

## **Word-Stress Location and the Order of Subject and Verb:**

### **Preliminary Data Analysis**

Hisao Tokizaki

*Sapporo University*

**ABSTRACT.** This article argues that word-stress location and the position of subjects are correlated in the world's languages. An analysis of the data in Dryer (2013b) and Goedemans and van der Hulst (2013a, b) in the World Atlas of Language Structure Online (WALS, Dryer and Haspelmath 2013) shows that languages with right-hand word-stress are more likely to have verb-subject order than languages with left-hand word-stress.\*

**Keywords:** word order, typology, linearization, externalization, stress

### **1. Typology of subject positions and Externalization**

The variation of word order in the world's languages has been one of the most discussed issues in linguistic typology. Especially, typological studies have investigated the order of head and its complement, verb and object (VO/OV), adposition and noun phrase (P-NP/NP-P), and so on (see Song 2018: Ch. 10) for an overview of the literature). Compared to the head-complement order, the order of subject and verb (SV/VS) has not been paid much attention in linguistic typology.

In generative grammar, the subject of a clause is assumed to be the specifier of T (or some functional element); a basic sentence has the structure [<sub>TP</sub> Subj [<sub>T'</sub> T VP]]. In the minimalist program of linguistic theory (Chomsky 1995, 2012), which assumes that the word order parameter is an Externalization parameter, the order of subject and verb as well as the order of head and its complement are claimed to be determined by the linear order of the specifier of T and T' at the Externalization or at the interface between syntax and phonology. However, the mechanism of Externalization has not been made clear.

In this paper, I argue that word-stress location and the order of subject and verb are correlated in the world's languages. In Tokizaki (2011) and Tokizaki and Kuwana (2013), I argued that word-stress location is correlated with the order of a head and its complement

---

\* This article is based on my presentation at the 7th workshop on the phonological externalization of morphosyntactic structure held at Hokkaido University on September 14, 2018. I would like to thank the participants of the workshop, who were kind and brave enough to come to Sapporo just a week after a big earthquake in Hokkaido. This work was supported by JSPS KAKENHI (Grant Number 15H03213).

(e.g. verb and object). Here I discuss the possibility that stress-order correlation also holds in the case of subject-verb order. In Section 2, I briefly review some studies on subject-verb order in generative syntax and linguistic typology. In Section 3, I illustrate a typology of word stress location by Goedemans and van der Hulst (2013a, b). Section 4 presents an analysis of the data in Dryer (2013b) and Goedemans and van der Hulst (2013a, b) in the World Atlas of Language Structure Online (WALS, Dryer and Haspelmath 2013), which shows the correlation between word-stress location and the order of subject and verb. Section 5 concludes the discussion.

## 2. Previous studies on the order of subject and verb

### 2.1 Subject position in X-bar theory

Generative grammar, which has the X-bar schemata in (1), assumes that a head X (e.g. a verb) has its complement YP (e.g. its object) as its sister constituent as shown in (1a), while the subject has been assumed to be the Specifier (ZP) of a head X, i.e. T(ense) or Infl(ection), as shown in (1b).

- (1) a.  $X' = X \text{ YP}$  or  $\text{YP } X$   
 b.  $\text{XP} = \text{ZP } X'$  (or  $X' \text{ ZP}$ )

In (1b), the predicate-subject order ( $X' \text{ ZP}$ ) is parenthesized because this order is not very common in the world's languages. As the X-bar schemata, Chomsky (1986: 3) presents (2), "where  $X^*$  stands for zero or more occurrences of some maximal projection and  $X = X_0$ ."

- (2) a.  $X' = X X''^*$   
 b.  $X'' = X''^* X'$

The choices of order here are for English. Lightfoot (1991: 6) formulates the X-bar schemata as (3), where a pair of curly brackets shows an unordered set.

- (3) a.  $X' \rightarrow \{X \text{ or } X', \text{ (YP)}\}$   
 b.  $\text{XP} \rightarrow \{\text{Specifier, } X'\}$

The schemata in (1) AND (3) show that in generative grammar, the order of subjects and verbs as well as the order of a head and its complement are parameterized in languages, i.e. SV or VS. Baker (2001) calls this parameter the subject side parameter, which he argues ranks lower than the head directionality parameter: VS order is mostly seen in VO languages.

### 2.2 Subject position in the world's languages

Despite the parametric formulation of the subject position in (1) and (3), it is well known that subjects precede predicates in most of the world's languages, as shown in (2b). According to Dryer (2013b), 1,193 (79.7%) languages have SV order while 194 (13.0%) languages have VS order; 110 (7.3%) languages have no dominant order SV/VS, as shown in (4), where non-SV orders are in bold.<sup>1</sup>

(4) Order of S and V (Dryer 2013b)

<u>Order</u>	<u>Number of languages</u>	<u>Percentage</u>
SV	1,193	79.7%
<b>VS</b>	<b>194</b>	<b>13.0%</b>
<b>NDO</b>	<b>110</b>	<b>7.3%</b>
Total	1,497	100.0%

Although 79.7% of languages have SV order, we need to explain why 20.3% (=13.0%+7.3%) of languages may have VS order.

If we take into consideration the position of object in transitive clauses, there are seven word order patterns, as shown in (5), where V>S orders (verb precedes subject) are in bold.

(5) Order of S, O and V (Dryer 2013a)

<u>Order</u>	<u>Number of languages</u>	<u>Percentage</u>
SOV	565	41.0%
SVO	488	35.4%
<b>VSO</b>	<b>95</b>	<b>6.9%</b>
<b>VOS</b>	<b>25</b>	<b>1.8%</b>
<b>OVS</b>	<b>11</b>	<b>0.8%</b>
OSV	4	0.3%
NDO	189	13.7%
Total	1,377	100.0%

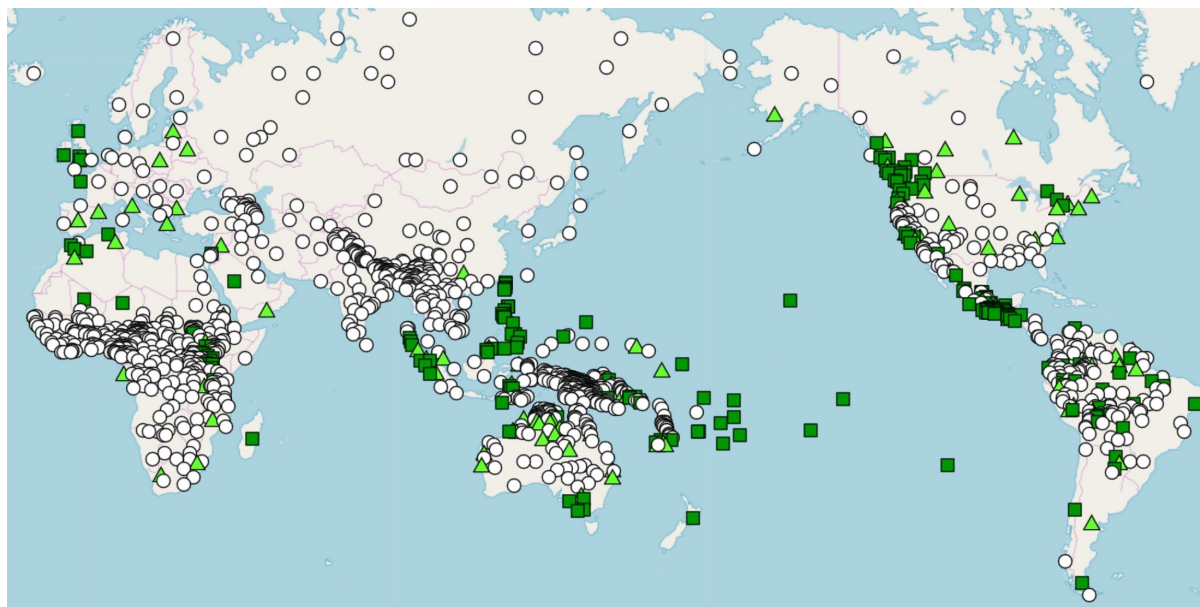
This table shows that in transitive sentences, subjects always precede verbs in 76.7% of languages (SOV 41.0%, SVO 35.4% and OSV 0.3%) while verbs precede subjects in 9.5% of languages (VSO 6.9%, VOS 1.8% and OVS 0.8%).

---

<sup>1</sup> Dryer (2013b) notes that “[i]n languages in which the position of subjects in intransitive clauses differs from that in transitive clauses, the map shows the order in intransitive clauses.”

It is interesting to see the geographical distribution of languages with SV, VS and no dominant order. The data in Dryer (2013b) and the World Atlas of Language Structure Online (WALS) give us the map in (6).<sup>2</sup>

(6) Order of Subject and Verb (Dryer 2013b, WALS 2013)

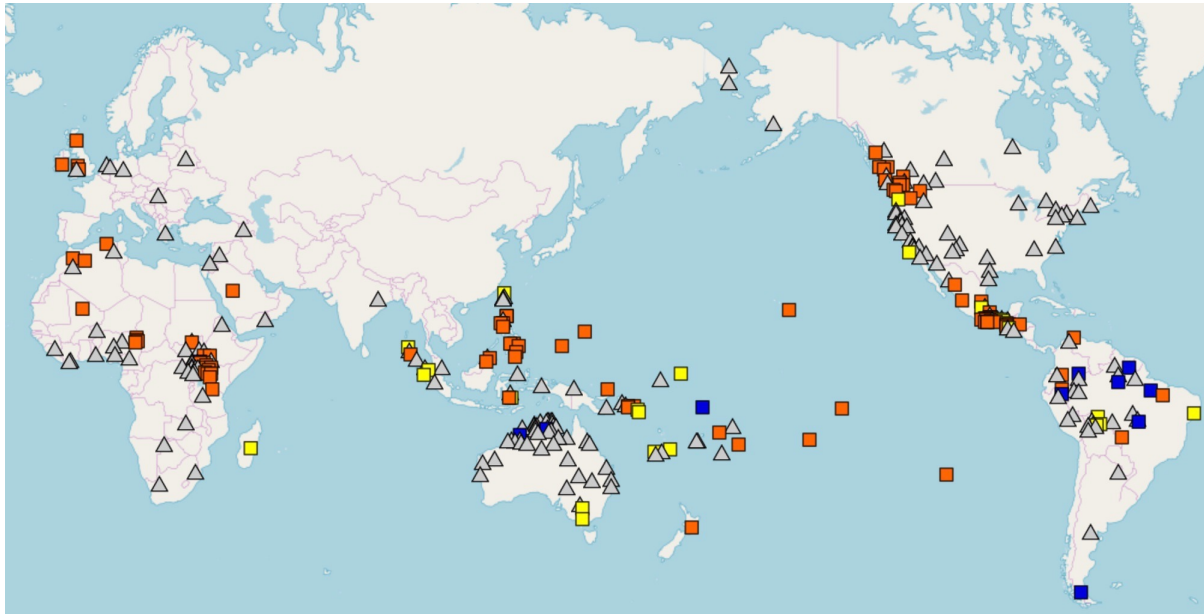


○	SV	1193
■	VS	194
▲	No dominant order	110

Dryer (2013b) points out that “the overall geographical distribution of VS order is rather similar to that of VSO and VOS languages.” The data in Dryer (2013a) and the World Atlas of Language Structure Online (WALS) give us the map of Order of Subject, Object and Verb in (7), where I show only languages that may have V>S order (VSO, VOS, OVS and no dominant order), omitting languages in which subject precedes verb (SOV, SVO and OSV), for the purpose of clarity.

<sup>2</sup> <https://wals.info>

(7) V>S orders and No Dominant order in Order of Subject, Object and Verb (WALS 2013, Dryer 2013a)



■	VSO	95
■	VOS	25
■	OVS	11
△	No dominant order	189

Comparing the map (6) Order of Subject and Verb with the map (7) Order of Subject, Object and Verb, it seems that the geographical distribution of VS (dark green squares) and no dominant order (SV/VS, light green triangles) in (6) are similar to the distribution of OVS (blue) and no dominant order (grey triangles) as well as VSO (orange) and VOS (yellow) in (7).<sup>3</sup>

### 3. A typology of word-stress location

One of the most comprehensive studies on the location of word stress in the world's languages is Goedemans and van der Hulst (2013a, b) and the subsequent database by Goedemans, Heinz and van der Hulst (2019). Here I use the data from Goedemans and van der Hulst (2013a, b) because they are available in the World Atlas of Language Structures (WALS) Online, which also includes the word order data from Dryer (2013b). Goedemans and van der Hulst (2013a, b) classify the system of word-stress location in the world's languages into two groups: fixed stress and weight-sensitive stress. Languages with a fixed

<sup>3</sup> For the rarity of OVS order, see Baker (2001: 167).

stress location have word stress in a fixed position in words, such as on the initial, second, third, antepenultimate, penultimate or ultimate syllable. The weight-sensitive stress system puts stress on a heavy syllable rather than on weak syllables in the stress window (possible stress locations) in a word; stress locations in the weight-sensitive stress system include left-edge (initial or second syllable), left-oriented (initial, second or third) right-oriented (antepenultimate, penultimate or ultimate), right-edge (penultimate or ultimate), and unbounded (anywhere in a word). Combining the fixed-stress locations feature (WALS #14, Goedemans and van der Hulst 2013a) and the weight-sensitive stress feature (WALS #15, Goedemans and van der Hulst 2013b) in Dryer and Haspelmath (2013), gives the table of word-stress locations in (8) (cf. Tokizaki 2019: 189).

(8) Word-stress locations and the number of languages

Initial 92	Second 16	Third 1	Antepenult 12	Penultimate 110	Ultimate 51
Left-edge (Initial or second) 37				Right-edge (Ultimate or penult) 65	
Left-oriented (One of the first three) 2			Right-oriented (One of the last three) 27		
Unbounded (Stress can be anywhere) 54					
Combined (Right-edge and unbounded) 8					
Not predictable 26					

The first row in (8) shows the fixed stress system, and the remaining rows show the weight-sensitive stress system.

In the next section, I argue that these word-stress locations correlate with the order of subjects and verbs.

#### 4. Correlation between word-stress location and the subject-verb order

Here I argue that languages with right-hand word-stress are more likely to have VS order (and no dominant order SV/VS) than languages with left-hand stress. If we count the number of languages with different word-stress locations (Goedemans and van der Hulst (2013a, b) and word orders (SV/VS/no dominant order) (Dryer 2013b), we get the table at (9) for languages with fixed stress location, and the table at (10) for languages with weight-sensitive stress.

## (9) Number of languages with fixed stress location

#Lang	Initial	Second	Antepenult	Penult	Ultimate	Sum
SV	<b>56</b>	5	5	<b>34</b>	24	124
NDO	<b>8</b>	1	2	<b>12</b>	4	27
VS	<b>5</b>	1	2	<b>17</b>	4	29
Sum	<b>69</b>	7	9	<b>63</b>	32	180

(NDO = no dominant order, i.e. SV/VS)

## (10) Number of languages with weight-sensitive stress

#Lang	L-edge	L-oriented	Unpredict	Unbound	Combin	R-oriented	R-edge	Sum
SV	<b>18</b>	1	14	29	4	14	<b>25</b>	105
NDO	<b>4</b>	0	1	4	0	4	<b>5</b>	18
VS	<b>2</b>	0	2	6	1	1	<b>8</b>	20
Sum	<b>24</b>	1	17	39	5	19	<b>38</b>	143

(L=left, R=right, Unpredict=unpredictable, Unbound=unbounded, Combin=combined)

First, let us compare the percentages of VS order in the major stress locations (left-hand and right-hand) shown in bold in (9) and (10). In languages with fixed stress (9), languages with penultimate stress are more likely to have VS than languages with initial stress (initial  $5/69=7.2\%$  vs. penultimate  $17/63=27.0\%$ ). In languages with weight-sensitive stress (10), languages with right-edge stress (penultimate or ultimate) are more likely to have VS than languages with left-edge stress (initial or second) (left-edge  $2/24=8.3\%$  vs. right-edge  $8/38=21.1\%$ ).

Next, let us compare the percentages of no dominant order SV/VS in the major stress locations. In (9), languages with penultimate stress are more likely to have no dominant order than languages with initial stress (initial  $8/69=11.6\%$  vs. penultimate  $17/63=19.0\%$ ). In (10), languages with right-edge stress are less likely to have no dominant order than languages with left-edge stress (left-edge  $4/24=16.7\%$  vs. right-edge  $5/38=13.2\%$ ). However, the difference between 16.7% and 13.2% is rather small.

Moreover, if we add languages with no dominant order to those with VS order in order to make a group of languages with non-SV order, the percentages again show that languages with penultimate stress are more likely to have non-SV order (VS or no dominant order) than languages with initial stress in the fixed stress system (initial  $(8+5)/69=18.8\%$  vs. penultimate  $(12+17)/63=46.0\%$ ), and that languages with right-edge stress are more

likely to have non-SV order (VS or no dominant order) than languages with left-edge stress (left-edge  $(4+2=)6/24=25.0\%$  vs. right-edge  $(5+8=)13/38=34.2\%$ ).

Furthermore, if we merge the fixed stress system (9) and the weight-sensitive stress system (10) and add up the total number of languages with left-hand stress (initial, second, third, left-edge and left-oriented) and the total of languages with right-hand stress (antepenult, penult, ultimate, right-oriented and right-edge), we get the chart at (11).

(11) Number of languages with fixed stress and weight-sensitive stress

#Lang	Left-hand stress	Right-hand stress	Sum
SV	80	102	182
NDO	13	27	40
VS	8	32	40
Sum	101	161	262

Again, this chart shows that languages with right-hand stress are more likely to have VS than languages with left-hand stress (left-hand stress  $8/101=7.9\%$  vs. right-hand stress  $32/161=19.9\%$ ) and that the former are more likely to have non-SV order (VS and no dominant order) than the latter (left-hand stress  $(13+8=)21/101=20.8\%$  vs. right-hand stress  $(27+32=)59/161=36.6\%$ ).

If we use a Chi-squared test for the correlation between word-stress location (left-hand vs. right-hand) and the order of subject and verb (SV, no dominant order and VS) in (11), the result is statistically significant with  $p < 0.05$ . Moreover, if we compare SV order with non-SV orders (VS or no dominant order), the result is statistically significant with  $p < 0.01$ . Thus, we can conclude that languages with right-hand stress are more likely to have VS order than languages with left-hand stress. Then, the order of subject and verb correlates with word-stress location in the world's languages.

## 5. Conclusion

I have argued that the order of subject and verb is correlated with word-stress location in the world's languages. An analysis of the data in Dryer (2013b) and Goedemans and van der Hulst (2013a, b) shows that languages with right-hand stress are more likely to have VS order than languages with left-hand stress. Then, we can conclude that word-stress location plays a role in deciding the order of subjects and verbs in addition to the order of a head and its complement (e.g. a verb and its object) (cf. Tokizaki 2011, Tokizaki and Kuwana 2013).



The next step is to consider why the correlation holds between word order and word-stress location. The basic hypothesis I would like to argue for is that word stress location projects up to phrasal stress location, which is in the most deeply embedded element in a structure (Cinque 1993). In intransitive sentences, the subject can be the most deeply embedded element receiving the main stress (i.e. [[SUBJECT ...] V] or [V [SUBJECT ...]]). Languages with right-hand stress (e.g. right-edge stress [WORD  $\sigma$   $\sigma$   $\underline{\sigma}$   $\underline{\sigma}$ ]) allow phrasal stress on the right-hand position in an intransitive sentence, where a subject can occur ([V [SUBJECT ...]]), but languages with left-hand stress do not. If this approach is on the right track, we can further substantiate the minimalist program, which assumes that word order is decided at externalization, not in syntactic computation.

### References

- Baker, Mark C. 2001 *The Atoms of Language: The Mind's Hidden Rules of Grammar*. New York: Basic Books.
- Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: The MIT Press.
- Chomsky, Noam. 1995. *The Minimalist Program*, Cambridge, MA: The MIT Press.
- Cinque, Guglielmo. 1993. A null theory of phrase and compound Stress. *Linguistic Inquiry* 24, 239-298.
- Dryer, Matthew S. 2013a. Order of Subject, Object and Verb. In Dryer, Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/81>, accessed on 2018-08-09.)
- Dryer, Matthew S. 2013b. Order of Subject and Verb. In Dryer, Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/82>, accessed on 2018-08-09.)
- Dryer, Matthew S. and Martin Haspelmath (eds.). 2013. *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (<http://wals.info/>, accessed on 2014-08-01.)
- Goedemans, Rob and Harry van der Hulst. 2013a. Fixed stress locations. In Dryer, Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/14>, accessed on 2018-08-09.)
- Goedemans, Rob and Harry van der Hulst. 2013b. Weight-sensitive stress. In Dryer, Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures*

- Online. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/15>, accessed on 2018-08-09.)
- Goedemans, Rob, Jeffrey Heinz and Harry van der Hulst. 2019. StressTyp2. <http://st2.ullet.net>.
- Lightfoot, David. 1991. *How to Set Parameters: Arguments from Language Change*. Cambridge, MA: The MIT Press.
- Song, Jae Jung. 2018. *Linguistic typology*. Oxford: Oxford University Press.
- Tokizaki, Hisao. 2011. The nature of linear information in the morphosyntax-PF Interface. *English Linguistics* 28, 227–257.
- Tokizaki, Hisao. 2019. Word stress, pitch accent and word order typology with special reference to Altaic. *The study of word stress and accent: theories, methods and data*, ed. by Rob Goedemans, Harry van der Hulst and Jeff Heinz. Cambridge: Cambridge University Press, 187-223.
- Tokizaki, Hisao and Yasutomo Kuwana. 2013. A stress-based theory of disharmonic word orders. In Theresa Biberauer and Michelle Sheehan (eds.) *Theoretical Approaches to Disharmonic Word Orders*, 190-215. Oxford: Oxford University Press.