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

Authors' names and affiliations:

Masatoshi Koizumi	Tohoku University
Yoshiho Yasugi	National Museum of Ethnology
Katsuo Tamaoka	Nagoya University
Sachiko Kiyama	National Center for Geriatrics and Gerontology
Jungho Kim	Tohoku University
Juan Esteban Ajsivinac Sian	Comunidad Lingüística Kaqchikel
Lolmay Pedro Oscar García Mátzar	Comunidad Lingüística Kaqchikel

Authors' e-mail addresses:

Koizumi	koizumi@sal.tohoku.ac.jp
Yasugi	yasugi@idc.minpaku.ac.jp
Tamaoka	ktamaoka@lang.nagoya-u.ac.jp
Kiyama	skiyama@ncgg.go.jp
Kim	kjungho7@gmail.com
Ajsivinac Sian	ajtzibsyan@yahoo.com
García Mátzar	lolmay07@gmail.com

First author's mailing address:
Masatoshi Koizumi
Department of Linguistics
Tohoku University
27-1 Kawauchi Aoba-ku
Sendai 980-8576
JAPAN
koizumi@sal.tohoku.ac.jp

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

Abstract:
The processing load of sentences with three different word orders (i.e. 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 confirmed 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 is not universal; rather, processing load in sentence comprehension is greatly affected by the syntactic nature of individual languages.*

* 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 eighth Workshop on Altaic Formal Linguistics, the second Formal Approaches to Mayan Linguistics, and the sixth Formal Approaches to Japanese Linguistics, as well as at the ZAS (Centre for General Linguistics, Berlin), MIT, Mie University, Hokkaido University, and Kwansei Gakuin 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, Jaklin Kornfilt, Alec Marantz, Shigeru Miyagawa, Wayne O'Neil, Umut Özge, David Pesetsky, Colin Phillips, Uli Sauerland, Peter Sells, Shigeo Tonoike, Hubert Truckenbrodt, 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 to MK) from the Japan Society for the Promotion of Science.

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

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

1. INTRODUCTION

It is known that 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 and Meng 1999, Mazuka, Itoh, and Kondo 2002, Kaiser and 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 human cognitive features that are more universal. What we refer to as INDIVIDUAL GRAMMAR THEORY in this paper posits that a language's syntactically determined basic word order has a low processing load in comparison to other possible word orders, whereas 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 and 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, i.e. VSO and SVO. This suggests that the preference for SO in sentence comprehension is not 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 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, Itoh, and Kondo 2002, Imamura and 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 and 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 have found that the left inferior frontal gyrus is more activated 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 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 and Kluender 2003 and Hagiwara et al. 2007 for Japanese, Erdocia 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 and Waters 1999, Just and Carpenter 2001, Grodner and Gibson 2005, Santi and 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' (2000) 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, in the 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).[1] 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 and Bever 1982, Bock and Warren 1985, Hirsh-Pasek and Golinkoff 1996, Primus 1999, Branigan, Pickering, and Tanaka 2008, Bornkessel-Schlesewsky and Schlesewsky 2009). The universal cognition theory, therefore, leads to the expectation 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 the 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 the individual grammar theory, OS word orders should be processed faster than SO word orders in these languages. However, the universal cognition theory predicts that the opposite should 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, in Section 3.

3. KAQCHIKEL

Kaqchikel is one of the 21 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 K'iche', Q'eqchi', and Mam (Tay Coyoy 1996: 55, Brown, Maxwell, and Little 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].[2] An example is given in (1) below.

(1) *Y-e'-in-to'*
IC-B3pl-A1sg-help
'I help them.'

Since Kaqchikel is a pro-drop language, (1) functions as both independent speech and an independent sentence.

Like its ancestor language, Kaqchikel's syntactically determined basic word order is VOS, but SVO and VSO are also possible (Rodríguez Guaján 1994: 200, García Matzar and Rodríguez Guaján 1997: 333, Tichoc Cumes et al. 2000: 195, Ajsivinac Sian et al. 2004: 162).[3] According to England (1991), these word orders are derived from VOS through reordering rules, as schematically shown in (2).[4]

(2)

Order	Derivation
VOS	[VOS]
VSO	[[V ___ S] REORDERED O]

SVO [REORDERED S [VO ___]]

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

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 the individual grammar theory, VOS sentences should have a lower processing load than VSO or SVO sentences in Kaqchikel. On the other hand, if SO triggers a lower processing load than OS regardless of the basic word order of the individual grammar, as suggested by the 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 Section 4 tested these predictions.

It should be noted at this point that, even though Kaqchikel has VOS as its syntactically basic word order, it is the SVO order that is most frequently used in this language (England 1991: 472, Rodríguez Guaján 1994: 201, Maxwell and Little 2006: 102, Kubo et al. 2012). According to Kubo et al. (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 order constitute 74.4%, 24.2%, and 1.4%, respectively. In fact, not only in Kaqchikel, but also in many Mayan languages, word orders in which subjects are preposed appear more frequently than the syntactically determined basic word order. Therefore, it has been pointed out that "syntactically determined word order" from the standpoint of syntactic complexity needs to be distinguished from "pragmatically determined word order," commonly used for pragmatic purposes, when examining the "basic word order" of Mayan languages (Broady 1984, England 1991). In psycholinguistic literature, it has been reported that there are cases where the frequency at which the words and sentence structures appear affects the sentence processing load (e.g. Trueswell, Tanenhaus, and Kello 1993). That is, speakers of the 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 will return to this issue in Section 5.

4. EXPERIMENT

4.1 PARTICIPANTS

Sixty-one native speakers (29 females, 32 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 wide range of central Guatemala highlands, without any concentration on a particular region. As there is considerable dialectal and idiolectal variation among Kaqchikel speakers, only the data of 22 speakers (10 females, 12 males) who had over 80% accuracy for the sentence processing (of the 36 target items and 36 implausible items to be explained in Section 4.2) were used in the final analysis. The speakers ranged in age from 20 to 62 years. The average age was 36 years, 5 months with a standard deviation of 13 years, 4 months.[5]

4.2 STIMULI

Semantically natural, grammatical transitive sentences were arranged into each of three word orders (VOS, VSO, SVO), as shown in (3). Thirty-six sets, for a total of 108 target sentences, were created in this way.[6] All the target sentences were so-called non-reversible sentences, with a definite animate subject, definite inanimate object, and action verb.

(3) a. (VOS) *X-Ø-u-chöy* *ri* *chäj* *ri* *ajanel*
CP-B3sg-A3sg-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, 36 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 a selectional restriction violation (e.g. *#Xuch'äj ri kaq'ïq' ri xta Selfa*. 'Miss Selfa washed the air.'). The total of 72 sets, consisting of 216 sentences, were counterbalanced and then categorized into three groups according to word order. Further, 60 filler sentences were added to each group. The sentences were recorded by a male native Kaqchikel speaker. The length of time duration of each sentence was edited in Praat ver. 5.1.31 to make an equal duration across the three word order conditions by slightly shortening the duration of some pauses between phrases. No particular order was edited significantly more heavily than the others. After the editing, all 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 were as follows: M = 3,002 ms, SD = 469 ms for VOS, M = 3,006 ms, SD = 468 ms for SVO, and M = 3,001 ms, SD = 470 ms for VSO.[7] 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 [$F(2, 70) = 0.527$, $p = .592$, *ns*.].

4.3 METHOD

A sentence plausibility judgment task (e.g. Caplan, Chen, and Waters 2008) was

administered using E-prime ver.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 plausible and to push a YES button (correct sentence) or NO button (incorrect 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 36 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 36 sets of semantically plausible sentences in the three word orders are reported in Table 1.

<INSERT TABLE 1 ABOUT HERE>

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 fixed variables that are of interest in the study over random effects that can be assumed as being randomly sampled from the population. In this study, we assumed the word order of Kaqchikel sentences as a fixed variable, and participant and item (i.e. stimuli sentence) as 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 [$F(2, 401.854) = 5.917$, $p < .01$, $\eta_p^2 = .029$]. Multiple comparisons by the Bonferroni method revealed that VOS (M = 3,403 ms) was processed significantly faster than SVO (M = 3,559 ms) [$p < .05$] and VSO (M = 3,601 ms) [$p < .01$]. No significant difference was found between SVO and VSO.

An LME ANOVA of error rates revealed a significant main effect of word order [$F(2, 787.001) = 15.169$, $p < .001$, $\eta_p^2 = .037$]. Results of 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 [$p < .001$] and SVO [$p < .001$].

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 confirms the traditional analysis in Mayan linguistics that VOS is the syntactically basic word order of Kaqchikel (García Matzar and Rodríguez Guaján 1997: 333). More importantly, it is consistent with the prediction of the individual grammar theory and contradicts the prediction of the universal cognition theory. In other words, the preference for SO word order in sentence comprehension reported in previous studies on SO languages is not universal. This study verifies for the first time that for speakers of OS languages, 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.[8]

In the individual grammar theory, there are three major factors that are generally considered 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. It is therefore less demanding in terms of working memory load, and hence is easier to process. 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.[9] For this reason, derived orders cause a higher processing load when their discourse-pragmatic requirements are not met, e.g. when presented out of context, as is the case in many processing experiments, including that of the present study (see Kaiser and Trueswell 2004 and Weskott et al. 2011, among many others). Finally, the syntactically basic order tends to be more frequently used than other orders.[10] 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. In Japanese, for instance, sentences with the syntactically basic word order, i.e. SOV, have simpler syntactic structures than the corresponding sentences with the other grammatically possible word order, i.e. OSV (Hoji 1985, Saito 1985). SOV sentences can 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 and Koizumi 2011). The production frequency of SOV is higher than that of OSV (97.2 % vs. 2.8%, respectively, according to Imamura and Koizumi 2011). Together, these three factors seem to make SOV sentences easier to process than OSV sentences in Japanese (Imamura and Koizumi 2008a, 2008b).

How about in Kaqchikel then? 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 and Rodríguez Guaján 1997: 334, Tichoc Cumes et al. 2000: 219-223, Ajsivinac Sian et al. 2004: 178-180). VSO is employed mostly when the object is "heavy" or "complex" (England 1991: 474, Rodríguez Guaján 1994: 203, but see also García Matzar and Rodríguez Guaján 1997: 341). Both of the 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 we have pointed out in Section 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, therefore, should facilitate the processing of SVO compared to VOS and VSO. Restricting ourselves to VOS and SVO for the moment, therefore, the syntactic complexity and discourse-pragmatic factors, on the one hand, and the frequency of usage, on the other hand, presumably work in the opposite direction: 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 of the subject in SVO crosses (two discourse participants associated with) V and O, whereas the movement of the object in VSO only crosses one element, i.e. S. The reaction time for VSO, however, was not significantly different from that for SVO in the current experiment. This is presumably due to the fact that the production frequency of VSO is lower than that of SVO. That is, the effects of syntax and frequency cancel out each other, yielding comparable processing loads for VSO and SVO. The relationship among word order, syntactic complexity, discourse-pragmatic requirements, production frequency, and processing load is summarized in Table 2.

<INSERT TABLE 2 ABOUT HERE>

As we have seen, SVO is more frequently used than VOS in Kaqchikel. A question naturally arises as to why this should be the case, despite the fact that SVO is not the syntactically basic word order and is harder to process than VOS. There are three conceivable reasons. The first has to do with the head-marking nature of Kaqchikel. 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 on the verb (e.g. selectional restrictions) can immediately be used to facilitate the processing of the subsequent region (Trueswell, Tanenhaus,

and Kello 1993, Altmann and Kamide 1999). Therefore, in Kaqchikel, having a verbal complex in the sentence-initial position may be advantageous in that it helps develop predictions about the upcoming subject and object, rendering the processing of the subsequent portion of the sentence easier. In the case of sentence production, in contrast, verb-initial word orders in Kaqchikel may be more disadvantageous than nominal-initial word orders such as SVO. 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 and 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.

A second possible factor, related to the first, for the preference of SVO in sentence production is concerned with "disambiguation." Since Kaqchikel is a head-marking language, nouns are not marked as subjects versus objects. Therefore, in sentences where both subjects and objects are animate nouns, such as the one in (4), it is ambiguous which of the nouns is the subject and which is the object.

(4)	N-Ø-u- to'		ti	ala'	ti	xtän.
	IC-B3sg-A3sg- helps		CL	boy	CL	girl

In such cases, speakers of Mayan languages often prepose subjects before verbs to avoid ambiguity and/or similarity-based interference (Skopeteas and Verhoeven 2009). Traditionally, in Kaqchikel, when subjects were preposed before the verb, the verb came to have a special form, called an agent-focus antipassive. However, in contemporary speech, the alternation of the verb form is optional, and subjects can be preposed without changing the verb form (Tichoc Cumes et al. 2000: 222, Ajsivinac Sian et al. 2004: 180). It is assumed that the frequency of transitive, active sentences in SVO word order has thus increased.

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 and Rodríguez Guaján 1997: 334).[11] 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: 219-223, Ajsivinac Sian et al. 2004: 178-180). 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.

As observed above, 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. On the other hand, in SO languages, the syntactically simple word order, which triggers a lower processing load, is also the frequently used word order; subjects appear in front as a topic, and thus, they are more economical. For example, in English, regardless of whether or not the referent of the subject is the topic of the conversation, word order remains fixed as SVO. Similarly, in Japanese, an SOV language, the subject is marked with the nominative case marker, and the object with the accusative case marker, in pragmatically neutral contexts. 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 (5).

(5) a. [S-nom O-acc V]
b. [S-top [___ O-acc V]]

[S-nom O-acc V] vs. [S-top [____ O-acc V]] in Japanese seems to be parallel to VOS vs. SVO in Kaqchikel. [S-nom O-acc V] in Japanese and VOS in Kaqchikel are syntactically simple and typically used in pragmatically neutral contexts. [S-top [____ O-acc V]] and SVO are syntactically more complex and used in contexts where the subject is a topic. The production frequencies of [S-top [____ O-acc V]] and SVO are several times higher than those of [S-nom O-acc V] and VOS, respectively (See Imamura and Koizumi 2011 for Japanese). 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 (5b) being string vacuous. In fact, Sato and Koizumi (2011) observed that Japanese speakers processed [S-top [____ O-acc V]] as fast as [S-nom O-acc V]. They argue that this is because the syntactic complexity favors [S-nom O-acc V], whereas the production frequency facilitates processing [S-top [____ O-acc V]], and that these two factors cancel out each other. Turning back to the processing of VOS and SVO in Kaqchikel, we see that the syntactic complexity and production frequency work in the opposite direction in Kaqchikel as well, as we have mentioned above. However, the difference in syntactic complexity between VOS and SVO in Kaqchikel is much larger than that between [S-nom O-acc V] and [S-top [____ O-acc V]] in Japanese, because in Kaqchikel SVO sentences, V and O intervene between the preposed S and its original position. Thus, unlike in Japanese, in Kaqchikel, the syntactic complexity factor has a larger impact on sentence processing than the

frequency factor, resulting in a higher processing load for SVO, as we have argued above.

The economic efficiency of linguistic performance must still be carefully examined in detail in future research. However, the reason that over 95% of world languages are SO and few have OS as the basic word order, as well as the reason that OS languages are relatively unstable, occasionally shifting to SVO/VSO (Gell-Mann and 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.

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 is not 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 productive 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 phenomenon of Kaqchikel is found in other OS languages as well, it might be one of the reasons that 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

Aissen, Judith L. 1992. Topic and focus in Mayan. *Language* 68:43–80.

Ajsivinac Sian, Juan Esteban, Lolmay Pedro Oscar García Mátzar, Martín Chacach Cutzal, and Ixkusamil Estela Alonzo Guaján. 2004. *Gramática descriptiva del idioma maya Kaqchikel: Rutzijoxik rucholik ri Kaqchikel ch'ab'äl*. Academia de las Lenguas Mayas de Guatemala, Comunidad Lingüística Kaqchikel.

Altmann, Gerry T.M. and Yuki Kamide. 1999. Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition* 73: 247–264.

Coon, Jessica. 2010. VOS as predicate fronting in Chol Mayan. *Lingua* 120:345–378.

Baayen, R. Harald. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. New York: Cambridge University Press.

Bader, Markus and Michael Meng. 1999. Subject-object ambiguities in German embedded clauses: An across-the-board comparison. *Journal of Psycholinguistic Research* 28:121–143.

Bock, J. Kathryn and Richard K. Warren. 1985. Conceptual accessibility and syntactic structure in sentence formulation. *Cognition* 21:47–67.

Bornkessel-Schlesewsky, Ina and Matthias Schlesewsky. 2009. The role of prominence information in the real-time comprehension of transitive constructions: A cross-linguistic approach. *Language and Linguistics Compass* 3(1):19–58.

Branigan, Holly P., Martin J. Pickering, and Mikihiro Tanaka. 2008. Contributions of animacy to grammatical function assignment and word order during production. *Lingua* 118:117–189.

Broady, Jill. 1984. Some problems with the concept of basic word order. *Linguistics* 22:711–736.

Brown, R. McKenna, Judith M. Maxwell and Walter E. Little. 2006. *¿La ütz 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: 257–275.

Caplan, David and Gloria S. Waters. 1999. Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences* 22:77–126.

Dryer, Matthew S. 2005. Order of subject, object and verb. *The world atlas of language structures*, ed. by Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie, pp. 330–333. Oxford, UK: Oxford University Press.

England, Nora C. 1991. Changes in basic word order in Mayan languages. *International Journal of American Linguistics* 57:446–486.

Erdocia, Kepa, Itziar Laka, Anna Mestres-Missé, and Antoni Rodriguez-Fornells. 2009. Syntactic complexity and ambiguity resolution in a free word order language: Behavioral and electrophysiological evidences from Basque. *Brain and Language* 109: 1–17.

García Matzar, Lolmay Pedro and Pakal B'alam José Obispo Rodríguez Guaján. (1997)

Rukemik ri Kaqchikel chi': Gramática kaqchikel. Guatemala, Guatemala: Cholsamaj.

Gell-Manna, Murray and Merritt Ruhlen. 2011. The origin and evolution of word order. *Proceedings of the National Academy of Sciences of the United States of America* 108:17290–17295.

Gibson, Edward. 2000. Dependency locality theory: A distance-based theory of linguistic complexity. *Image, language, brain: Papers from the first Mind Articulation Project Symposium*, ed. by Alec Marantz, Yasushi Miyashita, and Wayne O' Neil, 95–126. Cambridge, MA: MIT Press.

Goldin-Meadow, Susan, Wing Chee So, Asli Özyürek, and Carolyn Mylander. 2008. The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences of the United States of America* 105:9163–9168.

Grodner, Daniel and Edward Gibson. 2005. Consequences of the serial nature of linguistic input. *Cognitive Science* 29:261–90.

Hagiwara, Hiroko, Takahiro Soshi, Masami Ishihara, and Kuniyasu Imanaka. 2007. A topographical study on the event-related potential correlates of scrambled word order in Japanese complex sentences. *Journal of Cognitive Neuroscience* 19:175–193.

Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

Hirsh-Pasek, Kathy and Roberta Michnick Golinkoff. 1996. *The origins of grammar: Evidence from early language comprehension*. Cambridge, Mass.: MIT Press.

Hoji, Hajime. 1985. *Logical Form constraints and configurational structures in Japanese.* Doctoral dissertation, University of Washington, Seattle.

Imamura, Satoshi, and Masatoshi Koizumi. 2008a. Bunshori ni okeru jyouhou kouzou to tougo kouzou no kougosayou ni tuite [On the interaction between information structure and syntactic structure in sentence processing]. *Proceedings of the 136th conference of the linguistics society of Japan*, 212–217.

Imamura, Satoshi, and Masatoshi Koizumi. 2008b. Bunrikai ni okeru jyouhou kouzou to tougo kouzou no kougosayou ga syoujiru taimingu ni tuite [On the time course of the interaction between information structure and syntactic structure in sentence comprehension]. *Proceedings of the 137th conference of the linguistics society of Japan*, 92–97.

Imamura, Satoshi and Masatoshi Koizumi. 2011. A centering analysis of word order in Japanese. *Tohoku Studies in Linguistics* 20: 59-74.

Just, Marcel Adam and Patricia A. Carpenter. 2001. A capacity theory of comprehension: Individual differences in working memory. *Cerebral Cortex* 11:223–237.

Kaiser, Elsi and John C. Trueswell. 2004. The role of discourse context in the processing of a flexible word-order language. *Cognition* 94:113–147.

Kim, Jungho. 2012. Kankokugo kakimazegojyunbun-no puraimingu kooka [Priming effects in

scrambled sentences in Korean]. *Culture* 75, 141–156.

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(1):151–166.

Kubo, Takuya, Hajime Ono, Mikihiro Tanaka, Masatoshi Koizumi and Hiromu Sakai. 2012. How does animacy affect word order in a VOS language. Poster presented at the 25th Annual CUNY Conference on Human Sentence Processing, New York, March 16, 2012.

Kuno, Susumu. 1978. *Danwa no Bunpoo* [A Grammar of Discourse]. Tokyo: Taishukan.

Lewis, M. Paul (ed.). 2009. *Ethnologue: Languages of the world,* Sixteenth edition. Dallas, TX: SIL International. Online version: http://www.ethnologue.com/.

Lindsley, James R. 1975. Producing simple utterances: How far ahead do we plan? *Cognitive Psychology* 7:1-19.

Maxwell, Judith M. and Walter E. Little. 2006. *Tijonïk Kaqchikel oxlajuj aj - Curso de idioma y cultura maya Kaqchikel*. Editorial Junajpu.

Mazuka, Reiko, Kenji Itoh, and Tadahisa Kondo. 2002. Costs of scrambling in Japanese sentence processing. *Sentence processing in East Asian languages,* ed. by Mineharu Nakayama, 131–166. Stanford, CA: CSLI Publications.

Preminger, Omer. 2011. *Agreement as a fallible operation*. Doctoral dissertation, MIT.

Primus, Beatrice. 1999. *Cases and thematic roles.* Tübingen: Niemeyer.

Pye, Clifton. 1992. The acquisition of K'iche' Maya. *The crosslinguistic study of language acquisition*, volume 3, ed. by Dan Isaac Slobin, 221–308. Hillsdale, NJ: Lawrence Erlbaum.

Rodríguez Guaján, José Obispo. 1989. Orden básico del Kaqchikel del siglo XVI. Paper presented at the XI Taller Maya, Quetzaltenango, Guatemala.

Rodríguez Guaján, José Obispo. 1994. *Rutz'ib'axik ri Kaqchikel: Manual de redacción Kaqchikel*. Guatemala, Guatemala: Editorial Cholsamaj.

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* 38:150–176.

Saito, Mamoru. 1985. *Some asymmetries in Japanese and their theoretical implications*. Doctoral dissertation, MIT.

Sandler, Wendy, Irit Meir, Carol Padden, Mark Aronoff. 2005. The emergence of grammar: Systematic structure in a new language. *Proceedings of the National Academy of Sciences of the United States of America* 102:2661–2665.

Santi, Andrea and Yosef Grodzinsky. 2010. fMRI adaptation dissociates syntactic complexity dimensions. *NeuroImage* 51:1285–1293.

Sato, Yohei and Masatoshi Koizumi. 2012. A note on the factors affecting the processing load of scrambled and topicalized sentences. Ms. Tohoku University.

Sekerina, Irina A. 1997. *The syntax and processing of Russian scrambled constructions in Russian*. Doctoral dissertation, City University of New York.

Skopeteas, Stavros and Elisabeth Verhoeven. 2009. The interaction between topicalization and structural constraints: Evidence from Yucatec Maya. *The Linguistic Review* 26:239–259.

Slobin, Dan I. and Thomas G. Bever. 1982. Children use canonical sentence schemas: A crosslinguistic study of word order and inflections. *Cognition* 12:229–265.

Smith, Mark and Linda Wheeldon. 1999. High level processing scope in spoken sentence production. *Cognition* 73:205–246.

Stiebels, Barbara. 2006. Agent focus in Mayan languages. *Natural Language & Linguistic Theory* 24:501–570.

Sugisaki, Koji, Koichi Otaki, Noriaki Yusa, and Masatoshi Koizumi (2012) The Acquisition of word order and its constraints in Kaqchikel: A preliminary study. Paper presented at the 5th Generative Approaches to Language Acquisition North America, University of Kansas, October 13, 2012.

Tamaoka, Katsuo, Hiromu Sakai, Jun-ichiro Kawahara, Yayoi Miyaoka, Hyunjung Lim, and Masatoshi Koizumi. 2005. Priority information used for the processing of Japanese sentences: Thematic roles, case particles or grammatical functions? *Journal of Psycholinguistic Research* 34:273–324.

Tamaoka, Katsuo, Prabath Buddhika Arachchige Kanduboda, and Hiromu Sakai. 2011. Effects of word order alternation on the sentence processing of Sinhalese written and spoken forms. *Open Journal of Modern Linguistics* 1(2):24–32.

Tay Coyoy, Alfredo. 1996. *Análisis de situación de la educación maya en Guatemala.* Guatemala, Guatemala: Cholsamaj.

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

Trueswell, John C., Michael K. Tanenhaus and Christopher Kello. 1993. Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory and Cognition* 19:528–553.

Weskott, Thomas, Robin Hörnig, Gisbert Fanselow, and Reinhold Kliegl. 2011. Contextual licensing of marked OVS word order in German. *Linguistische Berichte* 225:3–18.

Notes

[1] There are more indications for the universal nature of the SO word order preference. First, Al-Sayyid Bedouin Sign Language arose within the last 70 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, based on the distribution of word order types in world languages, that the original word order in the ancestral language was SOV. Finally, Goldin-Meadow et al. (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.

[2] The following abbreviations are used in this paper. CL [Completive], IC [Incompletive], A [Set A ergative], B [Set B absolutive], 1 [First person], 3 [Third person], sg [Singular], pl [Plural], AF [Agent focus], DET [Determiner], CL [Classifier].

[3] Results of a word order acquisition study in Kaqchikel (Sugisaki et al. 2012) suggest that Kaqchikel-speaking 3-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.

[4] All six word orders that are logically possible are indeed allowed in many of the Mayan languages including Kaqchikel (England 1991, García Matzar and Rodríguez Guaján 1997: 333). SVO in particular is most frequently used. It has been suggested that this is due to the influence of Spanish (Maxwell and Little 2006), but the fact that all six word orders, including SVO, appear in 16th century Kaqchikel texts shows that SVO was used before the language had contact with Spanish (Rodríguez Guaján 1989 quoted in England 1991, García Matzar and Rodríguez Guaján 1997: 334).

[5] When all the 61 participants were included in the analysis, due to large variability, no word order preference or correlation between accuracy rates and word order preference was found.

[6] VSO ordered sentences were included in the test for 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 compared to 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.

[7] Sentence length is what is indicated by the software program Praat and is shown to two

decimal places.

[8] Given that VOS is preferred to SVO in Kaqchikel, one might wonder if, in ergative languages such as Kaqchikel, Absolutive-Ergative orders are preferred to Ergative-Absolutive orders, i.e., the Absolutive-Ergative word order preference is observed. It has been reported, however, that in Basque, an SOV ergative language with pro-drop, SOV (= Ergative-Absolutive-V) sentences are easier to process than corresponding OSV (= Absolutive-Ergative-V) sentences (Erdocia et al. 2009). This, together with the results of the present experiment, suggests that in ergative-absolutive languages as well as in nominative-accusative languages, the most preferred word order is the syntactically basic word order.

[9] This discourse-pragmatic requirement for derived word orders is related to their syntactic complexity: Since derived word orders are associated with syntactically complex structures and hence are harder to process, the language user would take the trouble to employ them only to achieve a specific goal.

[10] The higher frequency of the syntactically basic word order is also related to its syntactic complexity: Since the syntactically basic word order is associated with syntactically simpler structures, and hence easier to process than derived word orders, (other things being equal) it tends to be used more frequently.

[11] In Mayan languages, a focused element also occurs preverbally, sometimes simultaneously with a topic (Aissen 1992, García Matzar and Rodríguez Guaján 1997: 337 and 341, see also Stiebels 2006, Preminger 2011). This also seems to contribute to the high frequency of SVO in Kaqchikel and other Mayan languages.

Table 1

Reaction times and error rates of transitive sentences judged as semantically correct (*M* = mean, *SD* = standard deviation)

Word order	Reaction time (ms)		Error rate (%)	
	M	*SD*	*M*	*SD*
VOS	3,403	673	10.61	30.85
SVO	3,559	663	7.58	26.51
VSO	3,601	674	22.90	42.10

Note: n = 22

Table 2

Relationship among word order, syntactic complexity, production frequency, and processing load in Kaqchikel

Factor	Word order		
	VOS	SVO	VSO
Syntactic complexity	simple	complex	medium
Discourse-pragmatics	less restricted	restricted	restricted
Production frequency	medium	high	low
Processing load	low	high	high