

# A semantic and pragmatic explanation of harmony

Patrik Austin

To cite this article: Patrik Austin (2021): A semantic and pragmatic explanation of harmony, *Acta Linguistica Hafniensia*, DOI: [10.1080/03740463.2021.1987685](https://doi.org/10.1080/03740463.2021.1987685)

To link to this article: <https://doi.org/10.1080/03740463.2021.1987685>



© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 06 Dec 2021.



Submit your article to this journal [↗](#)




View related articles [↗](#)



View Crossmark data [↗](#)

# A semantic and pragmatic explanation of harmony

Patrik Austin 

University of Helsinki, Finland

## ABSTRACT

This paper introduces a semantically and pragmatically oriented typological generalisation, which is named the orientation principle. It entails that the position of connectives, as defined as a single lexical category including adpositions and conjunctions, provides an explanatory principle for a number of harmonic correlations in crosslinguistic data. A reanalysis of the data guided by this insight is proposed as an alternative to processing approaches.

**KEYWORDS** Typology; universals; harmonic correlations; syntax; semantics; pragmatics

## 1. Introduction

Song (2012) divides frameworks of cross-linguistic word order research into two main types. One branch consists of constituency parsing approaches and processing explanations of syntactic typology. A second, pragmatically oriented branch is also recognised although not given much consideration. Song agrees with Hawkins (1994, 240–241), who states that “pragmatics appears to play no role whatsoever” in cross-linguistic variation and is willing to limit the role of information structure, status, or packaging to within-language variation. While parsing approaches have made progress in the explanation of harmonic word-order correlations, e.g., VO&Prep versus OV&Postp (adposition type vs. order of object and verb) and VO&AuxV versus OV&VAux (order of main verb and auxiliary vs. order of object and verb), there is to date no known attempt to explain these on a pragmatic basis (Song 2012, 7).

There are, however, aspirations of providing proper explanations of some OV/VO tendencies on pragmatic principles. Haberland and Heltoft (1992) endorse the views of Bühler and Habermas, adding that Grice’s maxims of conversation can be interpreted as principles of rationality in communicative behaviour. Building on Greenberg (1963), they arrive at a two-principle explanation (*lightness* and *uniformity*) of the order of object pronoun and nominal object in relation to the verb (Haberland and Heltoft 1992).

**CONTACT** Patrik Austin  Patrik.Austin@helsinki.fi  Department of Languages, University of Helsinki, P.O. Box 24, FI-00014 Helsinki, Finland

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.  
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Vennemann's *natural serialization principle* on the other hand builds on logical expressions and semantics. It is one of the earliest attempts to explain why VO languages tend to have prepositions and OV languages tend to have postpositions and other features that cause the two types to appear to mirror each other (Song 2012). Vennemann argues that because languages change, it is natural for them to diachronically oscillate between the two possibilities of consistent serialisation (left-to-right or right-to-left; Vennemann 1974).

This paper adds a semantically and pragmatically oriented single-principle generalisation to be captured by a concept named *orientation*. Ways to expand the generalisation into a proper explanation will also be explored.<sup>1</sup>

## 2. The search for an explanation of harmony

The merit of Vennemann's (1974) natural serialisation lies in finding a way to account for the reflection symmetry between OV and VO languages. Unary predicates of first-order logic were employed, that is, expressions of the form  $P(x)$ . This was linked with a larger endeavour to describe language by means of formal logic, which eventually gave rise to formal semantics (cf. Vennemann *ibid.*).

Song points out that the predicate–argument structures in Vennemann's model represent what are more commonly called *dependencies* in nonlinear expression (Song 2012, 19). Natural serialisation research can in hindsight be seen as the beginning of what one might call a 'tweaking' approach to the problem of harmonic correlations – with the word 'tweak' referring to the meaning 'improve a mechanism or system by making fine adjustments to it' – especially in order for it to fit the data.

The question of what the data are, then, becomes vital in the race for the leading model. Although Vennemann and his colleagues wanted to show that the overall pattern is based on the logic of the predicate–argument relationship, it was initially unclear which part of the relationship, e.g., verb–auxiliary, was the head and which was the dependent. Adjustments were therefore made in conjunction with analysis of the data, until reaching a full picture.

Hawkins's (1983) critical examination of this research, however, revealed problems with the data itself, leading to the conclusion that the model was not accurate with respect to a scientifically correct view of typology. This turn led to two things that would dominate the enterprise in the following decades.

---

<sup>1</sup>I am grateful for guidance by Kees Hengeveld, Esa Itkonen, Hartmut Haberland, Lars Heltoft, and the anonymous reviewers.

First, model engineering was continued by linguists who were trained in the generative tradition. Examples of models which argue on the basis of tree structures include Hawkins (1994, 60, with further sources) and Dryer's (1992) Branching Direction Theory, which is based on concepts of phrasality and left/right branching as discussed in Government and Binding (Stowell 1981; Koopman 1984; Travis 1984, 1989; Chomsky 1988, 69).

Second, the concept of processing explanation (Hawkins 1983) was introduced to serve as a proper explanation of the dependency structure. The data were not considered to provide insight into an innate concept of syntactic structures. Whatever model was proposed, there would always be several exceptions. Newmeyer (1998, 2005, 105-113) argued that typological universals are irrelevant to generative approaches, which focus on Universal Grammar, because they have no psychological reality, i.e., a child has no access to them while learning a language natively.

Hawkins (2004, 1-14), on the other hand, argued for an alternative to Universal Grammar based on *ultimately* innate processing factors. His view is that language entails biases for reasons which are less directly biological. Thus, weak human universals, *processing preferences*, could be uncovered with an analysis of the natural grammars of the world.

Dryer (1992) is a milestone in the research because it provides particularly comprehensive sets of data analysis. The lead in explanation was, however, swiftly taken over by Hawkins's (1994) *early immediate constituents* principle (EIC) which argues that languages favour orderings where the immediate constituents (ICs) are arranged in a way that makes them appear as close to their mother phrase as possible. This means in practice that phrases of different length within the verb phrase (VP) are statistically ordered from shortest to longest in VO languages, and from longest to shortest in OV languages. It is hypothesised that this makes processing of the constituent structure quicker and easier for the human parser (Hawkins 1994).

It would seem plausible that one would be able to construct the correct head-dependency model as soon as all the relevant data are analysed correctly. This did not turn out to be so easy in practice due to asymmetries in the data. Like natural serialisation, various processing explanations could make a valid generalisation for reflection symmetries, but not so well for lack of symmetry. For EIC, a major issue was linked with the order of relative clause and its head noun. While the VO&NRel ordering is rather consistent,<sup>2</sup> OV languages are split between RelN and NRel (Song 2012).

---

<sup>2</sup>As the few exceptions, Dryer (2013a) mentions Mandarin and other varieties of Chinese, Bai (Tibeto-Burman, China) and Amis (Austronesian, Taiwan). Despite the rarity of VO&RelN, the ordering is stable throughout the accessible history of Chinese; a fact that has been considered problematic for typological explanation (Chappell and Peyraube 2007).

Hawkins (2004) returns with a revision of his theory where EIC is included in the new principle of *Minimize Domains* (MiD). There are quite a few different principles in the overall work, including *Maximize On-line Processing* (MaOP), which suggests that the function of *movement* is to ease gap structure processing. As a whole, MaOP conflicts with MiD, but this contradiction is exploited by Hawkins who demonstrates that when both MiD and MaOP are satisfied, as they are in VO languages, the model correctly predicts NRel. But, in OV languages, RelN satisfies MiD but not MaOP, and NRel satisfies MaOP but not MiD. Put together, the two principles correctly predict VO&NRel and OV&RelN/NRel (Hawkins 2004, 207).

These are all part of Hawkins's *Performance-Grammar Correspondence Hypothesis* (PGCH), which is indeed a triumph for the method. We do not know if the principles MiD and MaOP are anyhow real from a neurobiological point – Hawkins does not cite research based on them (see also Song 2012, 303) – but, as the model has gained more complexity, it now seems to work sufficiently well.

Song (2010, 2018) hails PGCH as the cutting edge of typological explanation but Song (2012) also provides a critical examination of problems faced by processing explanations. A persistent issue is that once a model is properly adjusted for OV/VO correlations, a chain reaction may cause false predictions to arise for other syntactic phenomena (see also Frey 2015). It is like the crooked house: as you fix one angle, a different angle gets twisted.

To sum up, some researchers like Song argue that Hawkins's PGCH is advanced enough to be considered the leader and the future of linguistics, while other syntacticians believe one should keep looking for answers elsewhere (Frey 2015; Abels 2015). The following sections will propose a different approach to the question as based on the principle of orientation of connectives.

### 3. Theoretical context

The theoretical foundation of this paper is in humanistic linguistics. Hjelmslev ([1943] 1969) reconstructs De Saussure's (1916) bilateral semiology as an algebraic system. Grammatical models based on Hjelmslev's device, such as Systemic Functional Linguistics, regard language as necessarily forcing a nonlinear meaning potential into a linear form (Davidse 1987; Butler 2003). This can be called the problem of linear language, or what Tesnière (1959) from his point of view aptly describes as an *antinomy* between the hierarchical and the linear form.

Because explanation in the present paper will set itself against models derived from a generative tradition of syntactic analysis, it may be useful to elucidate a difference in thinking about language between the semantic and pragmatic approach of this paper and the more biologically oriented

approaches.<sup>3</sup> Based on his literary exploration of the classical sources, Seuren (1998, 160–167) suggests that Louis Hjelmslev was the first to apply algebraic structures to general linguistics. Thus, although Hjelmslev ([1943] 1969) does not provide any tree diagrams, it is worth noting that Hjelmslev's *Prolegomena* exhibits noticeable similarities with generative grammar (Koerner 1978, 41f).

Hjelmslev ([1943] 1969, 6) argues for a purely semiological view of language, separating it from extra-linguistic factors, whether sociological, psychological or physiological. From the current perspective, one might add neurobiology to the list: we are essentially trying to understand the ways of language in its own terms, not as an expression of human neurophysiology.

The most important difference from generative grammar lies in Hjelmslev's notion that the content plane ('semantics') is organised according to the same principle as the expression plane ('syntax').<sup>4</sup> Thus, the relationship between content and expression can be described as geometric.<sup>5</sup> Davidse's point (1987, see above) was that the visible linear form is necessarily a matter of breaking a two-dimensional structure into a one-dimensional structure – one way or another. This is exactly the idea that will be exploited in the present paper to explain cross-linguistic harmonic tendencies. The prevalent 'orientation' of connectives in the data is derived from the non-linear semantic mapping.

Hjelmslev ([1943] 1969, 74, 76) explicitly states that a formal grammatical model should be used for typology to gain insight into the nature of language. An enterprise in this spirit is today also found in Functional Discourse Grammar (Hengeveld and Mackenzie 2008).

At the same time, following the Praguian tradition of functional explanation, languages are considered as a man-made tool for communication (Daneš 1987). The two ideas combined – algebraic and sociological – syntactic structures build on a mathematical system that can be thought of as existing independently of human mind; it is something that may be

---

<sup>3</sup>See Hawkins (1994, 2, 2014, 85–86) for different biological processing views.

<sup>4</sup>Hjelmslev starts his analysis by dissecting texts of a given language into the smallest distinct components on both planes (content ↔ expression). Components are then re-grouped (bottom-up) into inventories, which are in their turn re-grouped into higher-level inventories (syllables → words → sub-clauses → clauses → sentences → discourse – on the expression plane) until the analysis is exhausted. By means of this method, a grammar is compiled which, when the procedure is reversed, and generates (top-down) all correct combinations of dependencies (vertically) and components (horizontally) and thus all grammatical sentences of a language (Hjelmslev [1943] 1969, 9, 16–17, 42). On the content plane, the same method of dissection and vertical regrouping compiles a grammar that correspondingly generates all valid combinations of conceptual structure (Hjelmslev [1943] 1969, 69–71).

<sup>5</sup>By this, I mean that when a content–expression pairing is depicted, e.g., as two trees, these will have shapes that are not identical. The trees and their elements can be described as being in different geometric relationships with one other (cf. Guerrero et al. 2014).

uncovered, reconstructed, and exploited by intelligent beings. The social construction of language is assumed to occur largely unconsciously (Daneš 1987; Itkonen 2013).

We find two guidelines in the tradition of structural linguistics: a semantic orientation that considers linguistic structures as arising from the inner workings of the semiological system and a pragmatic orientation that considers them as arising from language use. Linguistic form has been regarded as reflecting *economy*; a compromise between easiness of expression and easiness of comprehension (Martinet 1955). An explanation based on simplicity and unambiguity will also be considered below.

#### 4. The orientation of connectives

The orientation principle is the typological generalisation that connectives are oriented to their semantic head. This is especially the verb, but it can also be a different part of the sentence. In other words, it means that connectives are aligned with their semantic head, but the term ‘orientation’ was chosen as the name of this principle because alignment is already in use in syntax research.

Connectives are defined here as a lexical category that encompasses adpositions and conjunctions. Although the orientation principle assumes a flatter structure than phrase structure grammar, it does recognise some phrase structure as relevant. Orientation of connectives means that the connective is placed – on a statistical basis – within the phrase (or word, in the case of affixes) – in a position where it is closest or ‘pointed’ to the semantic head. For example, in the sentence *John saw a man [with binoculars]*, the connective *with* is found within the phrase *with binoculars*. The semantic head of the whole phrase can be the verb (*saw*) or the object (*a man*). The orientation principle assumes that the connective is placed within the phrase as close as possible to the semantic head of the whole prepositional phrase, whichever that is; however, this is not to say that the whole phrase has to be placed as close as possible to the head.

This is compatible with the observation that VO languages like English tend to have phrase-initial connectives, and OV languages like Japanese tend to have phrase-final connectives. The same notion will be used to explain a number of harmonic correlations in cross-linguistic data. To exemplify how orientation works, consider the following English (VO) phrases:

- (1) a A homeless person is sleeping <on the bench.
- b A teacher <from Chicago is visiting.
- c They chose the bench <that was dry.

Here, the pointed brackets represent the dependency direction of the connective (*on, from, that*) back toward its semantic head (underlined). We see from the examples that the reference can at least be a verb or a noun. All of the above cases are ideal in that the connective actually physically *connects* the phrase with the reference. This is not a rule in real language situations because there is competition for the adjacent place. For example, in the sentence

(2) The chair told a joke <after the meeting,

the prepositional phrase *after the meeting* is pushed back by the object (*a joke*). The orientation principle does not explain why this occurs. It merely suggests that the canonical place is such that the connective is placed towards the semantic head of the phrase (as opposed to *The chair told a joke the meeting* <*after*).

Conveniently for our purposes, English has two different connectives for possessive expressions: the *of* preposition and the *'s* suffix. Their recursive use is consistent with the orientation principle with respect to the noun. Compare the following:

(3) a A friend <of a friend <of a friend.    b A friend's> friend's> friend.

(3b) is our first example of postpositional orientation which is statistically prevalent in OV languages. These make up approximately half of the world's languages when languages lacking a dominant order are excluded (cf. Cinque 2013, 70). An example of Japanese postpositional orientation is given in (4):<sup>6</sup>

(4) *Watashi ga kono yo de ichiban sukina basho wa*  
 I            SUBJ this world in most    favorite place TOP  
*daidokoro da to omo-u.*  
 kitchen    be.PRS QUOT think-PRS  
 'I think that my most favorite place in this world is the kitchen.'  
 (Yoshimoto 1991, 1)

Japanese also has postpositional subordinators such as the quotative particle in (4) and *nagara* ('while') in (5).

(5) *Kare wa tantan-to warai-nagara, ochitsui-te hanas-u hito datta.*  
 he TOP light-ly laugh-while calm-ly talk-PRS person be.PST  
 'He was a person who talked calmly, while laughing lightly.' (Yoshimoto 1991, 14)

<sup>6</sup>Examples (4) and (5) are from Yoshimoto's novel *Kitchin* ('Kitchen'), thanks to Hartmut Haberland (p.c.).



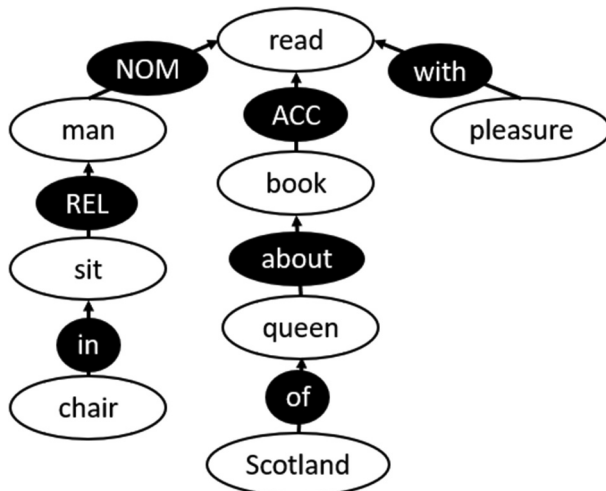
As the subordinator is attached to the verb, it is a matter of interpretation whether *nagara* is syntactically equivalent to the English conjunction *while*. This is a fundamental question for OV/VO typology because the mass-comparison of languages depends on the analysis of each individual language. We will accept here that postpositional suffixes may correspond to English conjunctions.

## 5. An explanation of orientation

The orientation principle is a typological generalisation: connectives are statistically oriented to their semantic head. The causes of orientation are currently not known, but two explanations will be sought in this section: a primarily semantic and a primarily pragmatic one.

From a semantic perspective, it is most conspicuous that the general pattern extracted from cross-linguistic data resembles maps or graphs used in formal semantics, especially a type that is neither exactly a phrase structure grammar nor a standard dependency grammar (cf. Koller 2015). The orientation principle predicts that the word order of the average language is organised in a way that is iconic with the graph below.

Figure 1 is read bottom-up for OV languages and top-down for VO languages. *Read* is the verb; the first column under it is the subject; the second column under *read* is the object; and the rest represents an optional adverbial. Thus, a full SVO reading is “AG man REL sit in chair



**Figure 1.** The average human language according to the orientation principle. A graph for the sentence *A man who is sitting in a chair reads a book about the queen of Scotland with pleasure*. Note that the agent (AG), the patient (PAT) and the relativiser (REL) are frequently unmarked.

read PAT book about queen of Scotland with pleasure”, and a full SOV reading is “chair in sit REL man AG Scotland of queen about book PAT pleasure with read.”

Thus, a semantically based explanation for the orientation of connectives is that syntax is derived from semantics. It is the linear form of the nonlinear conceptual mapping. Each permutation of the order of subject, object and verb is merely another possibility of linear organisation.

It needs to be noted that the structure of the semantic representation is not arbitrary. Figure 1 can also be expressed by means of set theory. When each role marker (connective) is considered as a set and each content word as an element within a set, the graph is organised in a logically consistent way, as in Figure 2.

This arrangement can indeed be argued to be logically consistent: if any of the variables denoting sets or elements was placed differently in the hierarchy, the meaning of the corresponding expression would also be different. We can name this type of explanation *structural*.

As for a functional or pragmatically based explanation of the orientation of connectives, a possible starting point is in examining how the arrangement of linear representation affects disambiguation of garden-path sentences, i.e., grammatically correct sentences that, when parsed word by word, may lead the reader or hearer to a false interpretation of the meaning of the sentence. The conjunction, when separating two predicate–argument structures, helps the hearer link each argument with the correct predicate. Compare the following examples. (6a) has a standard subclause-initial subordinator; (6b) uses *because* as a hypothetical clause-final subordinator.

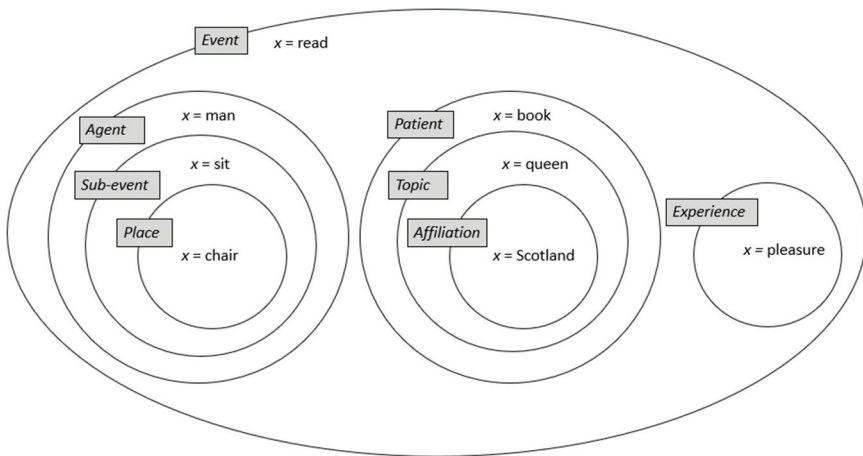


Figure 2. The logical organisation of the semantic representation from the point of set theory. Note that the main set (*Event*) is unmarked in the given example sentence (see Figure 1).

- (6) a I read *because* it makes me look wiser.  
 b \*I read it makes me look wiser *because*.

Although (6b) is ungrammatical in English, it seems hardly catastrophic considering all the variability among the world's languages: if such an ordering is normal in any given language, it merely suggests that the speakers of that language have learned to anticipate the sentence-final connective.

There are nevertheless two facts to consider. First, the beginning of (6b) *I read it . . .* gives rise to an initial misanalysis with *it* seemingly appearing as the object of the main clause. This is because the verb *read* can take two arguments. The question is not trivial: it is argued in machine translation research, for example, that syntactic and semantic processing should take place simultaneously to resolve lexical and structural ambiguities (Lytinen 1987). From this departure, it might be possible to argue for an economy explanation, as a reference to Martinet's (1955) concept,<sup>7</sup> because it suggests that a language entails a compromise between simplicity and clarity. (6a) and (6b) having exactly the same complexity, the difference from the current perspective is that (6b) is less economic because it has less clarity owing precisely to the displacement of the connective.

Second, this problem would seem to be statistically rare because SVO languages do not typically have subclause-final connectives as in (6b), but subclause-initial connectives as in (6a). This helps disambiguation whether or not it is the actual cause of the phenomenon.

As for prepositions, it is possible that their orientation, too, works for disambiguation. Alternatively, adpositional orientation could follow from the conjunctive orientation, which was explained above. Although a structural-functional approach offers an interesting starting point for the study of word-order universals, there do not appear to be studies to cite as either supporting or rejecting clarity issues as giving rise to syntactic solutions on a universal scale.

Some orderings are likely to make it easier for the hearer to perceive the intention of the speaker. To what extent it will be possible to argue for a specific processing preference is unknown. Languages tend to exhibit some variety, and English, too, has some postpositions. This may not necessarily lead to processing difficulty. For example, there does not appear to be evidence that it is more difficult for the brain to process a sentence like *we went to Spain two weeks ago*, with the postposition *ago*, than it is to process a sentence like *we went to Spain for two weeks*, with the preposition *for* (cf. e.g., Spivey, Joanisse, and McRae 2012).

---

<sup>7</sup>Related notions include Haiman's (1983) competing motivations of iconicity and economy, as well as economy versus faithfulness in Optimality Theory (McCarthy and Prince 1995); see also Vincentini (2003) for a history of the principle of economy.

It may be a more plausible explanation that word-order conventions are based on a consensus within the speech community, and a consensus is more easily reached when the connective links the attribute with its reference. Take the following examples with the prepositional phrase highlighted.

- (7) a My neighbour *at the White House* saw the President.  
 b My neighbour saw the President *at the White House*.

While there is no clear logical argument for the superiority of any of the above orderings, cross-linguistic data show that the post-verbal adjunct (7b) is more common in prepositional languages and that the pre-verbal adjunct (7a) is more common in postpositional languages (Dryer 1992). In English, conventionalised usage actually links the prepositional phrase to the subject and not the verb in (7a). The ultimate explanation might be that many people simply find it to make more sense this way.

## 6. Evidence and predictions

The idea that typology could conflate adpositions and conjunctions into a single category is taken up by Schmidtke-Bode (2009) who explains:

it has to be noted that it can be difficult to decide whether a particular marker is an adposition or a conjunction. As it turns out in more than a handful of languages, the division is not clear-cut, with many adpositions also regularly surfacing as conjunctions. English *after*, *before*, *since* and, importantly in our context, *to* all serve both functions depending on the larger construction they appear in. In fact, Huddleston (1984; 338–441) is not reluctant to suggest conflating prepositions and conjunctions into a single word class. (2009, 73–74)

In Schmidtke-Bode's study (2009, 80; an unbiased sample of purposive clauses), 59 out of 61 languages placed their subordinate conjunction either clause-initially or clause-finally. Conjunctions were most commonly found in VO languages, and there was a strong preference to place them in the clause-initial position. Thereby, they are expected to function frequently as bi-clausal connectives.

Regarding languages that use a subordinating affix rather than a subordinating conjunction, typically OV languages, 76 out of 96 use the affix as a connective between the subclause and the main clause. In 74 of the cases it appears subclause-finally. As for languages using adpositions as subordinators, 33 out of 39 use it as a connective between the sentences, either as a clause-final postposition (subclause-to-main clause) or as a clause-initial preposition (main clause-to-subclause; Schmidtke-Bode 2009, 79–80). Affixed subordination is found in Kewa (Engan; Papua New Guinea).

- (8) [*Ádo-la*] *pá-lua*.  
 [see-PURP] go-1SG.FUT  
 ‘I will go to see it.’ (Franklin 1971, 97)

As for the predictions, the orientation principle entails by definition that the order of object and verb correlates with adposition type (OV&Postp, VO&Prep), a well-established correlation pair. Conjunction type, on the other hand, has not been given equal attention in post-Greenbergian typology. Greenberg (1963) gave relatively little consideration to subordinators, and this possible oversight seems to have been inherited by Vennemann (1974) and the processing models that followed.

The data analyses that now serve as a standard reference (especially Dryer 1992 and later) have been organised with phrasal order in mind, not so much the connective, although Dryer (1992) includes interesting data on the place of the subordinator, whether clause-initial or clause-final. Based on a representative sample of forty languages, Diessel (2001) demonstrates that there is “a strong correlation between the ordering of main and adverbial clauses and the position of the adverbial subordinator” (Diessel 2001, 442). To be more precise, there are two major crosslinguistic ordering patterns. In languages where adverbial clauses have a final subordinator, adverbial clauses tend to precede the main clause. All 17 languages in Diessel’s sample that have adverbial clause before the main clause are OV languages. The second type is mixed with adverbial clauses appearing either before or after the predicate or main clause; these are for most part VO languages; and almost all such languages mark adverbial clauses with an initial subordinator (Diessel 2001, 442-443). In fact, Dryer (1992, 103) considers VO and clause-initial subordinator a strong correlation pair. These findings are taken as supporting the orientation principle. In VO languages the actual type of adverbial clause however plays a significant role in determining its place in the complex sentence as will be discussed in section 7 below.

The orientation principle has decisively less predictive power than generalisations based on phrasal length. This is because it can only make predictions for cases that include a connective in the first place. A difficulty for weight-based models is that generalisations which build on phrasal length tend to have too far-reaching consequences. As illustrated by the crooked house metaphor in section 2, one adjustment to fix an issue in one prediction may give rise to a different issue in a different prediction. An obstacle to surpass by any model is the question of non-correlation pairs. Here, the question is what data should be considered as correlative or non-correlative. One possible

interpretation is given by Cinque (2013) who lists *weak or non-tendencies* as follows, with corresponding examples from Dryer's (1992) noncorrelation pairs where applicable:<sup>8</sup>

- a. Adjectives with respect to noun (Adj&N; *tall + man*).
- b. Numerals with respect to noun (Num&N; *four + books*).
- c. Demonstratives with respect to noun (Dem&N; *that + man*).
- d. Intensifiers with respect to adjectives (Intens&N; *very + tall*).
- e. Negative particles with respect to verbs (Neg&V; *not + go*).
- f. Tense/aspect particles with respect to verbs (Tense&V).<sup>9</sup>

Because adjectives, numerals, demonstratives, intensifiers, negative particles and verbal particles do not usually involve a connective, the restricted predictive power of the orientation principle may be advantageous: no correlation – no prediction. A comparison of correlation with noncorrelation pairs is carried out here to help us see the big picture in syntactic typology: where connectives are involved, there are correlations – where not, there are noncorrelations.

Regarding the true correlations in addition to adposition type, Cinque (2013) gives a list of another six pairs of bidirectional correlation, which will be discussed one by one below.

### 6.1. *OV&VAux vs. VO&AuxV ('must read books')*

This is apparently related to *OV&VWant vs. VO&WantV* ('wants to read books'). Some languages may use connectives as infinitive markers, e.g., French *Je veux essayer de lire* ('I want to try to read'). Infinitive affixes are also considered as a type of purpose particle (Schmidtke-Bode 2009, 36–37) in Hungarian (Finno-Ugric, SOV/SVO), for example:

- (9) *Anna elküldte Péter-t [a könyv-et olvas-ni].*  
 Anna sent Peter-ACC [the book-ACC read-INF]  
 'Anna sent Peter to read the book.' (Kenesei, Vago, and Fenyvesi 1998, 56)

However, it is also a possibility that the infinitive marker should not be treated as a subordinator, and that the cross-linguistically prevalent place of the 'main' verb is derived from the canonical place of the object since it is the semantic object of the auxiliary (*I want: food*; cf. *I want: to eat*).

The canonical place of the object, for its part, follows from the orientation principle when an object marker is assumed. This is not a trivial point just because accusative markers are more common than nominative markers (Comrie 2013), but also because the difference between objects and

<sup>8</sup>Dryer's list excludes b. Num&N as a noncorrelation pair, but the example '*four + books*' is from Dryer (1992).

<sup>9</sup>See examples from Yapese and Kiowa in Dryer (1992).

prepositional phrases is not always clear-cut, either. The semantic object can be marked with a preposition, e.g., ‘look **at** me’ (versus ‘look **me** in the eye’). Another case in point is Spanish where the preposition *a* is used as an accusative marker to such an extent that Spanish is classified as a nominative-accusative language by Comrie (2013), as opposed to its ‘neutral’ relatives French and English.

### 6.2. *OV&PredCop vs. VO&CopPred*

An example of this is ‘Lesley is a teacher’, where *is* is the copula and *a teacher* is the predicate. This is another exemplification of the place likely being derived from the place of the object, with no apparent connective.

### 6.3. *OV&AdvV vs. VO&VAdv*

The canonical place of the manner adverb appears to be derived from the place of the prepositional phrase, which is a harmonic correlation, e.g., English: ‘the dog ran joyfully’ – ‘the dog ran **with** joy’; or, ‘she hit his picture furiously’ – ‘she hit his picture **in** fury’.

### 6.4. *OV&StAdj vs. VO&AdjSt*

The order of standard of comparison and adjective, e.g., English ‘bigger <**than** you’ is in line with the orientation principle when a postpositional connective is assumed for OV languages. An example of this comes from Mundari (India; Austro-Asiatic):

- (10) *sadom-ete* *hati* *maranga-e*  
 horse-**from** elephant big-3SG.PRS  
 ‘The elephant is bigger than the horse.’ (Hoffmann 1903, 110)

In this example, which is also used by WALS (Stassen 2013), the semantic adjective appears to be a verb. In such cases, though, the StAdj order is actually necessitated by the SOV order itself.

### 6.5. *OV&PPAdj vs. VO&AdjPP*

This correlation is predicted by the orientation principle when the VO ordering is as in the English ‘green <**with** envy’; the OV equivalent is expected to resemble ‘envy-**WITH**> green’ as in the corresponding Finnish<sup>10</sup> expression.

<sup>10</sup>Personal knowledge. The reverse ordering *vihreä kateudesta* is also grammatical (Hakulinen et al. 2004, §618). Finnish has the exceptional canonical SVO order with postpositions.

- (11) *kateude-sta vihreä*  
 envy-ELAT green  
 ‘green with envy’

### 6.6. *OV&PP-V vs. VO&V-PP* (‘goes to school’)

Assuming postpositions in OV languages and prepositions in VO languages, the orientation principle correctly predicts the order of verb and prepositional phrase; see example (4) in Japanese.

There are some further correlations in Dryer (1992) that are not included in Cinque’s list. The order of genitive and noun is quite clearly predicted by the orientation principle as exemplified by the two modes of English possessive (3). The order of main clause and indirect question does not correlate very strongly, but it nonetheless follows from the orientation principle if there is a question particle that functions as a bi-clausal connective as in the English indirect question: “I wish to know **if/whether** my assistance is required.” Furthermore, the order of main clause and subclause is in general predicted by the orientation principle, assuming that subordinators are subclause-final in OV languages and clause-initial in VO languages (cf. Diessel 2001; Schmidtke-Bode 2009).

Additionally, Cinque (2013) discusses a unidirectional correlation between the order of object and verb vs. the order of noun and relative clause. VO languages implicate NRel, and RelN implicates OV; but there are almost as many OV languages with NRel (Dryer 2013a). Assuming a subclause-initial relativiser, the orientation principle passes the test of VO&NRel (‘saw something **that** . . . ’). It would appear to predict a lack of a relativising subordinator in OV languages since there is no correlation.

This seems correct on the basis of examples of OV relativisation given by Dryer (2013a): relativisation is typically adjectival or could for instance be caused by movement whereby the NRel ordering follows from the SO order in both SOV and SVO languages.

To illustrate, the basic SOV order ‘dogs food love’ (for ‘dogs love food’) can be changed to ‘food dogs love’ for relativisation. It is noteworthy that reversing the subject and the object gives rise to NRel in both SVO and SOV languages (with food representing N), suggesting it may not be a proper case of an OV/VO correlation in the first place. This could bring the simple orientation principle close to the accuracy of Hawkins’s (2004, 2014) complex model. VO&NRel satisfies both orientation of connectives and the lack of a connective: OV languages have RelN if there is a subclause-final relativiser and NRel if there is not.

However, much of SOV relativisation is participial, and it remains to be firmly established whether participial suffixes actually correspond to proper relativisers. Although it had been once suggested that participial relatives are





- (13) a Mary<sub>VP</sub>[gave<sub>PP</sub>[to Bill]<sub>NP</sub>[the book she had been searching for since last Christmas]]  
 b Mary<sub>VP</sub>[gave<sub>NP</sub>[the book she had been searching for since last Christmas]<sub>PP</sub>[to Bill]] (Hawkins 2004, 26)

When we look at predictions made by EIC/MiD as regards harmonic correlations in particular, it is tempting, from the orientation perspective, to suggest that the simplest explanation of their accuracy is in the realisation that adpositions and conjunctions are statistically placed *at the boundary* between two phrases or sentences. This is not to say that Hawkins's model is redundant since it makes more predictions than just that, e.g., the dominant order of subject, object and verb.

But Song (2012, 286) suggests that there remain issues concerning the dominant S\O\V order. One of them is that EIC/MiD treats the order of determiner phrase (DP) in relation to noun as a harmonic correlation. The data, however, suggest a cross-linguistic tendency of placing the determiner before the noun regardless of language type (Song 2012, 246–247). Again, the orientation principle makes no predictions.

This brings us to a further issue to consider. As discussed above, Diessel's (2001) material supports a general pattern of subordination as predicted by both orientation and EIC/MiD. It is merely at the level of a closer examination that irregular patterns are uncovered. There is a bias towards sentence-initial adverbial clauses because OV languages tend to begin the sentence with an adverbial clause, while VO languages have flexible or mixed ordering. Diessel demonstrates that especially conditional subclauses and – to a lesser degree, temporal subclauses – often appear sentence-initially in English and other VO languages, too. Diessel's (2001, 443) example sentence *If you change jobs, you won't necessarily have to sell the farm* conflicts with the orientation principle which assumes that the connective should join the two sentences. It indeed does join them in the overall patterning of the subordination data, but conditional clauses do not correlate. Examples of a conditional clause preceding the main clause include Babungo (Niger-Congo; SVO) and Malayalam (Dravidian; SOV).

- (14) [*kí à gàŋtə mə*], *mə kə fá tí ghô*.  
 [if you help I], I give thing to you  
 'If you help me, I have to give you something.'  
 (Babungo; Schaub 1985, 40, modified)

- (15) [*avan vann-aal*] *parayaam*.  
 [he come-if] tell.FUT.MOOD  
 'If he comes, I shall tell (him).'
- (Malayalam; Asher and Kumari 1997, 87, modified)

From a pragmatic perspective one may simply suggest that there is a different communicative factor at play: a principle of urgency, iconicity or isomorphy which may override the orientation of connectives particularly in conditionals. Accounting for such a break from the generalised [C<sub>m</sub>IC] pattern in VO languages will be difficult for Hawkins, who argues that pragmatics or information structure plays no role in the explanation of syntactic universals. There would have to be a separate processing explanation for the typology of conditional clauses; one that overrides that of purpose clauses (e.g., *You won't necessarily have to sell the farm to change jobs*). However, as we have seen, adding an ad hoc solution just for conditional clauses could prove to cause serious repercussions for the overall model.

To summarise, there are similarities and differences between predictions made by the two principles. For lack of a complete impartial comparison, it is up to the reader which one they choose, if any. The fundamental difference, though, may not be as much in the generalising principle as it is in thinking about language. Hawkins (2004, 174) substantiates his rejection of Universal Grammar by reference to research on event-related brain potentials (ERPs; Kluender and Kutas 1993). Hawkins's interpretation is, quite correctly, that the researchers failed to find support for an innate syntactic constraint in *wh*-islands (cf. Chomsky 1977, 1993) and argue for processing effects instead. Hawkins takes this as evidence for explanations in terms of processing while still holding on to a biological view of language.

However, what Kluender and Kutas (1993, 573) uncovered were more precisely "lexical semantic processing effects". Attention was paid especially to the connectives *who*, *what*, *if* and *that* in the embedded clause boundary. These were linked with working memory: the results show that the semantic role or reference of the subordinator (especially the relativiser) was resolved as the relevant information became available in the sentence. The longer the distance, the more load on working memory. This is all compatible with the view of language as a system of meaning. Because semantic structures are logical in the sense that they can be learned and understood by means of inference, I fail to see a necessity for a specific biological component governing interaction of meaning and form.

What is more, maybe the biggest issue for Hawkins's model is in the very concept of processing difficulty (1994) or grammatical efficiency (2004). According to this standpoint, Chinese, due to its canonical VO&AdvV and VO&RelN ordering (cf. Dryer 1992; Dryer and Gensler 2013), exhibits frequent 'inefficiency', as do Finnic languages with SVO and postpositions. If the linguistic form is based on processing universals, one would not expect such patterns to be as stable as they are.

In contrast, a humanistic and systemic approach takes language itself as the starting point. It would be interesting for linguists to take a closer look into such language systems - ones that have been perceived to

exhibit inefficiency – in the future, to see what other structural differences they contain, as defined in terms of the interaction of syntax and semantics, something that might reinforce the exceptional patterns. This would be a matter of seeking deeper insight into language, now that major patterns in syntactic typology have been uncovered as more data and more analyses have been made available by the many pioneers in the field.

## 8. Conclusion

This article introduced the orientation principle, which is a semantically and pragmatically oriented generalisation of harmonic correlations. It was shown on a preliminary basis that this very simple principle makes roughly as good predictions as complicated processing explanations.

As Song (2012) argues, linguists are keen to find out which single-principle generalisation makes the most correct predictions in order for it to be used as the foundation of typological explanation. Song (2010, 2018) argues that the leading model has been provided by Hawkins (2004), but Song (2012) admits that the scientific premises of the processing explanation are yet to be confirmed by psycholinguists.

The orientation principle comes in as an alternative proposition. Dryer's (1992 and later) data, which is based on Greenberg's (1963) approach, has long served as a standard reference, emphasising the role of phrasal order in defining types of variation.

The findings of this paper, however, suggest that a reanalysis of the material should be undertaken to examine the role of the connective in harmonic correlations. This task could offer a ground for determining whether the participle phrase is a proper case of relativisation, or to be grouped with adjectival phrases.

## Abbreviations

ACT	active
ELAT	elative
MOD	modal

## Disclosure statement

No potential conflict of interest was reported by the author.

## Notes on contributor

*Patrik Austin* holds an MA in Scandinavian languages with pedagogy, philosophy, and English philology from the University of Helsinki where he is now doing a Ph.D. in general linguistics. The theme of his forthcoming dissertation is explanation of language in relation to typological universals.

## ORCID

Patrik Austin  <http://orcid.org/0000-0002-5015-4353>

## References

- Abels, Klaus. 2015. "Word Order." In *Syntax – Theory and Analysis*, edited by Tibor Kiss and Artemis Alexiadou, vol. 1, 1400–1445. Berlin: Walter de Gruyter.
- Asher, Ronald E., and T. C. Kumari. 1997. *Malayalam*. London: Routledge.
- Behaghel, Otto. 1932. *Deutsche Syntax: Eine geschichtliche Darstellung, Band 4: Wortstellung, Periodenbau*. Heidelberg: Carl Winter.
- Butler, Christopher S. 2003. *Structure and Function – A Guide to Three Major Structural-Functional Theories, Part 2: From Clause to Discourse and Beyond*. Amsterdam: John Benjamins.
- Chappell, Hilary, and Alain Peyraube. 2007. "Chinese Linguistics and Typology: The State of the Art." *Linguistic Typology* 11 (1): 187–211. doi:10.1515/LINGTY.2007.014.
- Chomsky, Noam. 1977. "On Wh-movement." In *Formal Syntax*, edited by Peter W. Culicover, Thomas Wasow, and Adrian Akmajian, 71–132. Orlando, FL: Academic Press.
- Chomsky, Noam. 1988. *Language and Problems of Knowledge: The Managua Lectures*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 1993. *Prospects for Minimalism*. Lecture given at the University of California, Irvine, March 8, 1993.
- Cinque, Guglielmo. 2013. "Word-order Typology: A Change of Perspective." In *Theoretical Approaches to Disharmonic Word Order*, edited by Theresa Biberauer and Michelle Sheehan, 47–73. Oxford: Oxford University Press.
- Comrie, Bernard. 2013. "Alignment of Case Marking of Full Noun Phrases." In *The World Atlas of Language Structures Online*, edited by Matthew S. Dryer and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology. Accessed 10 August 2020. <http://wals.info/chapter/98>
- Daneš, František. 1987. "On Prague School Functionalism in Linguistics." In *Functionalism in Linguistics*, edited by René Dirven and Vilém Fried, 3–38. Amsterdam: John Benjamins.
- Davidse, Kristin. 1987. "M. A. K. Halliday's Functional Grammar and the Prague School." In *Functionalism in Linguistics*, edited by René Dirven and Vilém Fried, 39–79. Amsterdam: John Benjamins.
- de Saussure, Ferdinand. 1916. *Cours de linguistique générale*. Paris: Payot.
- de Vries, Mark. 2004. "The Syntax of Relativization." PhD diss., University of Amsterdam.
- Diesel, Holger. 2001. "The Ordering Distribution of Main and Adverbial Clauses: A Typological Study." *Language* 77 (2): 433–455. doi:10.1353/lan.2001.0152.

- Dik, Simon C. 1978. *Functional Grammar*. London: Academic Press.
- Dik, Simon C. 1989. *The Theory of Functional Grammar, Part I: The Structure of the Clause*. Dordrecht: Foris.
- Dryer, Matthew S. 1992. "The Greenbergian Word Order Correlations." *Language* 68 (1): 81–138. doi:10.1353/lan.1992.0028.
- Dryer, Matthew S. 2013a. "Relationship between the Order of Object and Verb and the Order of Relative Clause and Noun." In *The World Atlas of Language Structures Online*, edited by Matthew S. Dryer and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology. Accessed 10 August 2020. <http://wals.info/chapter/96>
- Dryer, Matthew S. 2013b. "Order of Adjective and Noun." In *The World Atlas of Language Structures Online*, edited by Matthew S. Dryer and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology. Accessed 10 August 2020. <http://wals.info/chapter/87>
- Dryer, Matthew S., and Orin D. Gensler. 2013. "Order of Object, Oblique, and Verb." In *The World Atlas of Language Structures Online*, edited by Matthew S. Dryer and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology. Accessed 10 August 2020. <http://wals.info/chapter/84>
- Franklin, Karl J. 1971. *A Grammar of Kewa, New Guinea*. Canberra: Australian National University.
- Frey, Werner. 2015. "Word Order." In *Syntax – Theory and Analysis: An International Handbook*, edited by Tibor Kiss and Artemis Alexiadou, 514–562. Berlin: De Gruyter Mouton.
- Greenberg, Joseph H. 1963. "Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements." In *Universals of Language*, edited by Joseph H. Greenberg, 58–90. Cambridge, Mass.: MIT Press.
- Guerrero, Paul, Stefan Jeschke, Michael Wimmer, and Peter Wonka. 2014. "Edit Propagation Using Geometric Relationship Functions." *ACM Transactions on Graphics (TOG)* 33 (2): 1–15. doi:10.1145/2591010.
- Haberland, Hartmut, and Lars Heltoft. 1992. "Universals, Explanations and Pragmatics." In *Meaning and Grammar: Cross-Linguistic Perspectives*, edited by Johan van der Auwera and Michel Kefer, 17–27. Berlin: De Gruyter Mouton.
- Haiman, John. 1983. "Iconic and Economic Motivation." *Language* 59 (4): 781–819. doi:10.2307/413373.
- Hakulinen, Auli, Maria Vilkuna, Riitta Korhonen, Vesa Koivisto, Tarja Riitta Heinonen, and Irja Alho. 2004. *Iso suomen kielioppi* [Comprehensive Grammar of Finnish]. Helsinki: Suomalaisen Kirjallisuuden Seura. Accessed 23 September 2021. <http://scripta.kotus.fi/visk/sisallys.php?p=618>
- Hawkins, John A. 1983. *Word Order Universals*. New York: Academic Press.
- Hawkins, John A. 1994. *A Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press.
- Hawkins, John A. 2004. *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press.
- Hawkins, John A. 2014. *Cross-Linguistic Variation and Efficiency*. Oxford: Oxford University Press.
- Hengeveld, Kees, and J. Lachlan Mackenzie. 2008. *Functional Discourse Grammar: A Typologically-based Theory of Language Structure*. Oxford: Oxford University Press.

- Hjelmslev, Louis. [1943] 1969. *Prolegomena to a Theory of Language*. Translated by Francis J. Whitfield. Madison: University of Wisconsin Press.
- Hoffmann, Johann. 1903. *Mundari Grammar*. Calcutta: Bengal Secretariat Press.
- Huddleston, Rodney. 1984. *Introduction to the Grammar of English*. Cambridge: Cambridge University Press.
- Itkonen, Esa. 2013. "Functional Explanation and Its Uses." In *Functional Approaches to Language*, edited by Shannon Bischoff and Carmen Jany, 31–69. Berlin: Walter de Gruyter.
- Kenesei, István, Robert M. Vago, and Anna Fenyvesi. 1998. *Hungarian*. London: Routledge.
- Kluender, Robert, and Marta Kutas. 1993. "Subjacency as a Processing Phenomenon." *Language and Cognitive Processes* 8 (4): 573–633. doi:10.1080/01690969308407588.
- Koerner, Ernst Frideryk Konrad. 1978. *Toward a Historiography of Linguistics: Selected Essays*. Amsterdam: John Benjamins.
- Koller, Alexander. 2015. "Semantic Construction with Graph Grammars." In *Proceedings of the 11th International Conference on Computational Semantics*, 228–238. London: Association for Computational Linguistics.
- Koopman, Hilda. 1984. *The Syntax of Verbs: From Verb Movement to Rules in the Kru Languages to Universal Grammar*. Dordrecht: Foris.
- Lehmann, Christian. 1984. *Der Relativsatz*. Tübingen: Gunter Narr.
- Lytinen, Steven L. 1987. "Integrating Syntax and Semantics." In *Machine Translation – Theoretical and Methodological Issues*, edited by Sergei Nirenburg, 302–316. Cambridge: Cambridge University Press.
- Martinet, André. 1955. *Économie des changements phonétiques: Traité de phonologie diachronique*. Berne: Francke.
- McCarthy, John J., and Alan Prince. 1995. "Faithfulness and Reduplicative Identity." *University of Massachusetts Occasional Papers in Linguistics* 18: 249–384.
- Newmeyer, Frederick. 1998. "The Irrelevance of Typology for Grammatical Theory." *Syntax* 1: 161–197.
- Newmeyer, Frederick. 2005. *Possible and Probable Languages: A Generative Perspective of Typology*. Oxford: Oxford University Press.
- Schaub, Will. 1985. *Babungo*. London: Croom Helm.
- Schmidtke-Bode, Karsten. 2009. *A Typology of Purpose Clauses*. Amsterdam: John Benjamins.
- Seuren, Pieter. 1998. *Western Linguistics: An Historical Introduction*. Oxford: Blackwell.
- Song, Jae Jung. 2010. "Word Order Typology." In *The Oxford Handbook of Linguistic Typology*, edited by Jae Jung Song, 253–280. Oxford: Oxford University Press.
- Song, Jae Jung. 2012. *Word Order*. Cambridge: Cambridge University Press.
- Song, Jae Jung. 2018. *Linguistic Typology*. Oxford: Oxford University Press.
- Spivey, Michael, Marc Joanisse, and Ken McRae, eds. 2012. *The Cambridge Handbook of Psycholinguistics*. Cambridge: Cambridge University Press.
- Stassen, Leon. 2013. "Comparative Constructions." In *The World Atlas of Language Structures Online*, edited by Matthew S. Dryer and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology. Accessed 10 August 2020. <http://wals.info/chapter/121>
- Stowell, Timothy. 1981. *Origins of Phrase Structure*. Cambridge, Mass.: MIT dissertation.
- Tesnière, Lucien. 1959. *Éléments de syntaxe structurale*. Paris: Klincksieck.

- Travis, Lisa. 1984. "Parameters and Effects of Word Order Variation." PhD diss., MIT.
- Travis, Lisa. 1989. "Parameters of Phrase Structure." In *Alternative Conceptions of Phrase Structure*, edited by Mark R. Baltin and Anthony S. Kroch, 263-279. Chicago: University of Chicago Press.
- Vennemann, Theo. 1974. "Topics, Subjects, and Word Order: From SXV to SVX via TVX." In *Historical Linguistics, vol. 1: Syntax, Morphology, Internal and Comparative Reconstruction*, edited by John M. Anderson and Charles Jones, 339-376. Amsterdam: North Holland.
- Vincentini, Alessandra. 2003. "The Economy Principle in Language: Notes and Observations from Early Modern English Grammars." *Words Mots Palabras* 3: 37-57.
- Yoshimoto, Banana. 1991. *Kittchin* [Kitchen]. Tokyo: Kadokawa shoten.