affects the sentence-processing load (e.g. Trueswell et al. 1993, MacDonald et al. 1994). That is, speakers of a language are more proficient in sentence structures and words that are used frequently, and they are more likely to process these with speed and accuracy. It is thus interesting to observe how the production frequency influences sentence processing in Kaqchikel. We return to this issue in §5.

4. EXPERIMENT.

4.1. PARTICIPANTS. Sixty-one native speakers (twenty-nine females, thirty-two males) of Kaqchikel participated in the experiment, which was carried out in Guatemala. The place of origin and residence of the participants were distributed evenly throughout a

---

⁶ We assume that syntactic structure X is more syntactically complex than syntactic structure Y if X contains more (terminal and nonterminal) nodes (Pritchett & Whitman 1995, O’Grady 2007, Koizumi & Tamaoka 2010). We may also assume that dependency across a discourse entity increases syntactic complexity (Gibson 2000, Hawkins 1990, 1994, 2004).
wide range of the central Guatemala highlands, without any concentration on a particular region. Since there is considerable dialectal and idiolectal variation among Kaqchikel speakers, only the data of twenty-two speakers (ten females, twelve males), who had over 80% accuracy in sentence processing (of the thirty-six target items and thirty-six implausible items, explained in §4.2), were used in the final analysis. The speakers ranged in age from twenty to sixty-two years. The average age was thirty-six years, five months, with a standard deviation of thirteen years, four months.7

4.2. Stimuli. Semantically natural, grammatical transitive sentences were arranged into each of three word orders (VOS, VSO, SVO), as shown in 6. Thirty-six sets were created in this way, for a total of 108 target sentences.8 All of the target sentences were so-called irreversible sentences, with a definite animate subject, definite inanimate object, and action verb.

(6) a. VOS: X-Ø-u-chöy ri chäj ri ajanel.
   compl.-B.3sg-A.3sg-cut det pine.tree det carpenter
   ‘The carpenter cut the pine tree.’
   b. VSO: Xuchöy ri ajanel ri chäj.
   c. SVO: Ri ajanel xuchöy ri chäj.

Additionally, thirty-six transitive sentences that were grammatical but not semantically natural were arranged in each of the three word orders. They were semantically implausible mostly due to selectional-restriction violations (e.g. #Xuch’äj ri kaq iq ’ri xta Selfa ‘Miss Selfa washed the air’). The seventy-two total sets, consisting of 216 sentences, were counterbalanced and then categorized into three groups according to word order. Sixty semantically plausible and implausible filler sentences were then added to each group.

The sentences were recorded by a male native Kaqchikel speaker. In order to create equal durations across the three word-order conditions, the time duration of each sentence was edited in Praat (Boersma & Weenink 2010) by slightly shortening some pauses between phrases. No particular order was edited significantly more heavily than the others. After the editing, all of the test items were judged as natural in terms of prosody by our native Kaqchikel consultants. The averages and standard deviations of time duration for word order are shown in Table 1.9 A one-way analysis of variance (ANOVA) showed no significant differences among the word orders in terms of time duration between the onset and offset of the sentence ($F(2,70) = 0.527, p = 0.592, \text{n.s.}$). Similarly, there was no significant difference in time duration between the onset of the sentence and the onset of the third phrase (i.e. S of VOS and O of SVO/VSO; $F(2,70) = 1.443, p = 0.243, \text{n.s.}$).

4.3. Method. A sentence-plausibility judgment task (e.g. Caplan et al. 2008) was administered using E-prime (version 2.0, Psychology Software Tools). In this task, the stimulus sentences were presented in a random order to the participants through headsets. The participants were asked to judge whether each sentence was semantically valid.

---

7 When all sixty-one participants were included in the analysis, no word-order preference or correlation between accuracy rates and word-order preference was found due to large variability.
8 VSO-ordered sentences were included in the test for a comparative standard. In other words, VSO is neither the syntactically canonical order nor the order most frequently used. The production frequency of VSO is third after SVO and VOS. Note also that VSO is minimally different from VOS in that only the order of S and O are reversed, whereas SVO diverges from VOS not only in the order of S and O but also in the relative order of S and V.
9 Duration was measured in Praat and is shown to two decimal places.
plausible and to push a YES button (plausible sentence) or NO button (implausible sentence) as quickly and accurately as possible. The time from the beginning of each stimulus sentence until the button was pressed was measured as the reaction time.

4.4. Data compilation for analysis. Among the thirty-six sets of semantically plausible transitive sentences, only correctly judged items were analyzed. Answers that were given too quickly (500 ms and under) or too slowly (8000 ms and over) were recorded as missing values. Then, reaction times outside of 2.5 standard deviations at both the high and low ranges were replaced by boundaries indicated at plus and minus 2.5 standard deviations from the individual mean of each participant in each category. These procedures resulted in the loss of 0.3% of the data. The means and standard deviations of reaction times and error rates for the thirty-six sets of semantically plausible sentences in the three word orders are reported in Table 2.

<table>
<thead>
<tr>
<th>WORD ORDER</th>
<th>WHOLE SENTENCE (ms)</th>
<th>BEFORE THIRD PHRASE (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>VOS</td>
<td>3,002</td>
<td>469</td>
</tr>
<tr>
<td>SVO</td>
<td>3,006</td>
<td>468</td>
</tr>
<tr>
<td>VSO</td>
<td>3,001</td>
<td>470</td>
</tr>
</tbody>
</table>

Table 1. Mean durations of semantically plausible transitive sentences (‘yes’ items) for each word order ($M = \text{mean, } SD = \text{standard deviation; } n = 36$).

4.5. Statistical analysis. Statistical analyses were conducted on the basis of a linear mixed-effects (LME) model (e.g. Baayen 2008), which estimates the effects of the fixed variables that are of interest in a study over random effects that can be assumed to be randomly sampled from the population. In this study, the word order of Kaqchikel sentences was the fixed variable, and participant and item (i.e. stimuli sentence) were the random variables. PASW ver. 18.0J was used to conduct the analysis.

4.6. Results. An ANOVA of reaction times using an LME model showed a significant main effect of word order ($F(2,402) = 5.917, p < 0.01, \eta_p^2 = 0.029$). Multiple comparisons by the Bonferroni method revealed that VOS ($M = 3,403$ ms) was processed significantly faster than SVO ($M = 3,559$ ms, $p < 0.05$) and VSO ($M = 3,601$ ms, $p < 0.01$). No significant difference was found between SVO and VSO.

An LME ANOVA of error rates revealed a significant main effect of word order ($F(2,787) = 15.169, p < 0.001, \eta_p^2 = 0.037$). The post hoc test showed that no significant difference was observed between the error rates for VOS ($M = 10.61\%$) and SVO ($M = 7.58\%$), whereas the error rate for VSO ($M = 22.90\%$) was significantly higher than those for VOS ($p < 0.001$) and SVO ($p < 0.001$).

5. Discussion. The results of the sentence-processing experiment in Kaqchikel showed that VOS word order induced a lower processing load than the two SO word orders (SVO and VSO) for sentences presented in isolation. This supports the traditional analysis in Mayan linguistics that VOS is the syntactically basic word order of Kaqchikel.
More importantly, it is consistent with the prediction of individual grammar theory and contradicts the prediction of universal cognition theory. In other words, the preference for SO word order in sentence comprehension reported in previous studies on SO languages may not be universal. This study verifies for the first time that for speakers of an OS language, OS word order has a lower processing load. This finding does not deny the existence of universal reasons for the preference for SO, but it certainly demonstrates that grammatical factors of individual languages have a relatively greater influence on sentence-processing load.

According to the psycholinguistic literature on SO languages, there are three major factors that are generally considered, in individual grammar theory, to contribute to the lower processing load of syntactically basic word orders compared to other grammatically possible word orders: syntactic complexity, discourse-pragmatic requirements, and production frequency. First, the syntactically basic word order in a language, by definition, is associated with simpler syntactic structures than the other grammatically possible orders in that language (see n. 6). It is therefore less demanding in terms of working-memory load, and hence is easier to process (cf. Pritchett & Whitman 1995, Gibson 2000, Marantz 2005, Tamaoka et al. 2005). Second, the syntactically basic order can be felicitously used in a wide range of contexts, including the absence of any substantial context, whereas derived orders require a specific discourse context to be felicitous. For this reason, derived orders cause a higher processing load when their discourse-pragmatic requirements are not met, such as when they are presented out of context, as is the case in many processing experiments, including that of the present study (Kaiser & Trueswell 2004, Weskott et al. 2011). Finally, the syntactically basic order tends to be more frequently used than other orders. Since, other things being equal, more frequently used structures are processed faster and more accurately, the basic word order tends to be easier to process (Trueswell et al. 1993, MacDonald et al. 1994).

In Japanese, for example, sentences with the syntactically basic SOV word order have simpler syntactic structures than the corresponding sentences with the other grammatically possible word order, OSV (Hoji 1985, Saito 1985). SOV sentences can also be used in pragmatically neutral contexts, in contrast to OSV sentences, which are typically produced when the referent of the object is discourse-given (Kuno 1978, Imamura & Koizumi 2011). Finally, the production frequency of SOV is higher than that of OSV (97.2% vs. 2.8%, respectively, according to Imamura & Koizumi 2011). Together, these three factors seem to make SOV sentences easier to process than OSV sentences in Japanese (Imamura & Koizumi 2008a, b).

What is the case in Kaqchikel? In Kaqchikel, VOS is the syntactically basic word order, and therefore it is associated with simpler syntactic structures than SVO, VSO, or...
any other order. In terms of discourse-pragmatics, VOS can be used in various contexts, including a pragmatically neutral context, whereas SVO is frequently used in contexts where the subject is a topic (García Matzar & Rodríguez Guaján 1997:334, Tichoc Cumes et al. 2000:219–23, Ajsivinac Sian et al. 2004:178–80). VSO is employed mostly when the object is ‘heavy’ or ‘complex’ (England 1991:474, Rodríguez Guaján 1994: 203, but see also García Matzar & Rodríguez Guaján 1997:341). These syntactic and discourse-pragmatic factors presumably made the VOS sentences easier to process than the SVO and VSO sentences in the present experiment, which employed a sentence-plausibility judgment task with no specific context provided. As for the relationship between processing load and word-order frequency, however, Kaqchikel seems to be different from SO languages like Japanese.

As pointed out in §3, the production frequency of SVO is higher than those of VOS and VSO in Kaqchikel (SVO: 74.4%, VOS: 24.2%, and VSO: 1.4%, according to Kubo et al. 2012). The production frequency factor should, therefore, facilitate the processing of SVO compared to VOS and VSO. Restricting ourselves to VOS and SVO for the moment, the syntactic complexity and discourse-pragmatic factors, on the one hand, and the frequency of usage, on the other, presumably work in the opposite directions: the syntax and pragmatics favor VOS, whereas the frequency favors SVO. The former overwhelms the latter, resulting in the lower processing load of VOS. VSO is syntactically more complex than VOS and is less complex than SVO, because the movement (or binding) of the subject in SVO crosses (two discourse participants associated with) V and O, whereas the movement of the object in VSO crosses only one element, that is, S. The reaction times for VSO, however, were not significantly different from those for SVO in the current experiment. This is presumably due to VSO’s production frequency being lower than that of SVO. That is, the effects of syntax and frequency cancel each other out, yielding comparable processing loads for VSO and SVO. The presumed relationship among word order, syntactic complexity, discourse-pragmatic requirements, production frequency, and processing load is summarized in Table 3.

<table>
<thead>
<tr>
<th>FACTOR</th>
<th>VOS</th>
<th>SVO</th>
<th>VSO</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYNTACTIC COMPLEXITY</td>
<td>simple</td>
<td>complex</td>
<td>medium</td>
</tr>
<tr>
<td>DISCOURSE-PRAGMATICS</td>
<td>less restricted</td>
<td>restricted</td>
<td>restricted</td>
</tr>
<tr>
<td>PRODUCTION FREQUENCY</td>
<td>medium</td>
<td>high</td>
<td>low</td>
</tr>
<tr>
<td>PROCESSING LOAD</td>
<td>low</td>
<td>high</td>
<td>high</td>
</tr>
</tbody>
</table>

Table 3. Relationship among word order, syntactic complexity, discourse-pragmatic requirements, production frequency, and processing load in Kaqchikel.

A question naturally arises as to why SVO is more frequently used than VOS in Kaqchikel, despite the fact that SVO is not the syntactically basic word order and is more difficult to process than VOS. There are three conceivable reasons. The first has to do with the head-marking nature of the language. As mentioned earlier, Kaqchikel is a head-marking language that exhibits subject and object agreement markers on the verb. The verbal complex of a transitive sentence [Aspect-B-A-Verb stem] contains the information about the person and number of the subject and object. It has been shown in other languages (e.g. English) that information about the verb (e.g. selectional restrictions) can immediately be used to facilitate the processing of the subsequent region (Trueswell et al. 1993, Altmann & Kamide 1999). Therefore, in Kaqchikel, having a verbal complex in sentence-initial position may be advantageous in that it helps de-
velop predictions about the upcoming subject and object, rendering the processing of the subsequent portions of the sentence easier. From the perspective of production, in contrast, verb-initial word orders may be more disadvantageous than nominal-initial orders such as SVO in Kaqchikel. This is because, in order to initiate a sentence with a verbal complex, conceptual and grammatical information about the subject and object must have been activated and processed to a certain degree, prior to the beginning of the utterance. Again, in other languages, it has been shown that the complexity of the sentence-initial phrase is correlated with the time required to initiate the utterance (e.g. Smith & Wheeldon 1999), and that latencies are shorter for subject-verb utterances than for verb-only utterances (Lindsley 1975). For this reason, therefore, SVO may be less demanding than the verb-initial orders for Kaqchikel speakers, and hence it is produced more frequently than VOS and VSO. It is thus important to test in future research whether production latencies are indeed shorter for SVO than VOS.

A second possible factor for the preference of SVO in sentence production, related to the first, is concerned with similarity-based competition. Gennari and colleagues (2012) argue that when there is a temporal overlap in the planning of two conceptually similar nouns, the similarity leads to interference between the semantic information of the nouns. As a result, when the concept of one noun is activated, the concept of the other noun is inhibited, and the latter noun is mentioned away from the initially activated noun, or simply omitted in the sentence. Moreover, the effect of conceptual similarity interacts with language-specific grammatical constraints, and the actual instantiation may vary across languages. Kubo and colleagues (2012) examined how similarity-based competition influences speakers’ choices of sentence patterns in Kaqchikel. The production of VOS sentences is interesting because the most accessible element, an animate agent noun usually realized as the subject, must be retained in memory until the end of the sentence, and hence it potentially competes with other elements. If similarity-based competition arises between the subject and object in Kaqchikel, one of them must be realized away from the other. Since the object usually follows the verb in Kaqchikel, the increase in competition would lead to the decrease of VOS word order. Kubo and colleagues conducted two picture-description experiments to verify this prediction. In the first experiment, the animacy of the patient noun was manipulated (human, animal, inanimate object) such that similarity between the agent (human) and patient varied across conditions. The results showed that VOS sentences were produced more often with an inanimate patient than with an animal or human patient, as predicted by similarity-based competition. In the second experiment, the researchers examined the effect of an agreement morpheme on the verb by changing the number of the object noun. The results replicated the overall patterns of the first experiment. That is, VOS sentences were produced more often with an inanimate patient than with a human patient, even when the number of the subject was different from the number of the object. This indicates that ambiguity resolution is not the most influential factor in the choice of sentence pattern in Kaqchikel (cf. Skopeteas & Verhoeven 2009). Both of these results together indicate that native Kaqchikel speakers seem to be sensitive to the competition caused by the similarity of noun concepts involved in an event described in a sentence, and they select the sentence pattern that best resolves competition between nouns with similar concepts.

Finally, the ‘saliency of subjects’ may contribute to the frequency of SVO. It has been observed in many languages that subjects tend to become topics of conversation more easily than other immediate sentence constituents, and topics tend to appear at the beginning of sentences. Indeed, in Mayan languages, constituents that appear before verbs are often interpreted as the topic of the utterance, and the observation that space
for a topic is syntactically secured before verbs is widely supported (England 1991, Aissen 1992, García Matzar & Rodríguez Guaján 1997:334). This means that, although VOS is syntactically the basic word order used in pragmatically neutral contexts and induces a lower processing load, SVO is used more frequently in conversation because subjects are often preposed as the topic (Tichoc Cumes et al. 2000:219–23, Ajsivinac Sian et al. 2004:178–80). This last point leads to the expectation that SVO sentences may be easier to process given an appropriate context, which needs to be tested in future research.

We have observed that in Kaqchikel, SO word order, which causes a higher processing load, is used more frequently than OS word order, which induces a lower processing load, arguably for pragmatic reasons. If this is true not only of Kaqchikel, but also of other OS languages, it would mean that OS languages are less economical in terms of linguistic performance. In SO languages, by contrast, the syntactically simple word order, which triggers a lower processing load, is also the most frequently used word order; subjects appear in front as a topic, and thus these languages are more economical. For example, in English, word order remains fixed as SVO regardless of whether the referent of the subject is the topic of the conversation. Similarly, in Japanese, in pragmatically neutral contexts the subject is marked with the nominative case marker and the object with the accusative case marker. When the referent of the subject is a discourse topic, the subject is preposed and marked with the topic marker. In either case, the word order is SOV. This is schematically shown in 7.

(7) a. [S-nom O-acc V]
   b. [S-top [ _ O-acc V]]

The production frequency of [S-top [ _ O-acc V]] is several times higher than that of [S-nom O-acc V] (Imamura & Koizumi 2011), parallel to SVO and VOS in Kaqchikel. However, there is a crucial difference between Japanese and Kaqchikel. In Japanese, both [S-nom O-acc V] and [S-top [ _ O-acc V]] have SOV word order, and the difference in syntactic complexity between them is minimal, the topicalization in 7b being string-vacuous. In fact, Sato and Koizumi (2012) observed that Japanese speakers processed [S-top [ _ O-acc V]] as fast as [S-nom O-acc V]. Although the economic efficiency of linguistic performance must still be carefully examined in detail in future research, the reason why over 95% of the world’s languages are SO and so few have OS as the basic word order, as well as the reason why OS languages are relatively unstable, occasionally shifting to SVO/VSO (Gell-Manna & Ruhlen 2011), might be explained, in part, by the fact that the syntactically determined word order does not coincide with the pragmatically determined word order in OS languages.

To summarize, the present study suggests that Kaqchikel speakers prefer VOS to SVO in sentence comprehension despite the fact that SVO is more frequently produced than VOS. A justifiable criticism of our claims is that a single experiment with a single type of stimuli and a single task is not sufficient to firmly establish that VOS is indeed easier to process than the other orders in Kaqchikel. Admittedly, there are a number of shortcomings. For example, the test items of the form VNN are syntactically ambiguous between VOS and VSO, and can only be disambiguated semantically by relying on the animacy of the subject and object. It may be necessary to test whether the core results can be replicated with structurally unambiguous sentences. Furthermore, we em-

---

13 A focused element also occurs preverbally in Mayan languages, sometimes simultaneously with a topic (Aissen 1992, García Matzar & Rodríguez Guaján 1997:337, 341; see also Stiebels 2006, Preminger 2011). This also seems to contribute to the high frequency of SVO in Kaqchikel and other Mayan languages.
ployed reaction time in the sentence-plausibility judgment task to estimate processing load. While reaction time is one of the standard indices frequently used to measure task difficulty in psycholinguistic studies, it may be desirable to conduct follow-up studies with a different task and/or index (e.g., an ERP experiment with a picture-sentence matching task). The relationship between comprehension and production, as well as contextual effects on comprehension/production, also needs to be investigated in Kaqchikel and other OS languages. We hope to address these and related issues in future research.

6. Conclusion. The results of the sentence-processing experiment showed that VOS, which is the syntactically determined basic word order, has a lower processing load than SVO and VSO for Kaqchikel speakers. This revealed that the preference for SO word order in sentence comprehension observed in previous studies of SO languages may not be universal; rather, processing load in sentence comprehension is greatly affected by the syntactic nature of individual languages. Further, in Kaqchikel, SVO (one of the SO word orders) has a higher production frequency than VOS, which is the syntactically determined basic word order. That is, the most frequently used word order in Kaqchikel is one that triggers a relatively higher processing load; hence, the language may be less optimal in this respect in terms of linguistic performance. If this Kaqchikel phenomenon is also found in other OS languages, that might be one of the reasons why few OS languages exist in the world. Future studies should investigate processing load and syntactic frequency in other OS languages to test this possibility.

REFERENCES
Ajxivinac Sian, Juan Esteban; Lolmay Pedro Oscar García Mázpar; Martín Chacac Cutzal; and Ixkusamil Estela Alonzo Guajan. 2004. Gramática descriptiva del idioma maya Kaqchikel: Rutzijoxik rucholik ri Kaqchikel ch’ab’al. Guatemala City: Academia de las Lenguas Mayas de Guatemala, Comunidad Lingüistica Kaqchikel.
Brannigan, Holly P.; Martin J. Pickering; and Mikihiro Tanaka. 2008. Contributions of animacy to grammatical function assignment and word order during production. Lingua 118.117–89.
Broady, Jill. 1984. Some problems with the concept of basic word order. Linguistics 22.711–36.
Brown, R. McKenna; Judith M. Maxwell; and Walter E. Little. 2006. ¿La útz awäch?: Introduction to Kaqchikel Maya language. Austin: University of Texas Press.
Caplan, David; Evan Chen; and Gloria Waters. 2008. Task-dependent and task-independent neurovascular responses to syntactic processing. Cortex 44.257–75.
Coon, Jessica. 2010. VOS as predicate fronting in Chol Mayan. Lingua 120.345–78.


GREWE, TANJIA; INA BORNKESSEL-SCHLESEWSKY; STEFAN ZYSSET; RICHARD WIESE; D. YVES VON CRAMON; and MATTHIAS SCHLESEWSKY. 2007. The role of the posterior temporal sulcus in the processing of unmarked transitivity. NeuroImage 35.343–52.


Kim, Jungho; Masatoshi Koizumi; Naho Ikuta; Yuichiro Fukumitsu; Naoki Kimura; Kazuki Iwata; Jobu Watanabe; Satoru Yokoyama; Shigeru Sato; Kaoru Horie; and Ryuta Kawashima. 2009. Scrambling effects on the processing of Japanese sentences: An fMRI study. *Journal of Neurolinguistics* 22:151–66.


Roesler, Frank; Thomas Pechmann; Judith Streb; Brigitte Roeder; and Erwin Hennighausen. 1998. Parsing of sentences in a language with varying word order: Word-by-word variations of processing demands are revealed by event-related brain potentials. *Journal of Memory and Language* 38:150–76.


Tichoc Cumes, Rosalino; Juan Esteban Ajsivinac Sian; Lolmay Pedro Oscar García; Ixchel Carmelina Espantay; Martín Chacach Cutzal; and María Estela Alonso Guaján. 2000. *Runuk’ul pa rub’eyal rutz ’ib’axik ri Kaqchikel ch’ab’ ál: Gramática normativa del idioma Maya Kaqchikel*. Chimaltenango: Comunidad Lingüística Kaqchikel de la Academia de las Lenguas Mayas de Guatemala.


---

Koizumi
Department of Linguistics
Tohoku University
27-1 Kawauchi Aoba-ku
Sendai 980-8576, Japan
[koizumi@sal.tohoku.ac.jp]

[Received 15 April 2012; revision invited 19 September 2012; revision received 1 February 2013; accepted 4 October 2013]