

The true cause of word-order harmonies – a reply to John A. Hawkins

Keywords: typology, universals, harmony, syntax, humanism, sociobiology

Abstract

We answer Hawkins's question why functional grammar is 'functional', reciprocating the question to his processing explanation of word-order variation. Our own research uncovers a link between economy, information structure and the place of connectives. Based on this clue and background knowledge from neurolinguistic research, we uncovered a crosslinguistically prevalent pattern which offers a key to a simpler explanation of harmony. This gives us the opportunity to reject Hawkins's MiD as redundant, and to underline that all biological approaches to syntax remain without a scientific basis.

1. Introduction

Hawkins (1994) successfully challenged contemporary paradigms of word-order theory and has since then (2004, 2014) further developed his theory of language processing as an explanation for crosslinguistic grammatical and performance preferences. Hawkins's theory consists of a series of postulated processing principles. The overall theory is called 'The Performance–Grammar Correspondence Hypothesis' (PGCH; Hawkins 2014). This name is an apparent reaction to the generative idea that grammar fully depends on innate factors, a belief for which we fail to find any support in the research literature. We take it as granted under such circumstances that the fact that people use language as a communication tool is the default source of explanation for why languages have the properties they have..

Therefore, while Hawkins spends much of his effort arguing that grammar cannot be completely innate, we find no cause for objection to this and consider it agreed. Instead, we will focus on Hawkins's principles of Early Immediate Constituents

(EIC) which he subsumes under the new term MiD (Minimize Domains) in Hawkins (2004; see further discussion in section 4).

Hawkins's work is a special case in today's linguistics in that it has had a significant impact in generative grammar as well as in typology. Generativists have described Hawkins's processing hypothesis as "influential" (Abels 2015) or "very influential" (Frey 2015) even if it is yet to persuade the various frameworks to explicitly endorse CAS (see below) rather than innate view of language.

Reception in typology has similarly been mixed positive. Although there is a long and well-known dispute between generativists and mainstream typologists (at least since Greenberg, 1979), Croft (2001/2015) places processing explanation into the orthodoxy of what he calls "functionalism". Croft does not consider the case of harmony as solved as such but he (Croft 2003, 80) argues that it has been demonstrated to a reasonable degree that both 'diachrony' and 'processing' do play a role in explanation of word order and affix order crosslinguistically. We take these two terms, respectively, as meaning that language is (i) a CAS (Complex Adaptive System; Beckner et al. [incl. Croft] 2009) which adapts to a certain (ii) innate grammatical 'structure' (Hawkins & Gell-Mann 1992).

Hawkins believes that the cause of dominance and harmony distributions can be uncovered by comparing human language processing to how a computer parses phrase structure grammars. He has been critical of topicalisation as an explanans for word-order variation, claiming that so-called functional grammar of the Prague Circle and its derivatives may not be properly 'functional' (Hawkins 1994, 116—117).

Our reply to Hawkins begins by discussing the standard meaning of functional explanation (section 2). Section 3 will take a retrospective view to the dispute between Hawkins and the advocates of what used to be called functionalism. Hawkins's model

will be critically examined in section 4. To answer Hawkins's question we will test and analyse texts in four languages (section 5). We will then proceed to study the related typological data (section 6). Based on all of our discussions and findings, we will reject MiD as an explanation of word-order variation by proposing a simpler model with equally good predictions (section 7).

2. What is a functional explanation?

Functional mode of explanation explains structures by an appeal to their functions (Couch 2011). For example, the reason why a heart has the properties it has depends on the function of the heart of pumping blood. At the same time, it is often implied that the blood-pumping function is the reason for the existence of hearts in living organisms. Explanations of this type in various sciences have been discussed extensively in philosophy of science especially since Nagel (1961). The tradition of functional explanation is however much older, dating back at least to Aristotle's concept of teleology (Allen & Neal 2019).

Traditionally, the functional mode of explanation in linguistics is the same as in biology and social sciences (Daneš 1987; see also Couch, *ibid.*). Functional linguists explain linguistic forms or structures in relation to the task, function or purpose that they serve or as a means to an end. Following Jakobson (1969), Daneš makes the case that functional explanation is equivalent to teleological finalism. Language and its structures are based on their instrumental or functional value. While language is known to have different functions, it is the communicative function that is taken as a starting point for functional-structural linguistics. Functional linguists also interpret Saussure's bilateral conception of the linguistic sign functionally: linguistic form depends on its semantic task. (Daneš 1987.)

The functional concept has made its way into modern everyday language. Functional clothing, for example, may refer to weatherproof properties: the function of the water-repellent structure in a jacket is to protect its wearer from getting wet in the rain. We understand that people design such clothing on purpose. Daneš (ibid.) however points out that the concept of functionality itself is separate from the question of intentionality. Thus, functionalism as an explanation of the construction of language “evidently involves a high degree of unconscious components” (Daneš 1987, 8). Itkonen (2011) considers functional explanation as *rational explanation*, meaning that languages are intelligently created by people to serve their own goals; but construction processes are largely unintentional and unconscious (see also Haberland & Heltoft 1992).

Or, so one might have thought. What is said above appears to be regarded as dated information because Nichols (1984) argues that the concept of functionalism “has changed”; and instrumentality is now associated with “structuralism”. Based on Nichols, Croft (1995) further associates structuralism with Chomsky. This chain of argumentation apparently suggests that Aristotle, Darwin and the early Prague Circle were Chomskyan thinkers; and that Chomsky believes linguistic structures are best explained through their instrumental value, all which cannot be correct.

But functional explanation according to Croft (1995, 493) *has changed* to pertaining to phenomena “external” to what *Chomsky* calls I-language (I for ‘internal’); that is, as opposed to “formal explanation”. It is in other words claimed that “formalism” refers to a view of language as a non-adaptational innate phenomenon. Functionalism, along this line, refers to an idea of language as an adaptational but not

innate phenomenon; reminiscent of classical **Darwinian linguistics** (Darwin 1871, 60—62; Aronoff 2017).

For a brief comment on the above, the question of innatism is difficult because language could be innate but functional, or non-innate and non-functional. This does not mean that Darwinian linguistics cannot be functional. It is rather that we want to understand why it is believed to be functional. Functional explanation was frequently employed by Darwin (1859). But there is something that does not seem to add up. As pointed out above, the structuralist notion of functional explanation corresponds to its concept in biology. Croft (1993, 22) however argues that the structuralist conception is misguided – “a caricature of biological theory” – and proposes “an adaptational view of functionalism” which he states is furthest developed in Cognitive Linguistics (Croft 1993, 22). Croft (2001/2015) emphasises that the meaning of *functional* in linguistics is **distinct** from its use in philosophy.

Such a position has to be very challenging because discussions concerning functional explanation in philosophy of science tend to be heavily focused on the use of the term in biology. What Croft must be suggesting is that evolutionary linguists have a more correct understanding of biology than evolutionary biologists. By all means, Nichols and Croft are not alone with their views. Hopper (1987, 3) discusses the “Functionalist Fallacy” and, Haspelmath (1999, 185), the “Teleological Fallacy”. Keller (1997, 13) proposes developing *a new concept* of functional explanation which is “free of teleological pollution.” Pointing to a functional explanation of the giraffe’s neck in biology, Keller argues that the correct formulation may not be, for instance, that the function of the length of the neck is that of reaching higher leaves for food; that is, an appeal to instrumental or functional *value*. Instead, a “workable” functional explanation could be as follows:

The length of a giraffe's neck is a function of the fact that it lives from the leaves of trees. (Keller 1997, 13.)

Thus, 'function' in Keller's definition of functional explanation is synonymous with 'consequence', rejecting an instrumental reading. This is crucially different from Darwin's functional explanation where the function entails instrumental value and purpose, exactly as it is taken in functional linguists; and as rejected by the Darwinian linguists. Darwin writes:

If about a dozen genera of birds had become extinct or were unknown, who would have ventured to have surmised that birds might have existed which used their wings solely as flappers, like the loggerheaded duck (*Micropterus* of Eyton); as fins in the water and front legs on the land, like the penguin; as sails, like the ostrich; and functionally for no purpose, like the *Apteryx*. (Darwin 1859, 165—166.)

In other words, Keller suggests that the concept of function, as we have known it, should be removed from linguistics and natural sciences alike. The ideas of Hopper and Keller have given rise to "emergent" and "invisible-hand" linguistics, respectively. Dahl (2004) identifies both as representing the same idea of cultural evolution, linking them with early Darwinian linguistics and Dawkins's (1976) *memetics*, a theory of cultural units as 'mind-viruses'.

The argument made by evolutionary linguists is two-edged. It is claimed that classical functional linguistics is misguided and should be replaced with the Darwinian concept of natural selection. At the same time, the Darwinian concept of organic function should be removed and replaced with a non-functional evolutionary explanation; and this explanation is to be relabelled 'functional'. Why?

Let us illustrate the difference. The concept of systemic function appeals to the maintenance of a structure. A heart cannot be removed from an organism that requires it to live. The concept of natural selection, on the other hand, is meant to explain the emergence and maintenance of such functional structures. The evolutionary concept suggests that a heart has evolved through a series of structural changes which became prevalent in a population through natural selection.

Darwin's theory relies on proper functional explanation. The ultimate reason for hearts existing is not only the fact that there were *changes*, but because these changes turned out to be *useful*. This way they gave an evolutionary advantage to individuals with such properties, over others, in the given circumstances. However, what today's Darwinian linguists claim is that the functional mode should be abandoned in order for the explanation to be "functional". Certainly, things could not be looking much worse.

Let us dive deeper into the claims to understand where the problem lies. Firstly, it needs to be pointed out that the term "Darwinian linguistics" is fully justified. People commonly associate Darwin with biology, but his theory of natural selection actually encompasses social evolution: human and non-human. Following the success of *On the Origin of Species*, the principles of natural selection were soon applied to theory of language (Schleicher [1863] 1869; Lyell 1863; Müller 1870; Darwin 1871). Although challenged by structuralists such as Durkheim and Saussure, social Darwinism enjoyed great influence until the end of World War II after which it was banished from human sciences (Aronoff 2017).

Darwinian ideas, or what is sometimes called 'universal Darwinism' made a stealthy return in a 1970s sociobiological wave in the form of various cultural replicator theories, now most famously Dawkins's memetics which was eventually featured in *The Skeptic Dictionary of Pseudoscience* (Polichak 2002). Based on the scientific criticism

that memetics had received, especially that the memetic concept has not been properly elucidated, linguists opted for a new model, *Complex Adaptive System* (CAS; Frank 2008). CAS purportedly enables the study of the collective behaviour of self-replicating units, such as memes (Gell-Mann 1994). This is the same model that was proposed early on by Hawkins who considers it analogous to the selfish gene schema (Hawkins & Gell-Mann 1992, 12; cf. Dawkins 1976).

The problem of functional explanation continues to trouble evolutionary linguistics. According to Haspelmath (1999), evolution cannot be propelled by the needs of the individual. He takes the example of polar fish with antifreeze proteins, pointing out that just any fish cannot start producing such proteins simply because they are useful in freezing water. Transferring natural selection to culture, Haspelmath is adamant that people are incapable of linguistic innovation. This leads to some strange conclusions. In Haspelmath's example, people can invent a bicycle saddle, but not a word for it. This is due to "obvious reasons" which he leaves without an explanation. We believe it is due to a flaw in Darwin's theory. We understand that the evolutionary model absolutely excludes the concepts of intelligence, design and creativity which, nevertheless, are the driving principles of cultural change (cf. Itkonen 1999, 2011).

In fact, Haspelmath's interpretation of Dawkins (1976) is incorrect: the universal Darwinian concept also covers scientific and technological innovation (p. 190 et passim). Consequently, the claim should be that people cannot just invent a bicycle saddle so they will have to bike around without one until a suitable meme mutation becomes available.

As regards the compatibility of the social Darwinian and the functional mode of explanation, it is interesting to note Daneš's (1987, 8) point that people have no functions, in the sense that is meant with functional explanation "unless they are being

used (or misused or manipulated) by other humans to attain their goals.” The same would seem to be true of the hypothesised replicator population. In the CAS view, language is not an instrument for communication. More precisely, communication is not the purpose of language, but a *consequence of having language*. Language – a population of replicators – is agentive, not instrumental. As much as communication is vital for the reproductive needs of cultural replicators, it is actually the case that language uses people as an instrument, not vice versa. Applying Dawkins’s concept of people as vessels for the needs of the selfish meme, what we people think of as speech and sensory organs are actually the replicators’ genitals.

In so far as we have correctly understood the view of language as advocated by Hawkins & Gell-Mann (1992), and Hawkins’s (2014, 3) theoretical sources which include Haspelmath (1999) and Kirby (1999), his efficiency explanation cannot, then, be *functional* because language in this view does not have a function. It is people and communication that are functional for the needs of a given language, the purported quasi-parasitic population, or “organism” as Kirby (2011) calls it. But this is not functional linguistics, it is functional humanity.

We have come to what in the first place prompted our investigation into the schism between Hawkins and the advocates of Functional Sentence Perspective (FSP). Hawkins (1994) asks:

The Prague School theory of given-before-new ordering seems to be particularly *non-functional*: why should each sentence begin with what has already been presented, delaying the newest information till the very end? (Hawkins 1994, 116—117; italics in the original.)

We will answer Hawkins’s question in detail in section 5. Here, we would like to take the opportunity to reciprocate the question. Why is Hawkins’s word-order theory

functional? Especially as a reference to standard literature on the subject (e.g. Couch ibid.). We propose that this type of explanation will be called ‘adaptive’ or ‘adaptational’; but not functional.

3. Background of the dispute

A humanist view of man and language is different from sociobiology. Linguistic structures such as topicalisation effects on the word order depend on the speakers’ communicative strategies (Hajičová 1994, 260). In this perspective, people have an agentive role in social interaction. Functional explanation implies that language is constructed by the speech community according to its needs. This is basically because it is hard to find any other scientifically plausible cause.

Hawkins’s question, as quoted shortly above, appeared as a comment on pre-1990s topicalisation studies. Based on earlier insights, Mathesius, Firbas, Daneš and others developed a three-layer model of linguistic analysis that studies the interaction of syntax, semantics and pragmatics. This analytical model is called Functional Sentence Perspective (FSP) because it is a tool for observing how formal structures convey informational functions (Firbas 1987).

Using FSP led to the reiteration that old information, in a sentence, typically precedes new. Explanation by Firbas (1971) as for why this should be so is probably ‘functional’ mainly as a rather general reference to language as a tool for communication. The reason for old information to appear first falls out of the premises that, as the discourse evolves, what is said first becomes established, laying ground for what follows. The new information of the previous sentence is the old information of the following sentence. It is possible for sentence structure to go against the chronological grain of the story line, but since this is not the point made by Prague

functionalists, the observation may not require a specifically functional ('means to an end') explanation. Firbas's principle of Communicative Dynamism (CD), which argues that informative value gradually rises towards the end of a sentence, does not appear to be an explanation of word-order universals, but rather, an observation arising from descriptive linguistics.

An early instance of the dispute between structural-functional and Darwinian linguists appears in Li & Thompson (1974a) who argue against economy explanation of word order change. In hindsight, the need to reject structural linguistics could be linked to an attempt to revive a pre-Saussurean tradition of evolutionary linguistics (cf. Aronoff 2017), especially Schleicher's extreme proto-language reconstructions (cf. Saussure 1916; Norde 2009, 48—105). The overall claim of Li & Thompson (*ibid.*), that Chinese has become an SOV language, is deemed as mistaken (Givón 1988). But, prompted by Thompson, Givón (1979) publishes a criticism of Saussure's separation of synchrony and diachrony. A vindictive attack on Saussure follows in Givón (1995), linking structuralism with generative grammar and thereby postulating a *second way* which he calls *functionalism*. It is only in (2002) that Givón redubs his *functionalism* as 'bio-linguistics', promoting the view of "language as a biological adaptation".

Givón (1983, 1988) expands his criticism to the Prague Circle, especially Firbas (1966, 1974), arguing that the classical functional approach is outdated after having been rejected by more modern research by Givón and his colleagues. It is a group collectively known as the 'West-Coast functionalists' (Bybee 1999).

The motivation for claiming the name functionalism is subject to speculation. Some clues are given by Aronoff (2017) who laments Darwinism being banned from humanities after the Second World War, blaming its downfall especially on Saussure; and Keller (1994) who likewise points to the bad reputation received by social

Darwinism, prompting its modern-day followers to camouflage in humanist terminology.

Croft's (2000, 2006) evolutionary model, on the other hand, is based on Hull's universal Darwinism. Hull (1988, 448) suggests that 'a research tradition' can expand its living space by parasiting its competitor. A common reference (see e.g. Keller 1994; Haspelmath 1999; Croft 2000; Dahl 2004; Haider 2015; Fillmore & Baker 2015) in evolutionary linguistics is Dawkins (1976) who discusses mimicry as an effective parasiting strategy.

Li & Thompson's failed attempt to revive evolutionary linguistics is followed by an attempt by Givón to replace functional explanation with a type of 'cognitive' explanation (section 3.1). This is eventually scrapped and replaced by Hawkins with a mixed cognitive-generative approach (section 3.2).

3.1 CTU: a second failed attempt

Givón (1988) argues that the Praguian tradition of functional explanation must be discarded and replaced "with a much more complex, cognitively-based understanding" of function. Givón questions the traditional claim that new information in sentences follows old information, finding that it is actually the other way around. This is due to a cognitive bias which he names CTU for 'Communicative Task Urgency'. In this view, there is a human tendency to arrange sentences to temporarily prioritise most important information.

A central source for Givón (1988) is Payne's (1987) study of word order in Papago (*O'odham*, Uto-Aztecan). Payne considers O'odham as a free-word-order language, and finds a strong tendency of placing indefinite nouns and new information pre-verbally, and definite nouns and old information post-verbally, concluding that

discourse factors constitute the main explanation of word order in this language and others.

Payne's overall conclusion is however different from Givón's proposition. She finds that some North American languages rather systematically topicalise by fronting new information although Slavic languages tend to do it reversely. This is quite interesting from a structuralist point. What is it in the language *system* that makes languages work in opposite directions? Or is it just another case of making an inevitable arbitrary choice between left and right in the linear ordering?

For a biological approach to language, however, crosslinguistic study of discourse pragmatics only ever leads back to the same persistent problem. If there is a human tendency of placing new information first, then, why do not all languages work this way? Givón (1988) goes as far as suggesting that the whole concept of topic and focus should be abandoned in favour of what could be described as cognitive grammar. Gundel (1988, 229, 234—235) however rejects CTU altogether, arguing that the 'importance' element of Givón's cognitive concept is too vague to be studied, and that both the given-before-new and the 'first things first' principle are valid concepts from a cognitive and pragmatic point.

3.2 EIC: third attempt successful

Hawkins's EIC (Early Immediate Constituents) is best seen as a last-ditch effort to make scientific sense of sociobiology. Hawkins (1994) makes the point that both the nativist Universal Grammar and Givón's cognitive explanation have failed to account for word-order variation, and he offers a working compromise between the two. The challenge is great. For one thing, some meaningful innate element would be desirable to make the biological claim seem necessary. For another, having failed biological

explanation both ways, information structure would have to be discarded completely as a potential explanans of word order universals.

The detail and diligence of Hawkins's work is therefore admirable. Building on a generative tradition of processing explanation, he presents a way to explain observations made by typologists with a model which gives more accurate predictions than any other. What is more, although Payne (1988) makes the reasonable suggestion that the syntax of free-word-order languages should be understood in discourse-pragmatic terms, Hawkins justifies his claim of abandoning information structure by demonstrating that EIC works statistically even with free-word-order languages.

EIC, or the current concept MiD (Minimize Domains) is now a staple of linguistic typology (e.g. Dryer 2013cd). It is of course false, as is willingly demonstrated within generative syntax (Frey 2015; see also section 6 in this paper), but it has sufficed to buy more time than needed by evolutionary linguists to usurp territory from their arch enemy, structuralism-functionalism (see Croft 1993; Aronoff 2017; versus. Saussure 1916; Sériot 1999) by replacing it with the Darwinian concept (Croft 2001/2015). Perhaps due to confusion surrounding the label functionalism, this prompts only a few protests which fail to make waves, including Itkonen (1999, 2011), Andersen (2006) and François (2018).

Hawkins (1994) argues that Givón's CTU (new before old), rather than Prague's CD (old before new), should count as an example of functional explanation. He then continues to reject both CD and CTU on the grounds that both claim to be functional while making the opposite assumptions. According to Hawkins, neither can provide a valid explanation because both types of information patterning, old-before-new and new-before-old, are *attested*.

At the same time Hawkins rejects much of the innatism claim providing a counter-example to Chomsky (1965) who argued that grammaticality questions always depend on Universal Grammar. Hawkins (1994) demonstrates that a centre-embedding structure, which is ungrammatical in English, is grammatical in Japanese. Therefore, English grammar cannot be innate. Hawkins goes on to argue for his own acceptability or performance-based explanation for which he finds support from statistical significance in multilingual data.

Here, an obvious fallacy made by Hawkins lies in that EIC likewise discusses various features that are *attested* in the data, whether predicted by his model or not. It cannot therefore claim superiority over functional sentence perspective on such grounds. Because Chomsky's innate parameters must hold universally, they can be rejected with counter-examples, but Firbas's (1985) analysis does not suggest that CD works without exception. The claim made by Hawkins is that his solution is more *functional*; but this must be mere wordplay (cf. section 2: *What is a functional explanation?*).

In fact, it was our first task, performed on texts in four languages, to find out the ratio of given-before-new versus new-before-given information structure case by case. The Prague notion proved prevalent in all four. The share of given-before-new sentences is about 90% percent in Czech, English, Finnish and Swedish (section 5.2). FSP should therefore hold its place as a useful tool for descriptive and theoretical linguistics.

As pointed out above, FSP is an analytical device rather than a principle of functional explanation. The latter is provided by Martinet (1955) whose definition of linguistic economy, following a long tradition on the subject, suggests an unstable balance between clarity and simplicity as a universal force that shapes languages

(Vincentini 2003). For this reason, we combined FSP and economy in our own research to gain insight into word-order variation (section 5).

4 Minimize Domains (MiD)

To better understand Hawkins's proposition, its methodological and theoretical implications need to be separated. Hawkins advocates a view that brings together generative sentence analysis on the methodological level, and linguistic units as cultural replicators as a theory of language (section 2). This is a major difference from Chomsky and his colleagues (Everaert et al. 2015) who consider generative grammar as the study of the hypothetical innate language faculty or language organ. But unlike others who reject innate 'structure' altogether, Hawkins does leave room for it. MiD is a marriage proposal between the generative analysis and the social Darwinian theory.

MiD is one of three principles under Hawkins's (2004, 2014) Performance—Grammar Correspondence Hypothesis (PGCH) and is itself divided into two concepts: *Early Immediate Constituents* (EIC) and *Phrasal Combination Domain* (PCD; previously known as *CRD – Constituent Recognition Domain*; Hawkins 2004, 106). The first one is the explanation and the latter is the analysis. Henceforth we will use the terms EIC and MiD largely interchangeably. The whole PGCH on the other hand is not further elaborated here because it is seen as the expansion of Hawkins (1994) with increasing explanatory power to cover various linguistic data; while this article focuses on the problem of harmony.

We take it that a “reduced” innate structure that Hawkins refers to concerns the generative convention of placing the object of the sentence into the verb phrase (VP) since Hawkins (1994 and elsewhere) continues to use it as a crosslinguistic point of reference.

4.1 Why is the object ‘in’ the verb phrase?

What seems most emblematic of generative analysis is the S ((NP) (VP (NP))) tree (Figure 1). A canonical English transitive sentence consists of two noun phrases (NPs) and a verb phrase (VP). For example, a tree for ‘*The man hit the ball*’ places the object (the second NP) within the verb phrase. The subject (the first NP), in contrast, is analysed as independent of the VP.

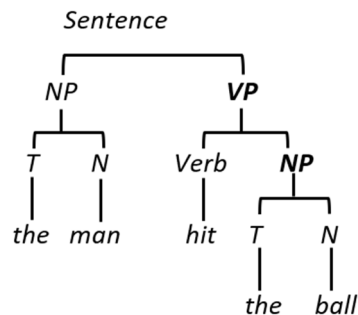


Figure 1. Chomsky (1957, 27; highlight added) places the verb and the direct object (the second NP) into a common verb phrase (VP).

As we understand from Everaert et al. (2015), generative grammar is the study of an innate grammatical structure. The VP (NP) structure thus forms the core of Universal Grammar and so, purportedly, the core of all ‘natural’ languages. But the decision to place the object – and not the subject – within the VP appears to be rather arbitrary from a historical point. It was never based on research evidence at all and, since MiD nonetheless depends on this structure, it may be that the most obvious explanation for Hawkins’s conception is that it is an effect caused by the analytical device itself.

The VP (NP) stipulation is found in the article *Immediate constituents* by Wells (1947) who gives the example [*the King [of England]*] [*opened [Parliament]*]. Based on Pike (1943; cf. Figure 2), and discussions with Zellig Harris, Wells (ibid.) explains that

this depends on the object being a special case of "expansion" of the verb, citing Bloomfield (1933). Pike (1943) comments his own diagram as follows:

here a layer of tagmemes may be indicated, in which the top layer is the full actor-action sequence and a lower layer is the feature of object and action. (Pike 1943, 68.)

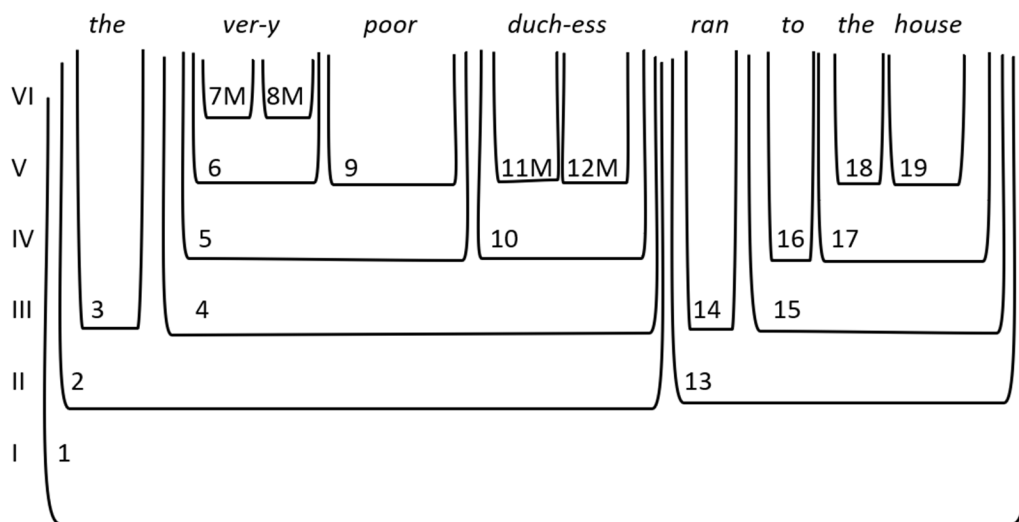


Figure 2. Pike (1943, 70) places the verb ran and the prepositional phrase 'to the house' into a common dependency, tagmeme 13.

The appeal of such analysis lies in the symmetry of the NP (subject) and VP which consists of the exact same number of layers or tagmemes. This perfect balance would of course be ruined if the subject was moved into the VP, or if the object was taken out of it. However, what Bloomfield actually wrote was that:

In English, taxemes of order appear in the difference between actor-action and action-goal, as in *John ran* and *catch John*; the difference between *John hit Bill* and *Bill hit John* rests entirely upon order. (Bloomfield 1933, 197.)

What we also observe in the early diagrams provided by Pike (Figure 2) and Chomsky (Figure 1) is that the lines seem to represent a standard school analysis of syntax where dependency structures are drawn under or above the text by the student.

According to Seuren (2015), Pike and his colleagues, who formed a school of thought which earned the nickname “God’s truth linguistics”, were Protestant missionaries whose main aim was to translate the Bible into preliterate languages. But they also believed that their analysis was a psychological or, indeed, “divine” reality (Seuren *ibid.*).

Nonetheless, the VP (NP) structure became crystallised in post-Bloomfieldian analysis. It looks as if Chomsky took whatever model was available at the time to claim it as a genetic structure. But there is admittedly something in the non-functionalist model of the Post-Bloomfieldian school that would work particularly well for Chomsky’s biolinguistics. That is, the eventual proof that grammatical sentences are generated irrespective of semantics; a rejection of functional grammar. The object of the sentence is not distinguished from the subject by any other virtue than that it occupies a different node in the IC-structure; something which only the trained expert is capable of discerning.

This could be used to promote Chomsky’s (1980) idea that *language* (i.e. Universal Grammar) cannot be learned by the child by deduction. The claim that the object – and not the subject – is in the VP is not logical. Nonetheless, it is *real*. Hence, the *logical explanation* is that it must be an innate structure and therefore universally present in all human languages.

Because Chomsky’s claim was never scientifically motivated, it is unsurprising that serious issues would appear as soon as the model was applied crosslinguistically (see Croft 2001 and 2003 for a detailed account), taking its toll on the credibility of the

innatism claim. It is also no great wonder that there has been a steady movement from generative grammar to Cognitive Linguistics. De Bot (2015) reports the decline of Generative Grammar in applied linguistics, revealing that universities now rarely hire Chomsky's supporters as professors. A 'usage-based' model, inspired by social Darwinism (see Boesch & Tomasello 1998, with a comment by Paterson), has taken over. This would have been an obvious scenario in syntax, too.

Here we come to understand the significance of Hawkins's research. Even if advocates of Chomsky's Minimalist Program remain true to innatism (Richards 2015), there are other syntactic models such as Lexical-Functional Grammar (Butt & Holloway King 2015), Head-Driven Phrase Structure Grammar (S. Müller 2015), Optimality-Theoretic Syntax (G. Müller 2015), Construction Grammar (Fried 2015; see also Goldberg 2006, Fillmore 2013, Bybee 2014) and others that have grown out of a rejection of many of Chomsky's claims while maintaining the VP (NP) stipulation. Chomsky (2002, 124, 2009, 36) has openly admitted that his theory has no basis in research evidence (see also discussion in S. Müller 2015, 962).

This seriously puts into question the whole meaningfulness of generative grammar which for decades has been the practice of designing solutions – “theories” – to make uncollaborative data fit the bias in the model. Like the Soviet space programme, it is not the machine that works but the genius of the engineer that *makes* it work, and it has a certain charm. Hawkins offers the generative programme a new *raison d'être* by providing crosslinguistic evidence for the cognitive reality of the VP (NP) structure, albeit with the cost of ditching the universal mode of explanation; a compromise which he nicknames “functionalism”.

If the desired structure is innate, then why are the data not consistent? This is where the Complex Adaptive System comes in. The language *organism* adapts to

human mind, but not perfectly, leaving statistical noise around the edges (Hawkins & Gell-Mann 1992, 14). Some may not be convinced, but there are other possibilities available. Osborne (2015) emphasises that Dependency Grammar, which was tested by Tesnière (1959) with free-word-order Latin, lacks the VP (NP) bias; while Van Valin asks:

what would linguistic theory look like if it were based on the analysis of Lakhota, Tagalog and Dyirbal rather than on the analysis of English? (Van Valin 1995, 461, as cited by Butler 2003, 121.)

Butler (2003) makes the point that functional grammar is unburdened by the problem of having to apply the purported VP (NP) structure where it does not fit. For example, there are OSV and VSO languages that separate the verb and the object by default, including Arabic, Filipino, and Celtic languages; and some that fix the subject to the verb: Spanish, Latin, Greek, Russian, Turkish; in fact, a majority of the World's languages (Dryer 2013a). Interestingly enough, the scientific plausibility of Darwinist typology now hangs on a generative thread: the one from the VP node to the NP node.

4.2 Hawkins's solution

Hawkins (1994) reiterates the cognitive reality of the IC tree, joining together a tradition of processing explanation in generative grammar and a heaviness explanation in functional typology. He observes that the conventionalised English V NP PP CP order within the verb phrase increases in length from left to right. The verb is typically a single word; the NP is on average longer than the verb but shorter than the PP which adds a preposition; and the CP (clausal phrase: a subclause), is the longest dependency. This is not only a question of grammar because Hawkins provides multiple analyses demonstrating that attested pragmatic choices in actual texts statistically prefer

organising the VP in a way that increases in length. Hawkins explains this as being based on processing efficiency.

In this view, the human brain is like the phrase structure parser whose first concern in sentence analysis is to figure out which way the object is ‘in’ the VP to draw a tree that computes. The constituency structure within the verb phrase supports it being done as quickly as possible. For example, the word order in the VP of the English sentence *I* _V[*gave* _{PP}[*to Mary* _{NP}[*the valuable book* _S[*that was extremely difficult to find*]]]] (Figure 3) is efficient because its IC structure (immediate constituents or dependencies within the VP) is recognised based on just four words, from *gave* to *the* (Hawkins 1994, 57). This can be seen in the diagram below, but also from the open brackets in the parsed sentence above, which all appear earlier from left to right than with any other ordering of these constituents within the VP.

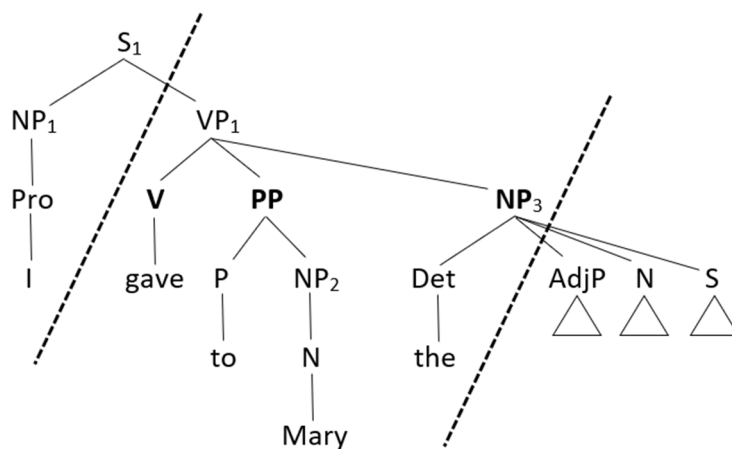


Figure 3. (Hawkins 1994, 60; highlight added to mark the relevant ICs). In the above ‘efficient’ word order, the IC structure of the verb phrase, consisting of a V, PP and NP nod, is recognised based on just four words between the dotted lines. This area was previously called the CRD (Constituent Recognition Domain) but is in Hawkins (2004) renamed as PCD (Phrasal Combination Domain). MiD argues that minimising this domain, as has occurred in the diagram, increases parsing efficiency.

A sentence of the exact same length, *I_v[gave_{NP}[the valuable book_s[that was extremely difficult to find_{PP}[to Mary]]]]* is not so common because the ICs have been arranged in an ‘inefficient’ way (see Hawkins 1994, 60 for a diagram). One will need to read eleven words, from *gave* to *to* to figure out the IC structure of the VP (Hawkins 1994, 57). MiD apparently requires the generative parse tree to be a cognitive reality of some kind. Hawkins concludes that, while processing efficiency does not necessarily constrain the grammar, it is an underlying factor for how word-order patterns become conventionalised and eventually grammaticalised, not only in English but in all languages.

Note that Hawkins does not discuss the possibility that the second word order is less common because it is centre-embedding which itself means more complexity (see further discussion in section 6.2; cf. Karlsson 2007). Yet another way to order the constituents might be ‘*I gave the valuable book to Mary that was extremely difficult to find*’ which we would consider ambiguous because the pronominal connective *that* appears to refer to *Mary*. But such issues relate to the economy principle (section 2 and 5), diverging from Hawkins’s processing explanation.

Based on English, the EIC principle would only work for less than half of the languages of the world. Hawkins, following a path of previous, similar but not quite so accurate models, doubles its explanatory power by studying SOV languages, such as Japanese which to an extent mirrors English word order. Parsing order is reversed from top-down to bottom-up in SOV. This does not actually mean that the sentence is read backwards, but if we examined the *closed* brackets right to left of sentence (1a) and (1b) below, we would observe an effect similar to the above English top-down parsing sentences.

Hawkins's Japanese example is somewhat different from the English analysis, so it may be necessary to quote his explanation directly. Hawkins takes up two different grammatical sentences for '*Mary said that John got married yesterday*'.

- (1) a. s₁[NP[Mary-ga] VP[š[s₂[kinoo John-ga kekkonsi-ta] to] it-ta]]
Mary-NOM [yesterday John-NOM got.married QUOTE] said
- b. s₂[š[s₁[Kinoo John-ga kekkonsi-ta] to] NP[Mary-ga] VP[it-ta]]
[Yesterday John-NOM got-married QUOTE] Mary-NOM said

Again, the reason for a weaker preference for (1a) is not because it is centre-embedding per se, but because it does not promote efficient constituency ordering. To start from the efficient (1b) alternative, Hawkins explains that the nodes

are parsed bottom-up, and once \bar{S} is constructed, the remaining ICs of the Matrix S2, NP, and VP, are attached within a very short and rapid viewing window: S2 {to_S, Mary-ga_{NP}, it-ta_{VP}}. (Hawkins 1994: 66—67.)

while the same does not quite occur in (1a), as indicated by the long line between *Maryga* and *it-ta*. Although the name EIC suggests that the constituency structure should be recognised as early as possible, it is rather the case in Japanese that the beginning of the sentence plays a minor role, leaving the vital structures to the end. For this reason the name ‘Minimize Domains’ (MiD) seems more appropriate, suggesting a closeness of IC nodes, whether towards the beginning or end of the sentence.

Hawkins does not provide diagrams for his Japanese examples. This could be because – as we can already see by comparing how the closed brackets are arranged right to left – the Japanese structures are not equivalent to the English ones. Parse trees are added here in Figure 4 for the reader’s convenience.

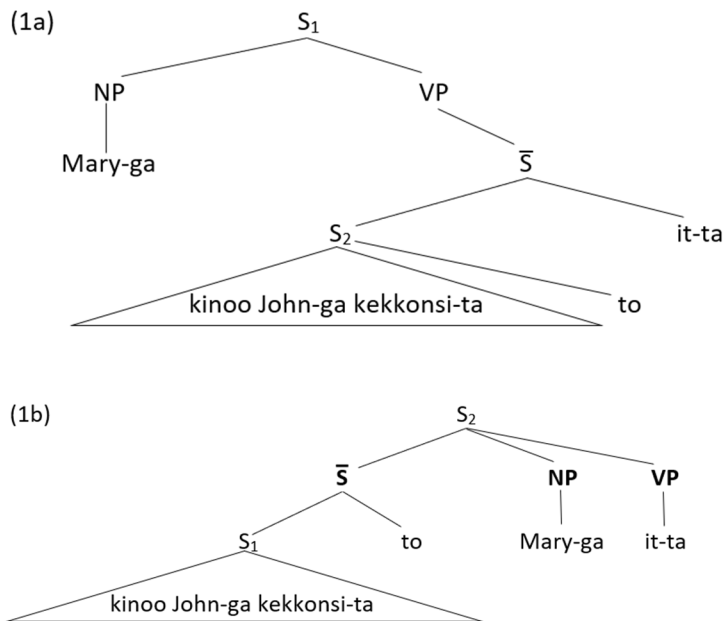


Figure 4. A tree diagram for sentence (1a) and (1b).

The first thing to observe is that the analysis does not work well for sentence (1b). Here the \bar{S} node is apparently meant to link the phrase in S_1 with the VP node, but we do not see \bar{S} being ‘in’ the VP anyhow, like the subclause was in the English sentences. The ICs of (1b), as highlighted in the diagram, are direct dependencies of the S_2 node, while in the English examples, they were dependencies of the VP. The Japanese comparison depends on including the subject (NP: *Mary-ga*) which was not the case in English. In fact, the typological data purportedly supporting MiD (section 6) are supposed to ignore the subject altogether, linking the place of the adposition and the subclause with the OV/VO order. But the desired effect in Figure 4 is caused by moving the subject only.

It was Hawkins’s starting point that a simple heaviness principle will not work because the verb is statistically the shortest and the object the longest of the three main elements (subject, object, verb; Hawkins 1994, 20). Heaviness would then predict that languages will either have the VSO or OSV order which is not correct. Therefore MiD excludes the subject, something which works well with the typological data but is not

supported by Hawkins's Japanese examples. Suffice it to say at this point that we are not fully convinced of the accuracy of MiD.

It could be the case that there is something else at play that could provide a more plausible explanation for the observations made by Hawkins and his colleagues (cf. especially Dryer 1992). We carried out our own text analyses in different languages to find clues of what might actually be happening underneath (section 5). Based on our findings we examined the related typological data, defining a different principle that makes equally good predictions as MiD (section 6.3). Finally, we compared the two, rejecting MiD because it is the model that makes the most complex claims for the same predictive accuracy (section 6.6).

5 Own research

We carried out an analysis and an experiment of a text available in various languages. Our first task was to find out whether CD (given before new) or CTU (new before given) was the prevalent theme—rheme order in Czech, English, Finnish and Swedish (section 5.2). As all four texts supported CD, we then performed an experiment to find out whether the given-before-new ordering is specifically functional, by inverting the information structure pattern (section 5.3).

We will present our empirical research in a rather compact form due to the mostly theoretical orientation of this article, and given space limitations. In fact, the theoretically-inclined reader may wish to jump directly to section 5.4. Sections 5.1 to 5.3 may be interesting for those who want to know more about our study of information structure. It becomes evident as our study progresses that functional discourse analysis based on the economy principle, our chosen approach, raises interesting questions everywhere one looks and in all languages. There will nonetheless be little effort from

our part to answer them all in a scientifically rigorous way. We will limit our focus to serve our goal which is to provide a reply to Hawkins's question (section 2).

5.1 Material

We analysed a text published by the European Commission (EC 2014) which is available in 23 languages although only four were selected for our study. The English title of the text is *Organic farming – a guide on support opportunities for organic producers in Europe*. It is a brochure which provides information especially for farmers interested in converting their animal or plant production to organic.

One reason for choosing this text was because it appeared first on a list of similar brochures on an official document website and seemed to cater for our needs. We were especially interested in a text which is moderate in length but with a rather high frequency of complex sentence structures, although in a naturally flowing form. Thus we could manage the necessary amount of data in a rather compact file. Each language version provided close to 200 cases when the complex sentences were split into clauses: Czech (194), English (171), Finnish (177), Swedish (217). These cases were inserted into spreadsheets to conveniently calculate sums and percentages based on our analysis.

Because the early discussions between proponents of CD and CTU had involved Czech and English, our preliminary idea was to limit our observations to these languages. However, it was in particular the Swedish version of the text that raised our interest due to its very high frequency of the relative pronoun *som*. There are as many as 45 tokens of the word in the text, excluding footnotes and such, 43 of which function as the relative pronoun (and two as a preposition). By comparison, the similar Danish relative pronoun *som* was observed to appear 25 times in the Danish version of the text which was then not further analysed, although it is also a very high frequency: the

English *which* only appears 14 times; but there are of course other relativisers, too. We nonetheless used the Swedish text for sandboxing to uncover the most consistent way of analysis; but since the results from Swedish were also interesting, we eventually included them in our account.

Because Swedish and English seemed to form a kind of pair, a Finnish version was added to support results from a Czech analysis. This way we had two languages, English and Swedish, with a strict word order and grammatical articles, and another two, Czech and Finnish, with a flexible word order and no articles. As regards language families, English and Swedish are Germanic; Czech is a Slavic (or Slavonic; Indo-European) language, and Finnish is a Finno-Ugric language; but language structuration was thought of being more important than language genealogy.

5.2 Counting the GN to NG ratio

We made a discourse analysis of the text(s) with the task to extract the given-before-new (GN) and new-before-given (NG) ratio. The GN versus NG ratio was analysed in the Czech, English, Finnish and Swedish material. Complex sentences were split into clauses each containing its own finite predicate. Arguments of each predicate were separated to be analysed as thematic—rhematic elements (cf. Hajičová 1994). A fully precise definition of theme—rheme was initially not quite crucial for our analysis of the order of given and new information which were relatively easy to recognise in the text, at least in principle. Any concept having appeared earlier was marked as given information. To give an example, *the European Commission*; *organic farmers*; and *the common agricultural policy (CAP)* first appeared in an opening summary or abstract, and were thereafter always considered as given information.

The most common type was GN, but NG sentences were also rather frequent. Question sentences were one common type to have the NG ordering, and there were a

few questions in the material, e.g. “*What should organic producers know?*” The relative pronoun, on the other hand, was a frequent predictor of the GN ordering, e.g.

- (2) [This regulation ...] **which** lays down *detailed rules for its implementation*.

Here, given (old) information is highlighted with bold type and new information with italics. Although all four languages have a similar information structure, sentences in the different versions are often not quite equivalent to each other because writers of the different versions had structured the text in somewhat different ways. Below are examples of GN sentences with a relative clause in different languages, applying the publisher’s language coding (cs = Czech; en = English; fi = Finnish; sv = Swedish).

- (3)

- cs [udržitelné zemědělské postupy,] **které** jsou *prospěšné pro klima a životní prostředí*.
 ‘[sustainable agricultural practices] **which** are *beneficial to the climate and the environment*.’
- en [conversion] during **which** *the provisions concerning the organic production* have been applied.
- fi [alat ...] **joilla** kasvaa *kerääjäkasveja tai typpeä sitovia kasveja*
 ‘[areas...] **which** are *covered with catch crops or nitrogen fixing crops*’
- sv [kommissionens förordning...] **som** innehåller *närmare regler*
 ‘[an EC regulation...] **which** contains *detailed rules*’

As is seen from these examples, the predicate was not considered as a candidate for a thematic element, only its arguments were (see further discussion below). There are in general many types of cases and orderings, with an English example of GN, NG, GG and NN below.

- (4)

- GN **These equivalent practices** shall be based on *agri-environment schemes and certification schemes*

NG *Support for conversion to organic agriculture* is foreseen **under the CAP**.

GG **Processors and retailers** can also join **the EU organic scheme**

NN *The most relevant sub-programme, in terms of content and financing, would be the one on climate change mitigation and adaptation and biodiversity.*

The last argument was new information in the context despite being marked with a definite article. In contrast, the theme or topic ‘*Processors and retailers*’ in the above GG sentence was recognised as given information because these had already once previously appeared in the text. There were additionally simple G and N sentences, and more complex information structures with multiple arguments, such as GGN, NGN, GNN etc., for example

(5)

GGN [...while] **the SFVS operates on the principle of co-financing** *with the total EU budget of €150 million per school year.*

Here, the predicate *operates* has three arguments. Only sentences with a clear GN or NG structure were included in the main results to keep the analysis, and especially the conversion that will follow, as simple and clear as possible, but the overall number of G-first and N-first sentences will also be indicated below. There were additionally many incomplete sentences without a predicate, most often captions, which were systematically excluded from the results.

While the predicate itself can be considered as having its own role in the theme—rheme chain (Firbas 1971), there was only one instance in the text where the event was referred to anaphorically. To simplify the analysis, we decided to disregard the eventual thematic role of the verb altogether, focusing on NPs and PPs only, which gave us a straightforward way of calculating a purely binary GN-to-NG ratio. A similar

stipulation was made by Payne (1987) and her sources. In principle, as any predicate describes an event, and each event in a discourse is normally unique, including the verb in a GN analysis would seem likely to bring about an avoidable bias.

An English example of an RP referring to a verb is: “The new Rural Development programmes have a more flexible structure **which** makes it easier for Member States to combine different types of rural development measures.” In this case we simply analysed the RP as referring to the noun ‘structure’ because there was enough ambiguity to justify it. But there was also a genuine case, here cited from the Finnish text (in parenthesis; with the proper antecedent underlined):

- (6) Uusi järjestelmä kannustaisi oppilaitoksia hankkimaan luomutuotteita (**mikä** on mahdollista ja käytäntönä jo nykyisessä kouluhedelmä- ja vihannesjärjestelmässä) [...]
- ‘The new system would encourage educational establishments to buy organic products (**which** is already a possibility and practice in the current school fruit and vegetable programme) [...]

That was, then, not considered significant enough to necessitate a more complicated analysis for the whole text.

Another stipulation made to hopefully ease crosslinguistic comparison was to elucidate information structure in cases where given information was clearly implied but not made visible, by repeating it in the corresponding place of the second sentence, e.g. ‘*If you are a farmer and wish to join the European organic scheme [...]*’. Because there are two predicates, the structure is split into two sentences, repeating the theme here in curly braces:

- (7) GG If **you** are a **farmer**
 GN and {**you**} wish to join *the European organic scheme*,

In the English text, for example, such bracketed cases accounted for seven GN sentences out of 77 in total, or 9 percent.

Quite frankly, analysing the given versus new information structure is not always quite simple because there are so many different types of cases in real texts. We must admit that many sentences could be analysed differently, and that we could have made different methodological choices. We did in fact do initial sandboxing with a much more detailed analysis discarding the classical theme—rheme division and simply focusing on new versus old words. We then rejected this approach because it resulted in some very complex structures (GGNGGNG and the like), and we felt that the classical theme—rheme concept was more helpful, and that repeating the theme where it was omitted increases contextual coherence and could avoid false biases especially when applied to strongly “pro-dropping” samples. That is, in texts in languages that place the anaphora into the verb or the conjunction, or omit it altogether assuming a conventionalised reading.

In our experience, a different analysis could show a lower share of GN sentences, likely settling at around 75 percent. Our suggestion is nonetheless that, any which way, there could be no reasonable analysis of a regular text in any of these languages that could show a higher NG than GN ratio.

| | GN (of applicable cases) | NG (of applicable cases) | All G-first cases | All N-first cases |
|---------|--------------------------|--------------------------|-------------------|-------------------|
| Czech | 80 (88.9%) | 10 (11.1%) | 119 (82%) | 36 (18%) |
| English | 77 (91.7%) | 7 (8.3%) | 102 (76.1%) | 29 (23.9%) |
| Finnish | 96 (93.2%) | 7 (6.8%) | 120 (79.7%) | 33 (20.3%) |
| Swedish | 114 (95.0%) | 6 (5.0%) | 141 (76.8%) | 43 (23.2%) |

Table 1. Results of the GN:NG analysis show a preference of given before new information in all four languages. GN = given-before-new information structure (sentences); NG = new-before-given information structure, with the GN:NG ratio in percentages in brackets. All G/N-first cases include sentences with simpler or more complex information structure that what were selected for the main results in the second and third column from the left.

Results from our analysis in table 1 show that the GN ordering is clearly prevalent in all four languages. What was perhaps surprising in the results is that English shows no divergence from the other three languages. Givón (1988) considers it as having a canonical NG ordering, while Payne (1987) did not take a stance on English, but implies agreement with Firbas (1964), that Czech has dominant GN. We expected the GN:NG ratio to be at least a little lower in English, but there is no clear indication of any difference (in this relatively small sample). Thus, it is no wonder that Givón's CTU was not accepted in its time (section 3.1) and that it was subsequently replaced by EIC/MiD.

We will use this analysis as a basis for the next phase, inverting the information structure of all of the GN sentences to NG.

5.3 Inversion of information structure

To recap, Hawkins (1994) asks why GN is supposed to be functional. We have discussed the meaning of 'functional' in section 2 based on standard sources on the topic. Our answer will consequently be based on demonstrating that the prevalent GN ordering supports linguistic economy.

Our experimental approach means examining what will happen if the word order of sentences in our text now containing a GN order is changed to NG. A conversion of the Czech text was not done at this time. We carried out an experiment to see whether economy of English, Finnish and Swedish sentences would be affected by the change. This was done simply by inverting all of the GN sentences to NG, e.g. "**Support** is also foreseen *for the maintenance of organic agriculture*" is changed to '*for the maintenance of organic agriculture* is **support** also foreseen'. An ordering that sounds most natural is preferred, although it will be ultimately ignored if the converted sentences seem

unusual. Instead, based on the economy principle, it will be assessed whether reversing the information structure causes an increase of lexical/morphemic complexity, or syntactic—semantic ambiguity in the sentence itself or in an adjacent sentence.

As expected, the analysis of English and Swedish showed an increase in complexity and ambiguity when the word order was changed. Complexity increased typically when the GN theme, which contains the ‘given’ subject, was moved after the rheme, which contains the ‘new’ object, because this requires the addition of an agent marker. The following Swedish example has the same meaning as the English sentence below, each cited from our material.

- (8a) sv original: **en jordbrukare** måste odla *minst två grödor*
 conversion: *minst två grödor måste odlas av en jordbrukare*
- (8b) en original: **a farmer** must cultivate *at least two crops*
 conversion: at least two crops must be cultivated by a farmer

The added elements are underlined here. In addition to the agent-marking preposition *av* (by), and a passive -s or participial (-ed) suffix is added to the verb. English also adds the copula (be). The number and quality of any obligatory additions was however not considered as essential, but the fact that the converted expression is lexically/morphologically more complex than the original. That is, there are more words or morphemes in absolute terms. Each conversion which required the addition of any element was marked as having increased complexity, compared to the original.

A second way that decreased economy of sentences was the increase of ambiguity. This sometimes occurred when a relative pronoun containing given information was moved behind new information from its sentence-initial place. All such sentences became ungrammatical, e.g.

(9)

(original:) [funding possibilities,] **which** aim to increase consumer awareness[.]
 (conversion:) [funding possibilities,] to increase consumer awareness **which** aim[.]

This lower converted sentence is actually marked as being fine from the economy point.

Since the plural intransitive verb *aim* cannot refer to the singular NP *consumer awareness*, ambiguity is not strictly speaking increased. In fact, moving the relative pronoun does quite typically not increase ambiguity in absolute terms, even if sentences become ungrammatical, due to various reasons such as the one seen in the above case.

There were many conversions nonetheless resulting in increased ambiguity, e.g.

(10)

original: [challenges] **that** were posed by increased competition and fluctuations of markets.
 conversion: [challenges] by increased competition and fluctuations of markets **that** were posed.

This particular conversion was deemed as adding ambiguity because the phrase *that were posed* now seems to refer to *markets* rather than *challenges*. While it was rather obvious from the onset that reversing the information structure would decrease economy due to increased complexity and ambiguity, somewhat surprisingly, this did not quite occur in Finnish which was clearly less affected by word-order change (table 2). To be precise, conversion of Finnish sentences did increase ambiguity, but not complexity. This appears to be due to systematic case marking which allows complication-free word-order inversion, e.g.

(11)

original: **jäsenmaat** voivat soveltaa *myös muita käytäntöjä*
 'member states can also apply other policies'
 conversion: *myös muita käytäntöjä* voivat **jäsenmaat** soveltaa
 'other policies can also be applied by member states'

As can be seen above, unlike the English translation, the Finnish conversion does not have any increased complexity compared to the original.

| | Unaffected | Increased complexity | Increased ambiguity | Ambiguity due to RP shift |
|---------|------------|----------------------|---------------------|---------------------------|
| English | 49 (64.5%) | 17 (22.1%) | 12 (15.6%) | 12/12 |
| Finnish | 84 (85.7%) | 0 | 14 (14.3%) | 14/14 |
| Swedish | 77 (67.5%) | 22 (19.3%) | 17 (15.0%) | 17/17 |

Table 2. Unaffected = number of GN sentences with economy not affected by a conversion to NG. ‘Ambiguity due to RP shift’ = number of cases where ambiguity increased due to moving the relative pronoun. Ambiguity increased similarly in all three languages, but complexity only in English and Swedish. The sum of English and Swedish percentages exceeds 100 due to, respectively, 1 and 2 cases with both increased ambiguity and complexity.

The only common denominator for all three languages is shown to be limited to concerning **the relative pronoun**. Moving the relative pronoun away from its clause-initial position caused referential (syntactic—semantic) ambiguity.

5.4 Discussion of results

Based on our research above, we are now ready to answer Hawkins’s question: *why is the GN ordering functional?* We linked functional linguistics with the principle of economy in section 2. Our research shows that reversing the GN ordering reduces economy by adding complexity or ambiguity. Therefore, the answer is that GN is functional in given language systems because it supports disambiguation without the increase of complexity.

Many details will have to be left for further research due to the complexity of the issue. Also, results from our study are clear, but reasons causing them can only be a subject to speculation or hypothesising. It can nevertheless be said that they do not seem to support sociobiological concepts in any way. As regards Darwinian linguistics and

generative grammar alike, there does not appear to be anything in the material to necessitate a non-intelligent cause for linguistic structure and variation. The simplest explanation is that language is functional because it is a tool of communication, and people make it so to serve their own needs. It is not so much to prove the null hypothesis, but to ask: *how else could it be?*

Our study furthermore does not support the generative concept of wh-movement. In our Czech, English, Finnish and Swedish material, there was no wh-movement whatsoever, in absolute terms. Instead, wh-words functioned as conjunctions, appearing clause-initially, whether they were question words or subordinators. Wh-words in GN sentences always carried old information. When moved after the rheme, an experiment we carried out in English, Finnish and Swedish, relative clauses became prescriptively ungrammatical.

More interestingly, it was only this type of conversion that caused an ambiguity increase (of ca. 15 percent) while other types of GN-to-NG changes made no impact. Our conclusion based on this must be that *grammar* supports disambiguation. In contrast, it remains fully unclear to us why the object, when encompassed in a wh-word (e.g. *what*, *who[m]*) should ever be fronted in the first place, according to Darwinian linguistics and generative grammar.

Out of the three tested languages, only English and Swedish saw an increase in complexity due to a GN-to-NG conversion. This however neither affected grammaticality nor disambiguation because the two languages have a specific mechanism of prepositional marking available for this type of word-order change. The canonical word order in these languages thus supports economy because the agent is only marked when necessary. The GN-to-NG conversion caused a case-by-case increase

of around 20 percent of complexity, suggesting it is more economic in terms of lexical/morphological complexity than systematically marking the agent.

The classical structuralist concept of different languages doing the same job in different ways – that is, expressing predicate—argument-structure in an economic way – is neatly illustrated by comparison with Finnish which rather systematically marks the subject—object distinction. While this would seem to be an unnecessarily costly way, Finnish subject-last sentences are clearly simpler. Such sentences should be equally vital for all three languages since all use the same word-order strategy to help link the relative pronoun with the correct antecedent; that is, by moving *the antecedent* next to the relative pronoun regardless of canonical agent—patient order.

It can be argued in the current perspective that the English and Swedish strategy is optimal for simplex sentences, but the Finnish strategy is optimal for complex sentences, especially in the rather prototypical case where new information is elaborated in a relative clause. At this point however we will not proceed to examine whether one of the strategies is more optimal than the other in absolute terms. Instead, we believe we have found enough information to continue in the next section to systematically reject EIC/MiD as a useful concept for language sciences.

6 A rejection of MiD

In the last section we provided an answer to Hawkins’s question why given-before-new ordering is functional in four European languages. In this section we will proceed to dismiss his processing explanation of word order harmonies in worldwide typological data by demonstrating that there is a simpler principle that makes equally good predictions. Thereby MiD will be overruled per Occam’s razor. We will start our quest by enumerating weaknesses of MiD.

6.1 *Weak points of MiD*

The main problem in Hawkins's model is the same as in other versions of sociobiological linguistics: the absence of the null hypothesis. As the objective is to explain language universals or statistical tendencies in crosslinguistic data, since the beginning of Chomsky's biolinguistic revolution, the starting point always seems to be that there has to be a biological explanation of some kind. This is despite no evidence being presented of a single linguistic form being biologically-based.

In scientific practice, on the other hand, the null hypothesis serves as a starting point and as a general reference. There is plenty of historical evidence of people having created linguistic forms: just take Shakespeare and Noah Webster as examples. Therefore, the simplest explanation for any linguistic phenomenon is that it is made by people because this requires no further sources than what we already have. Identifying language as a man-made creation is the starting point of the null hypothesis, and any biological or quasi-biological explanation (such as Darwinian linguistics, see section 2) is necessarily making an *alternative* hypothesis. The technical difference is that, while the alternative hypothesis must be supported by relevant evidence in order to qualify as a scientific argument, the null hypothesis in fact does not require any further evidence to be acceptable as such.

This kind of thinking seems to be totally alien to Culbertson & Kirby (2016) who consider Occam's razor as a cognitive "bias". So, every time sociobiology fails to find evidence – and this is literally *every time* – it is somehow taken to prove itself.

Hawkins believes that

There is clearly a biological and an innate basis to human language. That much is obvious. The issue is: precisely what aspects of language and of language use are innately determined? Our species has grammars and others don't, so it is tempting to assume simply that grammar is innate [...] (Hawkins 2004, 10.)

As was the case with Haspelmath (1999, as discussed in section 2), it is forgotten that the purpose of science is not to accept the “obvious”, but to question it. It may be true that our species has grammars and others do not, but dissimilarities do not end there. For example, we have cars while other species do not (as much as we now know of our universe). This may obviously have *something* to do with biology because if we were not anyhow *biological* we would probably not drive cars; or do anything else for that matter. But to what extent should this be relevant for humanities?

The logic of sociobiology is illustrated by the following example. We look outside the window and observe that all people drive on the left side of the road, whichever way they are travelling. As this is unlikely a mere coincidence, the question will be whether driving on the left is a genetically determined characteristic of *homo sapiens*. As it turns out that there are other places in the world where people actually drive on the right, the innate hypothesis is rejected in favour of CAS: a system of cultural replicators which adapts to its environment, causing some regions to have right-hand traffic and others to have left-hand traffic.

Following the principle of natural selection, it will be conjectured that people driving on the left have a higher survival rate and produce more healthy offspring in some regions, giving these individuals an evolutionary advantage and thus helping the left-hand CAS to become dominant. It may then be observed that this left-hand ‘adaptation’ is statistically more prevalent on islands, subcontinents and small continents, while the right-hand adaptation has spread primarily across the larger continents.

In much of today’s human sciences the whole notion of free will, creation or intelligent design seems unthinkable as an explanation of human behaviour. As it is stated by Bybee, Christiansen, Croft and their colleagues:

Linguistic patterns are not preordained by God, genes, school curriculum, or other human policy. Instead, they are emergent[.] (Beckner et al. 2009, 18; italics added.)

Here, the concept of social construction is equated with religious creationism, which is wrong; language does not “emerge” without human creativity. If this principle does not work in biology, it only means that human sciences and natural sciences are not analogous.

As regards MiD, it was needed when predictions made by Chomsky’s innatism and Givón’s cognitive explanation failed the test of crosslinguistic predictability (section 3). Hawkins’s solution was to “reduce” the innate grammar by adding a second sociobiological principle, CAS, to cater for all needs. In place of just one unsubstantiated alternative hypothesis, there are now two. This is not standard scientific practice: the null hypothesis – that biology plays no specific role – has not been elaborated in a satisfying manner or, in fact, it has not been elaborated at all.

The main weakness of Hawkins’s crosslinguistic analysis was discussed in section 4.2. MiD seemed to rely on ignoring the subject and only focusing on the hypothesised verb phrase as relevant for word order harmonies, as is standard practice in typology (Dryer 2013acd). However, we noticed that the subject was elementary for Hawkins’s analysis of Japanese, and that the analysed IC depth varied between the languages. We suggested that the full parse tree approach seemed excessive as one could have more simply just counted the open or closed brackets in linear representation.

As regards typological predictions made by MiD, it makes a very accurate prediction of OV/VO + adposition harmony. OV languages are clearly more likely to have postpositions than prepositions, e.g. Japanese, while VO languages are clearly more likely to have prepositions than postpositions, e.g. English (Dryer 2013c).

Predictions for OV/VO in relation to the order of the noun and the relative clause are weaker. The VO&NRel order (head noun before relative clause, e.g. English) is a very strong tendency over VO&RelN (relative clause before head noun) crosslinguistically. But same cannot be said of OV&RelN (e.g. Japanese) which can either be described as a weak tendency or as one that is limited to Asia (Dryer 2013d).

6.2 What is known about language processing?

Hawkins (1994, 2004, 2014) makes strong claims concerning language processing, processing efficiency and processing difficulty, but the question remains how much is really known about such concepts.

Today's cognitive linguistics is dominated by the competition of two contrasting sociobiological approaches. Schwarz-Friesel (2012) discusses two combative dogmas in the field that make extreme generalisations of how the brain and mind work. One is the Chomskyan or generative approach which claims that language is compartmentalised and thus unaffected by general processing principles. The other one is the Lakoffian—Langackerian branch which goes to the other extreme by claiming that language processing follows purely from the general principles.

We suspect that the motivation for both stipulations is to make people believe that intelligent construction of language is totally impossible, thus pretending to justify the case for sociobiological linguistics. For the Chomskyan camp, intelligent creation and intelligent learning of syntax is not possible because language and mathematics are locked into two separate mind modules, with no connection in-between. Fortunately, Chomsky has been clear that his theory is not based on any evidence whatsoever (Chomsky 2002, 124, 2009, 36; cf. Robbins 2017).

For Langacker (2010), language is conceptualisation, and for Lakoff & Johnson (1999), the mind operates on cognitive metaphors. Hence, language cannot be logically

motivated because mathematics is just another metaphor among all the rest (Lakoff & Núñez 2000). Mathematics or logic cannot function as the universal measure of anything, so our economy approach must be doomed to fail.

The mind-view of so-called Cognitive Linguistics (CL; capitalised) is however counter-evidential. It is well known that, even though these are not modules, the brain has several specialised regions for language processing. Language is not just processed according to general principles, but differently from other types of information such as visual input (Schwarz-Friesel *ibid.*); and that, though these are not the same, there is an indisputable connection between mathematical, analytical and linguistic processing (Zhou et al. 2018). What is cognitive linguistics supposed to be, anyway? Peeters (1998) makes a reasonable suggestion.

It would seem to me that the latter is the sort of linguistics that uses findings from cognitive psychology and neurobiology and the like to explore how the human brain produces and interprets language. In other words, cognitive linguistics is a cognitive science, whereas Cognitive Linguistics is not. Most of generative linguistics, to my mind, is not truly cognitive either. (Peeters 1998, 226—227.)

Butler (2008) urges functional grammarians to seize the ‘cognitive’ trend with a synthesis of linguistics, psychology and neurology. This is easier said than done – the practice we are literally witnessing in current linguistics – because, as aptly summarised by the following passage on neurolinguistics,

though progress is being made in this field, very little is known for certain about the neurological aspects of language. (Encyclopaedia Britannica 2015.)

This is no wonder because neurolinguistics needs a model of language as a starting point for neuro-syntactic research, and this is typically either provided by generative grammar (e.g. Schlesewsky & Bornkessel-Schlesewsky 2013) or CL (e.g. Lai,

Howerton & Desai 2019). But the two models themselves are not supported by brain science, so much of today's cognitive trend cannot be based on more than mere speculation. Therefore, while working out our model (section 6.3), it is difficult to claim true cognitive adequacy when no one knows what that might be exactly. There is however something that can be done. Guided by Schwarz-Friesel's point that both the Chomskyan and the Lakoffian—Langackerian model are *too extreme* to be supported by research, the middle road would seem to be a reasonable way towards a more useful model of language.

From a syntactic point, this could be illustrated by the case of grammatical rules. Grammar is not just conceptualisation, but a set of conventional rules. The meaning of 'rule' in *generative grammar* is of course different from the usual meaning of the word, referring to mathematical rewrite rules (Chomsky 1957). It must be noted that these are simply the instructions of how to draw a parse tree, and it is the tree in particular that seems to be the creation of syntacticians rather than that of the unconscious mind, as far as seems possible to demonstrate scientifically (see further discussion in 6.5).

We also know that the brain is good at making associations (Eysenck 2012). Therefore, a plausible model could be one with a rather flat linear structure. The point of verbal thinking seems to be that, rather than thinking in pictures or diagrams – although of course capable of visualising things – people adopt a conventionalised linear language for verbal thinking. A proper cognitive grammar can be based on associations. Some structure may be real, although unlikely innate, but it is not certain what centre-embedding is exactly, for example. People do not have nonlinear representations at their disposal when they hear and read language. Centre-embedding structures appear as inserts in the linear form, and this is the reality that people will have to operate on.

As regards MiD, Hawkins (1994) observes that centre-embedded structures can be grammatical, but they tend to be less common than left or right embedding. His conclusion based on this is that grammaticality questions are not governed by a universal grammar, but emerge statistically from processing preferences which are explained in terms of placing the IC nodes of the VP closely together so that the tree can be easily constructed – hypothetically speaking. It must be noted from the economy perspective that centre-embedding structures (i.e. clauses with inserts) are more complex than left or right embedding. When the subclause is centre-embedded, the sentence structure is $A_1 B A_2$ with the main clause A divided into two parts. Thus having the full sentence divided into three parts is more complex than $A B$, or $B A$, each of which only contains two parts in all. Therefore the fact that centre-embedding is less frequent is predicted anyway.

A second way to address centre-embedding is with an appeal to psycholinguistic research. Papagno & Cecchetto (2019) take centre-embedded relative clauses as a standard example of the type of syntactically complex sentences that load on short-term memory. We believe that many of Hawkins's findings are more naturally explained through this concept rather than the early immediate constituents (section 6.3 and 6.5).

The question is, could there be a model of sentence analysis that works on flat linear structure: on language itself? And if that is possible, can this same model make equally good typological predictions as MiD? It would at least make a lot of sense to say that if grammarians stopped studying things that are not known to be there and accepted standard scientific concepts including the null hypothesis, there should be a more solid ground for achievement in syntax, typology and cognitive linguistics.

6.3 A non-biological explanation

We must welcome MiD as the best explanation of harmony so far (Song 2010), but have raised issues on its theoretical foundation (section 2, 4 and 6) and would prefer a non-biological explanation of word-order variation, for the lack of any evidence for a biological cause of linguistic forms and patterns. Our chosen approach follows a tradition of rational explanation of language universals with a focus on pragmatics (Haberland & Heltoft 1992).

We studied crosslinguistic data related to OV/VO with RelN/NRel and NPostp/PrepN (order of head noun and adposition) to reiterate that, while information structure does play a role in word order variation, it is unlikely to serve as a primary principle. We did not extend our own research to Japanese, for example, but observed that Japanese lacks relative pronouns and uses a special particle for topicalisation (Kim 2012). To what extent the findings from our research of information structure can be applied to Japanese and other OV languages will be the subject of a different study.

We moved to examine grammars of different types of languages, case studies and mass comparisons to find a pattern that indeed provides a simpler explanation of harmony. This examination is guided by background knowledge from neurology (section 6.2). To start with a factual statement, the pattern we uncovered is that languages around the world, much more often than not, have grammars that place adpositions and conjunctions, or what we collectively call connectives, towards the element that is referred to by the phrase this connective belongs to.

We will make the case in the following subsections that this principle makes equally good predictions as MiD. The rest of this subsection will give some examples of what is meant by our principle of *alignment*; while the next subsection (6.4) will continue to speculate on what the observed phenomenon might depend on. As was suggested by our own research, the relative pronoun is a type of connective which is

preferably placed immediately after its antecedent, making the two easy to associate.

The principle of alignment entails that prepositions are used in a similar way.

Like Hawkins, we will use English as a starting point. In the following two examples, the open pointy brackets mark the referential direction from the phrase to its antecedent.

- (12) a. A man <**who** wears glasses is sitting in the front row.
 b. A man <**with** glasses is sitting <**in** the front row.

The point of reference of each connective phrase is underlined above. In 12b, the restrictive prepositional phrase ‘*with glasses*’ corresponds to the relative clause ‘*who wears glasses*’ in 12a. The second PP in 12ab corresponds to informing that *sitting* is an action which occurs in the front row. The difference between prepositional and postpositional alignment can be illustrated with English possessives, namely the preposition *of*, and the genitive suffix ‘*s*’ which corresponds to postpositions in crosslinguistic data (e.g. Dryer 1992).

- (13) a. the cousin <**of** the landlord <**of** a friend
 b. a friend’s> landlord’s> cousin

These two mirroring chains can be described as being semantically logical or coherent because changing the order would change the meaning (see 6.4 for further discussion). Adpositions are used to physically link concepts with one another in a way that makes sense from an objective perspective. The order could be different: a given language may apply the same meaning for a structure resembling ‘*of the cousin of the landlord friend*’ or ‘*a friend landlord’s cousin’s*’. Such orderings, as long as they are consistent, could also be described as ‘logical’; but having the particle pointing away from their reference is markedly less common in the overall data.

A separate connective for each case, for both left and right alignment, is typically not available in a given language. Instead, the data indicate the canonical place of adposition and conjunction, and this is fixed with the verb. It is either that the fixed place of the connective arises as a consequence of the OV/VO order, or the other way around. But once the alignment is decided, there will be few exceptions, even when relating to pre-verbal elements like the English subject. Here, the referential direction is constantly leftward although not towards the verb:

(14) The sheep <**that** is <**in** front <**of** the cow <**that** is running (...)

This is because SVO is verb-medial, so the referential direction of the preposition will not always be towards the verb. Same is true for clause-initial adverbials of time, place and manner. It is in principle possible in *verb-final* languages that the postposition always points towards the verb. The following Turkish SOV sentence is the translation of the above English phrase.

(15) *Koş-an ineğ-in ön-ün-de-ki koyun*
run-(y)An cow-GEN front-POSS3sg-dat-rel sheep
(Özge, Marinis & Zeyrek 2010)

The referential structure of the Turkish phrase is

(16) V **Postp**> N **Postp**> N **Postp**> N (...V)

linking content words with connectives. The predicate will appear at the end of the sentence. As Hajičová (1994) argued that patterns of information structure depend on the communicative strategy chosen by the speaker, we linked information structure with the relative pronoun (section 5) which we then identified as a type of connective. But the conventionalised use of other types of connectives, to which our definition includes the adposition, can also be the result of people's effort to successfully convey their messages. Thus, language may be a rational or intelligent design made by people. What

it could be alternatively, no one knows, based on the lack of evidence for any other cause.

Word order is typically free or flexible for marked sentence parts, and since the PP is marked with an adposition, there could be several possibilities of placing it. The problem is how to recognise which sentence part the free element refers to. In principle, there is no necessary difference between saying ‘*The lady at the shop saw the president*’ or ‘*The lady saw the president at the shop*’. However, it seems that consensus in the speech community is typically reached when the connective points to its reference: The man <**at** the shop; versus: saw <**at** the shop. The way the adposition is placed apparently gives the hearer a clue of what it refers to. It is ideally as close to the reference as possible, which leads to a problem of ordering complex sentences when there are several candidates for the closest position.

6.4 Generalisation, explanation, hypothesis, and theory

The subsections below will propose explanations for the prevalence of alignment in crosslinguistic data. First, to try to keep things clear, we will briefly discuss some terminology. Egré (2015) makes the point that the word ‘explanation’ is used in many ways. In this paper, we will follow a four-way distinction:

(i) The first level of explanation is the *typological generalisation*. In the case here we are discussing generalisations under the headline *harmony*, e.g. OV + postposition versus VO + preposition.

(ii) The second level is the explanation of the generalisation, or a *principle*. For us, it is the *alignment principle*: harmony of the order of object and verb, and the place of the adposition, for example, depends on the connectives being aligned with their point of reference, especially the verb.

(iii) For speculation as for what causes languages to have alignment, the word *hypothesis* can be used. We will discuss different hypotheses that may account for the emergence and stability of the alignment. The traditional terms formal and functional explanation are used to group types of hypotheses.

(iv) The highest-level term, *theory*, is here reserved for theory of language. There are three: biological determinism (or innatism); social Darwinism (or CAS); and the theory of social construction – a humanist approach – which is loosely advocated in this paper.

Note that ‘theory’ has traditionally been used for methods of sentence analysis, but these are today preferably called *syntactic models* (Kiss & Alexiadou 2015) or *grammatical models* (Butler & González-García 2014), to avoid confusion.

As much as the alignment principle can be considered a fact – it is indeed the simplest explanation of harmony (see 6.6) – different hypotheses may be proposed to explain the emergence and stability of alignment. The starting point for speculation especially in today’s American linguistics is to ask in what way the explanation is “biological”: is it the biological reference “literal”, as in Chomskyan linguistics, or “not literal”, as in what Croft (2003) believes is functional typology; and what we have identified as Darwinian typology (section 2).

The way such discussion is expected to go is to state that alignment cannot be genetic or innate because it is not an absolute universal characteristic of language. Therefore, alignment is not taken for a cognitive ‘structure’ but for a cognitive ‘bias’. This we can all agree on: the typological data show that alignment is a linguistic bias, and language is of course ‘cognitive’ in some sense. However, this explanation is already included in the alignment principle (ii), and now we are looking for a hypothesis

(iii). We will be interested to learn in due time what it might be from a social Darwinian point.

Let us remember in the meanwhile that there is more to linguistics than sociobiology and look for hypotheses in the paradigm of nurture. This search will involve the structuralist concepts of formalism and functionalism which are very different from what we have sometimes learned, e.g. from Noonan (1999) and many others in the volume Darnell et al. (1999) *Functionalism and Formalism in Linguistics*. What we would rather consider as formal and functional explanation are discussed in the following two subsections. 6.4.1 and 6.4.2.

6.4.1 A formal explanation

Structuralism or structural linguistics is associated with Saussure ([1916] 1959) who also presents a version of functional sentence analysis or the study of the interaction of form and function. Post-Saussurean structuralists went two different ways, with differences between the so-called functionalists and formalists (Daneš 1987). For Hjelmslev (1943, 14—21), formalism represents the idea that linguistics is a formal discipline, meaning that the appropriate analytical model of empirical data is purely logical or operational, not *real* in itself. This type of formalism, then, is the opposite approach to generative grammar which proposes that the linguistic analysis is a cognitive and biological reality (section 4.1).

But the technical difference between Hjelmslev and the (proper) functionalists depended on the different usage of the word ‘function’ (Daneš *ibid.*). In brief, functional linguistics is based on the standard notion of functional explanation (section 2), while Hjelmslev made much use of the term ‘function’, like many European and American structuralists, without an appeal to what is generally referred to as functional *explanation*. In other words, one may go as far as to say that Hjelmslev’s conception of

language was *non-functionalist*, which also seems true of today's Darwinian linguistics (compare section 2 with Daneš 1987, 4—5).

The difference from the latter is, again, that formalism-structuralism is not a sociobiological approach to language, but a logical-mathematical one. Harder (2005) explains that Hjelmslevian formalism was abandoned by the end of the 1980s when it was considered as having been replaced by generative grammar. Everyone may not see it that way, but it is true that European formalism seems to have vanished. That is why things are now becoming truly interesting, because, looking for the simplest explanatory hypothesis for the alignment principle, what comes up foremost is a type of formal-structural explanation.

To illustrate, we will employ the term *iconicity* with a specific meaning. Precisely, languages in the data of Dryer (1992, 'genera') are *iconic* in respect with Figure 5 below which depicts the alignment of connectives across languages, borrowing English and Latin-based morphemes for illustrative purposes.

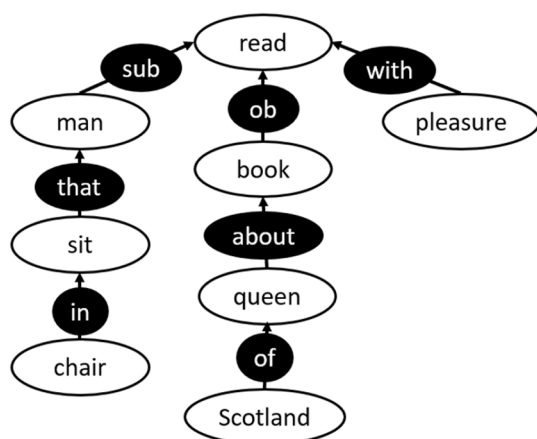


Figure 5. The non-linear matrix of human syntax. VO languages are read top-down, and OV languages bottom-up. Arrows indicate the alignment of the connective. The meaning of the expression in English is ‘*A man who sits in a chair reads a book about the Queen of Scotland with pleasure*’.

The diagram allows the reading of all six combinations of the subject, object and verb. For example, the prototypical SOV order starts from *chair* and reads the three columns bottom-up, one by one, leaving the predicate *read* till last. *Sub* (for agent) and *ob* (for patient) are omitted in many languages. A prototypical SVO reading is top-down from *sub* to *chair*, then from *read* to *Scotland*, *with* and *pleasure*. The place of the adverbial (*with pleasure*) varies in and between languages of the same canonical word order.

We will now use the term iconic for a second link. Figure 6 depicts a logical expression based on set theory, with the same concepts as in the above diagram.

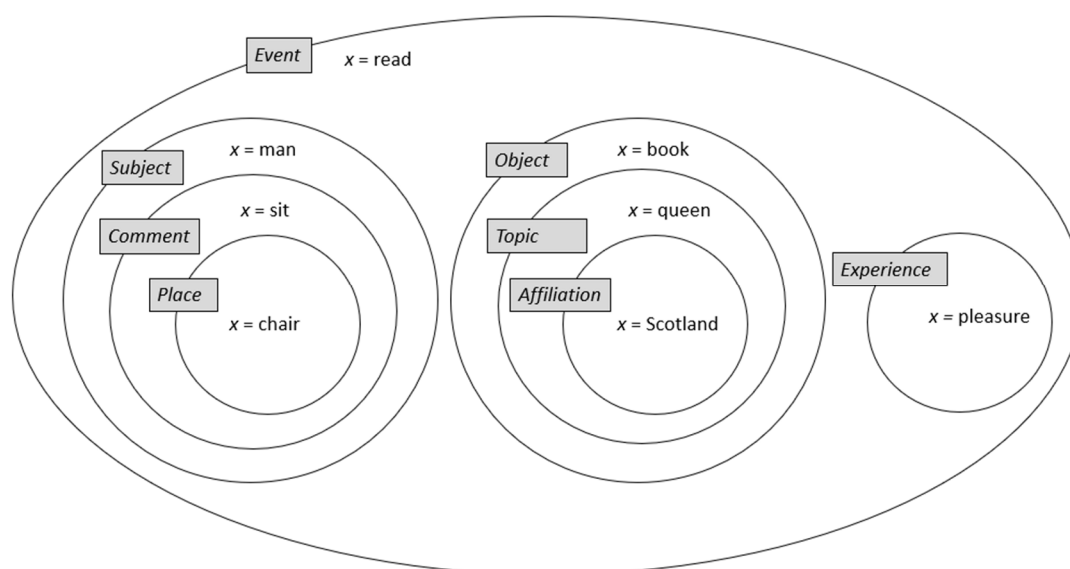


Figure 6. A logical expression based on set theory. Here the text boxes name the sets (circles) which correspond to connectives in Figure 5 (except for the predicate which is unmarked in Figure 4). Elements within the sets, marked here with *x*'s, correspond to content words in Figure 5.

This diagram reads as follows: there is a set named *Event* which entails the subsets *Subject*, *Object* and *Experience*; and the element (*x*) 'read'. In *Subject*, there is the subset *Comment* and the element 'man'. In *Comment*, there is the subset *Place* and the element 'sit'. In the subset *Place* there is the element 'chair', etc.

There are of course other ways to express relations mathematically, e.g. with logical predicates. Thus, the relation between *Event* and *read* is EVENT(read), and similarly: SUBJECT(man), COMMENT(sit), PLACE(chair) etc. Relations between the sets can be expressed as transitive predicates: IN(read,event), IN(man,subject), IN(sit,comment), IN(chair,place). Whichever way is used, the point remains that the prototypical language in Figure 5 is iconic with the logical expression in Figure 6.

The import of the formal-structural hypothesis or explanation is then that language is fundamentally a logical system since the prototypical human language is iconic with set theory. Each human language in its turn is iconic although not identical with the prototypical language. We see this by comparing Figures 5 and 6 with Dryer's (1992) data (see section 6.6).

This type of explanation may seem like a chicken-and-egg situation because logical languages are historically based on human languages (cf. e.g. Peirce [1903] 1998, 178). The above expression is however genuinely logical because moving any element in Figure 5 into a different place in the chain changes the meaning of the expression, as based on Figure 6, and this is what is expected to happen in human languages, too.

Thus, the case of alignment, and thereby harmony, appears to have been solved. The point of formal-structural explanation is that a specific functional explanation is redundant because the observed phenomenon falls from its substantial premises. It may be the least favourite hypothesis in today's linguistics. Sociobiological approaches have been clear that logic cannot play any part in theory of language (see 6.2); and even in the structuralist tradition, formalism seems to have collapsed under pressure from other approaches, including continental functionalism.

The great thing about science is however that it is not a popularity contest, but a matter of proof. The formal-structural hypothesis is extremely powerful because it is the closest thing to the null hypothesis: if it can be demonstrated that the syntax of the languages of the world statistically falls from the logical premises of the semiological system, this will be quite impossible to argue against without real evidence to the contrary. The explanation from the point of social construction is that speech communities favour orderings that are considered as making most sense, and the statistical basis of such judgments is, as we see, objectively logical.

This should by no means be interpreted as meaning that formal-structural explanation means null results that add nothing to our understanding of language. Quite the contrary: there has been interesting philosophical speculations over the subject matter before, involving extra-terrestrials. It was the American structuralist Hockett (1955) who suggested in the *Astounding Science Fiction* magazine that if there are other civilisations in our solar system, they are likely to have similar languages to ours due to logical necessity. This was strongly disputed by Chomsky (1980) who claimed that a “Martian scientist” visiting our planet would have to conclude that Earthlings essentially speak the same idiosyncratic language which only occupies a tiny random strip in the endless space of logical possibilities.

Contrary to such belief, any specifically human factor underlying harmony must be rather meagre because iconicity with set theory explains harmony of the order of object and verb and the place of the adposition with the accuracy of ca. 90 % (cf. Dryer 2013c). But this should by no means lead to the conclusion that (proper) functionalism will have lost its use in linguistics. We shall continue to discuss the plausibility of a functional hypothesis in the next subsection.

6.4.2 Functional explanation

Economy has not been studied extensively in psycholinguistics which means it may be impossible to appeal to research evidence at this point. It is on the other hand clear that formal-structural explanation has its limits, and especially as regards other issues in syntax excluding harmony, functional approaches are likely to find an important place in the whole explanation of word-order variation. While uncertainty due to lack of research limits our knowledge, there are possibilities for some interesting speculations regarding alignment hypothesis that involve a principle of specific instrumental usefulness.

As regards alignment, our own research in section 5 showed a link between economy and the place of the relative pronoun which suggests that harmony is at least in part motivated by the competition of simplicity and disambiguation. At this time there is no further link to other connectives. The alignment principle may however be divided into two questions. One is the fact that connectives of a given language are typically directed in one way rather than the other, as discussed above.

The other is that the place of each connective seems to be quite fixed grammatically, and it is especially this question that may require a functional explanation. The formal-structural hypothesis suggests that connectives are aligned with the word they refer to, but this works optimally in verb-initial (VSO, VOS) and verb-final languages (SOV, OSV). If directing connectives towards either periphery of the sentence is to be avoided, the commonness of SVO is unexpected. In English, for example, adverbials are frequently placed sentence-initially, even if it is not statistically the canonical place of the PP, e.g. '*After the dinner we went to see the show*' with the adposition *after* pointed away from the verb which is the point of reference of the phrase *after the dinner*.

If this is a problem, it could be fixed by simply moving the adposition to the end of the phrase, but the choice between adposition type is typically not available. Based on the economy principle, grammars are parsimonious, favouring one solution for one or many problems. It could be that the choice between placing the adposition before or after its head is fully arbitrary, initially. Once the place of the connective is set, it will likely become immobile, leading to uniformity of phrases.

This could be because having a choice between a separate preposition and postposition for each case would be a grammatical complication. What is more, simply moving the one connective may not add enough clarity to be worth the trouble. To illustrate, a purpose clause could move its connective depending on whether it is placed before or after the main clause. For example, '*I like swimming (because it's healthy)*' cannot be '*(It's healthy because) I like swimming*', with *because* functioning as a postposition, as would be confusing nonetheless.

Moving adpositions might similarly cause as many problems as they solve, e.g., ambiguity in the sentence '*Mary saw a man (in the park)*' could not be avoided by saying, hypothetically, '*Mary saw (the park in) a man*', with *in* functioning as a postposition for making the point that it was the man who was in the park. Instead, the sentence structure offers possibilities for disambiguation as it is. In the last case the adjunct *in the park* can be moved sentence-initially to clarify that *Mary* was in the park. Either way, the result is not perfect, so full clarity can only be reached by adding contextual information, e.g. '*From her window, Mary saw a man who was standing in the park*', and the like.

It can consequently be argued that the economy principle frequently overrides the alignment principle. It seems likely more useful to have fixed patterns with some asymmetry rather than perfect symmetry. Likely reasons for this is that rather rigid

grammatical rules make languages easier to learn, use and understand. There is also a good possibility that the patterning eases perception of language. Connectives are typically short, unstressed particles separating content words. This patterning can be illustrated with the following SOV sentence in Japanese.

- (17) *Watasi-wa tomodati-to resutoran-de supagetii tabe-ta-yo*
 1SG-wa friend-with restaurant-LOC spaghetti eat-PAST-FP
 ‘I ate spaghetti with a friend at a restaurant.’ (Nakagawa 2016, 78)

Here, the postpositions *wa*, *to* and *de* divide all four NPs. If we divide the structure by placing a horizontal line after each postposition, we will see a pattern of separating content words from one other:

I-NOM| friend-COM| restaurant-LOC| spaghetti eat-PAST-FP

In English, the verb separates the subject and the object, and as for the rest, we see a pattern of separating content words with prepositions.

I ate spaghetti |with a friend |at a restaurant.

It is most common that connectives are located around the centre of the sentence rather than in the periphery. This phenomenon may be a by-product of the verb-directed alignment, but it would seem to work particularly well for differentiating between sentence parts and sentences, with the clausal connective added. It is an alternative explanation of the data, that the canonical word order crosslinguistically avoids placing the connective to the periphery where they are not useful for creating word patterns that indicate phrasal and clausal boundaries. This explanation is economical in the sense that, instead of adding specific syntactic particles for the job, words carrying semantic tasks which are already present are arranged in a way that works for disambiguation and ease of perception without added complexity.

Research is needed in syntax, typology and psycholinguistics to shed light on functional motivation of harmony; with the side note that this will unlikely involve CAS or other models of natural selection, for reasons which are further discussed in the following subsections.

6.5 Discussing psycholinguistic evidence for MiD

It is a challenge for this article to make full justice to evidence supporting Hawkins's theory which is based on decades of research by numerous scholars. It takes three books to account for all of it (Hawkins 1994, 2004, 2014). Due to space limitations we will only discuss what we believe are the most central pieces of evidence for MiD while being prepared to answer further questions by request.

Hawkins claims to be supported by a significant body of psycholinguistic research. Schwarz-Friesel (2012) is however sceptical of the type of 'cognitive linguistics' done within the generative movement; and as we examine Hawkins's psycholinguistic evidence more closely, his claim of being supported by "numerous experiments" (2004, 173—174) turns out to be a rather approximate reference to surveys by various generativists asking people to for grammaticality judgments (cf. Kluender & Kutas 1993). This depends on generative grammar being taken as psycholinguistics or cognitive linguistics as it is by its advocates, while other people might consider it as just the study of grammar, without the hidden equation of 'grammar' and 'the mind'. To use an unambiguous term, the research in question does not provide any *neurolinguistic* evidence, so Hawkins's claim boils down to stating that generative grammar cannot be complete nonsense because so many people are involved in doing it.

More interestingly, Hawkins also cites genuine neurolinguistic research, notably Kluender & Kutas (1993) who combined the generative method with ERP brain

response measurement. Now, this is a type of study which has true potential to validate the VP (NP) stipulation. The claim is that, in English relative clauses with the OSV order, e.g. ‘I don’t know *what Mary saw* ____’, generative grammarians argue that there is an invisible cognitive structure that places the object *what* ‘into’ the VP. That is, into the ‘gap’ (____) after the verb *saw*. With ERPs, it should be possible to bring evidence for the reality of such a structure by demonstrating an otherwise inexplicable ease of processing in cases where the object and the verb are placed distantly apart, suggesting a kind of cognitive wormhole from the one to the other.

What Kluender & Kutas (ibid.) actually found was that this does not occur at all: the further the words are apart, the more effort it takes to link them, as is fully expected without assuming so-called Universal Grammar. That is, as actual brain studies became available, it was soon shown that the claim of VP (NP) as a cognitive reality is false, thus effectively nullifying generative grammar for once and for all. The conclusion of neurologists is that syntactic structures are not innate. Today the researchers remain more “agnostic” about explanatory frameworks (Barkley, Kluender & Kutas 2015).

But Hawkins makes two unexpected conclusions based on the 1993 study. The first one is that, since grammar is proved not to be innate, he (1994, xi) “reduces” the role of Universal Grammar. We understand this as suggesting that it is *precisely* the VP (NP) structure that is taken to be innate, rather than endorsing a full parameters framework. But this cannot be correct. Second, Hawkins (2014) additionally suggests that word-order variation is based on language being a population of organisms struggling for life by citing Haspelmath (1999) and Kirby (1999), thus seemingly eliminating any possibility that language is constructed by the speech community as would have been the obvious conclusion once the alternative hypothesis was rejected.

We summarise this subsection by stating that the claim of MiD being supported by serious research in psycholinguistics is unfounded. Hawkins (2014, through Hawkins 2004) cites notably one study which fails to agree with his theory. Many more could undoubtedly be found in the same vein. We will move on to examine linguistic evidence for MiD in the next subsection.

6.6 A comparison of MiD and the alignment principle

We argue that our alignment principle makes equally good predictions as Hawkins's MiD although not exactly the same ones, while being simpler. This means MiD is overruled by the alignment principle per Occam's razor.

The difference between predictions made by the two lies in the concept of *heaviness*. While it can be argued that MiD is a specific type of a heaviness explanation, the principle of alignment does not assume heaviness at all: that is, phrasal length within a sentence steadily increasing or decreasing from left to right (cf. LIPOC: Dik 1989). This is not to say that heaviness plays no role in explaining word order as a whole but, more specifically, that the alignment principle seems to be a sufficient explanation of *harmony* in typological data. The difference between the two principles is illustrated by the following comparison of sentence structure predictions.

| | |
|------------|------------------------------------------------------------------------------------------------|
| MiD: | Verb, (2-word phrase), (3-word phrase), (4-word phrase) ... |
| alignment: | Verb (Connective, 4-word phrase), (Connective, 2-word phrase), (Connective, 3-word phrase) ... |

In other words, phrasal complexity in itself is not a predictive factor for the alignment principle. Due to limitations and uncertainties in the crosslinguistic data, it may currently not be possible to know exactly which principle makes precisely the most accurate predictions when each and every language is properly examined. MiD is argued to be the best model available until now (Song 2010). But alignment appears to

be as good in explaining OV/VO harmonies making the correct predictions for Dryer (1992) table 1 adposition type; table 5 Order of noun and genitive; table 7 Order of adjective and standard in comparative construction; table 9 Order of verb and adpositional phrase; table 12 Order of copula and predicate; table 14 Order of ‘want’ + verb; table 28 Order of content verb and auxiliary verb; table 30 Order of question particle and sentence; table 32 Order of adverbial subordinator and clause; and table 35 Order of plural word and noun; all assuming a postpositional connective in OV languages and a prepositional connective in VO languages.

Alignment also makes equally good predictions as MiD for table 2 Order of noun and relative clause, including correct predictions for Eurasian SOV data. Our preliminary conjecture as for why it does not work with non-Eurasian SOV languages would be that the relativiser may not be overtly marked in many of these because relativisation could also be expressed by changing word order, as in the following example from Somali (Cushitic; Saeed 1999).

- (18) *buugágga nimanku keenàan*
 books-the men-the¹ bring
 ‘the books which the men bring’ (Saeed 1999, 213)

In this case, as is also possible in English (*‘the books the men bring’*), the canonical SO order is changed to OS to mark relativisation, necessarily resulting in the NRel pattern which is not predicted by alignment iff a relative particle is used, which is not the case here. Instead, a different principle in the structural-functional tradition might be available for this kind of phenomenon, namely markedness or marking *the other* (Andersen 1989). It is an example of a different principle that may be needed for the

¹ *nimanku* is overtly marked as a subject (Saeed *ibid.*).

whole explanation of word order. Hypothetically, because the subject, included in both intransitive and transitive predicates, is more frequent in discourse than the object, it supports the subject coming foremost, so the object may be marked as ‘the other’ by word order.

The SO order thus becomes *the norm* for indicative sentences. This order can again be changed to mark *the other* case which depending on the language could entail things such as topicalisation or relativisation. However, in order not to get fully side-tracked, one will have to accept at this point that we do not have sufficiently detailed information on what the differences between relativisation in SOV languages of Eurasia and other parts of the world depend on, and merely state that alignment makes as good a prediction as MiD for OV/VO + RelN/NRel.

Predictive differences between MiD and alignment are observed as regards table 17 Order of noun and adjective; and table 19 Order of demonstrative and noun. Assuming no connectives marking adjectives and demonstratives, the alignment principle makes no predictions for these generalisations, so MiD (Hawkins 2004) has clearly more explanatory power. Whether this is a good thing or bad may be a matter of taste because there is no clear correlation in the data, either.

The question of simplicity should be sufficiently straightforward because EIC/MiD is notoriously complicated (cf. Croft 2003, 74; Figure 3 in this article; and especially Hawkins 1994, 59). As regards Occam’s razor, the main point is however that alignment avoids making unsubstantiated biological claims that MiD makes, including the VP (NP) structure, “reduced” UG, CAS, and claims of processing efficiency based on these. As stated in the previous section, there does not appear to be any rigorous evidence for any of it. This is in contrast especially with our formalist

hypothesis (section 6.4.1) which follows from the null hypothesis so closely it should stand due to current lack of clear indication to the contrary.

Let us continue to compare MiD with alignment bearing in mind it would be vital for MiD to make superior predictions because it is more complicated. Hawkins's starting point is in English where he notes that option (b) is statistically preferred over (a) in each of the following cases:

- (19) a. **That** their time should not be wasted is important (for us all).
 b. It is important (for us all) **that** their time should not be wasted.
- (20) a. **To** dine in a good restaurant is delightful (for us all),
 b. It is delightful (for us all) **to** dine in a good restaurant.
 (Hawkins 1994, 190; bold type added.)

These examples make it possible for us to suggest that the preference actually concerns the placement of the connective (highlighted here) which is towards the phrase it refers to. Placing the connective *that* (19b) to the boundary of the two clauses may also be argued to help the hearer link each phrase with the right verb.

In the typological data of Schmidtke-Bode (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, as in Hawkins's (upper) example above. We take this as supporting the alignment principle. Our own research (section 5.3) additionally found that moving the connective from this place may result in increased syntactic ambiguity.

Regarding languages that use a subordinating affix rather than a subordinating conjunction, most 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).

Thus, our explanation for the preference of (b) over (a) in case (19) and (20) above is that connectives are more often located at the clausal intersection, directed towards the phrase they refer to. Our notion is also supported by Hawkins’s material, see table 3 below. *Order A* languages, like English, place the preposition to the clausal intersection rather than the right periphery, while *Order B* languages, like Japanese, place the postposition to the clausal intersection rather than the left periphery. These orders are deemed as *efficient* by Hawkins. Looking for less efficient languages, we find languages like Finnish in *Order C* and Chinese in *Order D* (cf. Hawkins 1994, 281). These are, then, implied to have “inefficient” grammars, causing their speakers “processing difficulty”.

| Order | Genera | Percentage |
|--------------|--------|------------|
| A [V [P NP]] | 161 | 41.4% |
| B [[NP P] V] | 204 | 52.4% |
| C [V [NP P]] | 18 | 4.6% |
| D [[P NP] V] | 6 | 1.5% |
| Total | 389 | |

Table 3. Grammaticalised $_{VP}\{V\}_{PP}\{P\}_{NP}\}$ orders by language family (Hawkins 1994, 255—257, based on his personal communication with Dryer); crosslinguistic positioning of a verb and object within VP, and the order of adposition and noun within the prepositional phrase. In the data, DO and PP are assumed to occupy the same position relative to the verb. In Dryer (1992), most VO languages (98.3%) fall within Order A while most OV languages (87.5%) fall within Order B.

We make no such claim, but instead assume that the dominant patterns are such that they win consensus within the speech community more easily as ‘making sense’ in a way (see section 6.4.1). As regards the exceptions, before implying that speakers of certain languages actually struggle with sentence processing, one should look into what may be causing the asymmetry.

In the case of Finnish, for example, it should be noted that there is some uncertainty in interpreting the data. Matsumura (1982) suggests that Finnish has a dual system of relativisation. One way is typical for VO languages and the other typical of OV languages, likely reflecting a relatively recent shift from SOV to SVO. As it stands, the postposition is directed away from the verb, but towards the nominal head of the participial phrase which is interpreted as corresponding to the RelN pattern. Compare the following examples of Turkish (21, SOV) and Finnish (22, SVO with postpositions).

(21)

- a. *[kitab-ı al-an] öğrenci*
 book-ACC buy-PTCP student
 ‘the student who bought the book’
- b. *[öğrenci-nin al-dığ-ı] kitap*
 student-GEN buy-NMLZ-3SG book
 ‘the book which the student bought’ (Comrie & Kuteva 2013)

These are translated into Finnish as follows:

(22)

- a. *kirjan ostanut opiskelija*
 book-ACC buy-PTCP student
 ‘the student who bought the book’
- b. *opiskelijan ostama kirja*
 student-GEN buy-PTCP book
 ‘the book which the student bought’ (pk)

In both languages, the suffixed postpositions are directed towards the head noun as is typical of SOV languages. There does currently not appear to be any research available

to answer our question but, admittedly, if it will turn out that asymmetrical ordering of connectives causes no disadvantage for perceiving or processing Finnish and other western Uralic languages, it would seem to support the formal hypothesis (4.6.1) rather than the functional one (4.6.2).

As for Mandarin Chinese, the question is whether the connectives in question are genuine prepositions like in English since these have also been likened to verbs (Li & Thompson 1974b).

Regarding relativisation, Hawkins's (1994, 272) analysis of a large number of languages seems to suggest that MiD makes slightly more accurate predictions than our principle. The obvious problem here is that a great majority of the languages fall into *Structure ABCD* (Table 4) as predicted by MiD and alignment alike, and the rest concerns a very limited number of data. Thus the quick and easy way to refute Hawkins's advantage is simply to point to the lack significance of his findings.

| Structure | EIC efficiency | Attested language numbers |
|-----------------|----------------|------------------------------------------|
| A [N Adj [C S]] | 100% | Extensive, e.g. Romance, Arabic, Ewe |
| B [Adj N [C S]] | 100% | Extensive, e.g. Germanic, Greek, Finnish |
| C [[S C] N Adj] | 100% | Extensive, e.g. Basque, Burmese |
| D [[S C] Adj N] | 100% | Extensive, e.g. Telugu, Turkish, Lahu |
| E [[S C] Adj N] | 83% | Attested, e.g. Lushei, Dyrbal |
| F [Adj N [S C]] | 83% | Attested, e.g. Yaqui, Hurrian |
| G [N [C S] Adj] | 63% | None |
| H [Adj [C S] N] | 63% | None |
| I [N [S C] Adj] | 63% | None |
| J [Adj [S C] N] | 63% | Marked variant in Lahu, Chinese |
| K [[C S] N Adj] | 38% | None |
| L [[C S] Adj N] | 38% | None |

Table 4. The twelve logically possible ways of ordering a relative clause ([C S] or [S C] within an NP. ‘EIC efficiency’ is titled *L-to-R IC-to-word ratio (aggregates)* in the original (Hawkins 1994, 272).

However, we would indeed like to elaborate why IC-to-word ratios may appear to be a potentially useful concept for analysing relativisation patterns although, in reality, these are most likely quite irrelevant. In Table 4, we see that most languages fall into structures with 100 % efficiency according to EIC/MiD. The rest can be divided into three groups: *Structures EF* with 83 % efficiency and some examples; *Structures GHIJ* with 63 % efficiency and some non-dominant examples in *J* only; and *Structures KL* with 38 % and no examples found at all.

To begin with, a persistent question throughout the work on MiD relates to the concept of centre-embedding, as discussed in section 6.2. It has been noted from the economy perspective that centre-embedding structures are more complex than left or right-embedding ones. When the subclause is centre-embedded, e.g. *’*Big, that stands on the hill, house*’, the sentence structure is $A_1 (B) A_2$ with the main clause *A* divided into two parts. This is more complex than $A [B]$ or $[B] A$. Hawkins has successfully demonstrated that, while centre-embedding is in many cases grammatical, it is statistically less frequent. This is however expected in any case, so the rarity of the purported 63 % efficient group (GHIJ) is more simply explained by having more complexity and load on memory without assuming the EIC principle.

Regarding *Structure J* with some non-dominant examples, on the other hand, it was suggested above that Finnish, for example, has a dual system of relativisation. This question reveals some inconsistency in the data or its interpretation because prenominal (RelN) relativisation is considered participial if not downright adjectival in Finnish grammar (Hakulinen & Karlsson 1979, 373; Hakulinen et al. 2004). But if we accept

that the participial suffix corresponds to a relativiser, we basically find *Structure J* in the English phrase ‘(a) *calm peace-lov-ing person*’. This only depends on what is meant with relative clause and relativiser; but it does not make the fact that a corresponding structure is attested in some of the world’s languages (excluding English) as anyhow surprising. To sum up, we neither consider the absence of *Structure G, H* and *I* nor the limited presence of *J* in the cited data as necessarily supporting MiD.

This leaves us with the question why Structures EF (83 % efficiency) are attested while Structures KL (38 % efficiency) are not. Starting from EF, the languages cited by Hawkins are SOV languages whereby we are looking at large overall data. As Dryer suggests,

relative clauses resemble adjectives in terms of their relationship to the nouns they modify, as reflected by the fact that relative clauses are often referred to as *adjective clauses*. (Dryer 2013b; italics in the original.)

But the place of the adjective does not have a strong link with the order of object and verb (Dryer 2013e). Therefore it is not highly surprising if some languages organise their participial or adjectival relative phrases in any which order. We consider Hawkins’s findings, that EF are attested in the cited data, as expected without MiD.

This only leaves us needing to understand why Structures KL are not attested at all. They mirror Structures EF, so they should be expected to be equally rare or common in the data of *VO languages*, unless one assumes MiD. This gives us an opportunity to take a little time to demonstrate the advantages of a systemic approach to typology which means investigating how structures work in relation to one another.

L implies a structure corresponding to ‘*who loves peace young person*’ for the English sentence ‘a person who loves peace’. To put this into a context, consider the sentence ‘*John is a young man who loves peace*’ which has the structure SVO[SVO].

The corresponding sentence along the *L* pattern would be ‘*John is who loves peace a young man*’ with the structure SV[SVO]O. It is necessarily centre-embedding and therefore more complex, anticipating a lower frequency by default.

Of course, all languages do not have relative pronouns, and some have relative particles. In a language that has relative particles, *Structure L* is rendered as ‘*REL loves peace young man*’ with the VOS order; or ‘*REL peace loves young man*’ with OVS. Now, it might simply be the case that we are short of material because these are both rare canonical word orders, and only a minority of languages relativise with a relative particle anyway, according to Dryer (2013d).

All in all, it is impossible to conclude that the absence of KL in the data must be anyhow related to the purported IC-to-word ratio. Instead, what we do see in the prevalent Structures ABCD is that the relativiser typically functions as a connective physically linking the relative clause with its antecedent, and we have discussed a formal and a functional explanation for the phenomenon. Any of these will give a simpler, sufficient explanation for Hawkins’s observations, making MiD redundant.

Of course, we could be totally mistaken; maybe the findings in Table 4 are significant after all, as strongly argued by Song (2010) who believes it is precisely these small findings that makes MiD the cutting edge of typology. His conclusion is that

Most word order studies have of late revealed a strong inclination towards processing as a major avenue of explanation. Future word order typology—along the lines of Hawkins (1994, 2004)—will increasingly be processing-based [...] This truly is the hallmark of an evolving scientific endeavour. (Song 2010, 278—279.)

But one should not jump into such conclusions before asking what it is that Hawkins has actually proved. MiD is based on the idea that the SVO verb phrase VP [NP PP CP] is mirrored by SOV languages with [CP PP NP] VP. Hawkins does not substantiate the

ordering, but we suspect the PP should actually follow the NP in many SOV languages. This would require the PP to be shorter than the object in order for MiD to work.

According to MiD, information structure is irrelevant, but we believe that if the CP is a relative clause, it should be placed next to its antecedent whereby it is either the CP or the NP/PP that will have to move, depending on the contextual semantics. If correct, such issues would suffice to sink the model while the alignment principle remains unaffected.

The starting point of EIC was to focus on the VP, ignoring the subject, but Hawkins cannot resist temptation to make a full explanatory model for the order of subject, object and verb. This is supposed to work based on his findings that the subject in SOV languages is actually longer than the object (Hawkins 1994, 334). Even though we have just learned that the new-versus-given ordering does not play a role in word order, it is now claimed that

SOV languages should exhibit more new-before-given scores, correlating with long before short. Hence, subjects should be more given and definite in SVO languages, less so in SOV languages. (Hawkins 1994, 334.)

The subject—object word ratio is set to 3:2 in order for EIC to work, but Hawkins does not show us how such discourse would actually look like. To take an SVO sentence, e.g. ‘*Mary saw a man from her window*’, one would expect the subject of the next sentence rather likely to be *she* (Mary), *he* (the man) or even *it* (the window). But SOV languages, according to Hawkins’s explanation, will instead pack a lot of new information *to the subject*, apparently leaving the above given candidates to the object. The reader is left wondering how this might actually work.

The biggest question mark hangs over Hawkins’s concept of processing efficiency itself. It is claimed that parse trees that mix left and right branching are

avoided because they cause people processing difficulty. Languages however have exceptions, and English for instance has some postpositions, too. While both MiD and alignment agree that this causes some irregularity in the patterning, MiD goes further by claiming that such structures are more difficult to parse. The claim is then, that English speakers have no problem parsing ‘*We will visit London in three weeks*’, but will have difficulty with ‘*We visited London three weeks ago.*’ Is this true?

If MiD truly is interesting, it should come back when the necessary evidence is there. Until then, the claim of its relevance for language sciences does not seem justified. Based on our inquiry, the sole purpose of MiD appears to be in joining all sociobiological forces together in order to survive in the academia in spite of a lack of evidential basis.

7 Conclusion

We scrutinised Hawkins’s explanation of word-order harmonies in crosslinguistic data, namely the EIC principle (Early Immediate Constituents) or MiD (for Minimize Domains, from Hawkins 2004 onward). Our first concern was that, unlike Hawkins claims, his explanation mode does not seem to be functional although it does represent an adaptive or adaptational explanation of some kind. We linked this issue with an apparent flaw in Darwinian evolutionary theory where the same explanation mode works either for nature or culture, but not both. We requested a clarification concerning Hawkins’s claim of advocating functional explanation, with reference to standard literature on the subject.

We linked MiD with two sociobiological theories of language. One is generative innatism which suggests that the VP (NP) structure is genetically encoded. We found no evidence for it, but observed that MiD, like generative grammar, depends on the

structure being a representation of cognitive reality of some kind. Based on the literature, we concluded that such an idea is not supported by serious research.

The second sociobiological theory is CAS or language as a Complex Adaptive System which we further linked with memetics. CAS is a theory of the natural selection of so-called cultural replicators, similar to what Dawkins calls “mind-viruses”. We could not identify any research validating language as a CAS and regretted the absence of the concept of null hypothesis in sociobiological linguistics. As the innatism claim was proven false, it should have been acknowledged that the null hypothesis, that the speech community creates its language, *is confirmed*. Instead, Hawkins patched the failed innatism hypothesis with another questionable claim, CAS, suggesting that language is a quasi-biological organism or a population of organisms that adapts to the already-refuted VP (NP) structure.

Nonetheless, MiD was successful as an explanation of word-order variation because it provided a common framework for English (VO) and Japanese (OV). Generalisations made on this basis were subsequently extended to large crosslinguistic data. We however identified asymmetries in the way VO and OV languages were analysed, suspecting that Hawkins’s findings could depend on a different cause.

To properly address Hawkins’s question why functional grammar is ‘functional’, we carried out an analysis of the information structure of Czech, English, Finnish and Swedish text, and an experiment on the last three. Our findings linked functional sentence perspective and the functional economy principle with the place of the relative pronoun whose primary function was observed to be connective. We then studied connective architecture crosslinguistically uncovering a pattern which provided a simpler explanation for observations made by Hawkins and his colleagues.

The real cause of word-order harmonies depends on the necessarily linear structure of language which demands an initially arbitrary choice of placing a linguistic element either to the left or to the right of its so-called head. In the language system, this arbitrary stipulation is systematised to make it possible for the speaker and hearer to associate the elements in a consistent way. It then leads to further consequences relating to word order. There does not seem to be a perfect solution to the problem of conveying nonlinear meaning in linear form, but exceptions aside, the preference is shown to be very strong in crosslinguistic data of canonical word order.

It is especially connectives that are statistically placed to the boundary of the phrase and directed towards what the phrase containing them refers to, whether leftward in VO languages or rightward in OV languages. We named this the principle of alignment and found that the simplest explanation for it is formal-structural: the prototypical language arising from the crosslinguistic data is iconic with set theory. Therefore, it seems human grammars are fundamentally logical systems.

A second possibility, a functional explanation was discussed. Uniformity of patterning was suggested as a simplified way to ease perception of language. Functional explanation was deemed as having genuine potential for the basis of future research in syntax, typology, and cognitive linguistics.

We compared MiD and our own alignment principle concluding that MiD is redundant because alignment makes equally good predictions with fewer assumptions. We found that functional sentence perspective and the economy principle remain integral for explaining how language works, endorsing a systemic view of language. The concept of rational explanation is important to understanding that languages are the functional (i.e. instrumental) creation of their speech communities even though the social construction of language may occur largely unconsciously and unintentionally.

As any research within theoretical linguistics, our findings essentially support the null hypothesis that language is the intelligent design of the speech community. While there remains the possibility that biological matters are required to explain a small percentage of linguistic structures, linguists are yet to learn what such structures might be and where the necessary evidence will be available to be seen.

References

- Abels, Klaus. 2015. Word order. In T. Kiss, and Alexiadou, A. (eds.) *Syntax-Theory and Analysis* (Vol. 1). Walter de Gruyter GmbH, and Co KG.
- Allen, Colin and Jakob Neal. 2019. Teleological Notions in Biology. In E. N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy*.
<<https://plato.stanford.edu/archives/spr2019/entries/teleology-biology/>>
- Andersen, Henning. 1989. Markedness theory—the first 150 years. Tomic, Olga (ed.) *Markedness in synchrony and diachrony*, De Gruyter.
- Andersen, Henning. 2006. Synchrony, diachrony, and evolution. Thomsen, O. N. (ed.) *Competing models of linguistic change: Evolution and beyond*. John Benjamins Publishing.
- Aronoff, Mark. 2017. Darwinism tested by the science of language. Bower, C., L. Horn, and R. Zanuttini (eds.) *On Looking Into Words (and Beyond)*. Berlin: Language Science Press.
- Barkley, C., R. Kluender, and M. Kutas. 2015. Referential processing in the human brain: An Event-Related Potential (ERP) study. *Brain research*, 1629, 143–159.
- Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Croft, W., ..., and Schoenemann, T. (2009). Language is a complex adaptive system: Position paper. *Language Learning*, 59 suppl. 1, 1–26.
- Bloomfield, Leonard. 1933. *Language*. New York: Holt.
- Boesch, Christophe, and Michael Tomasello. 1998. Chimpanzee and human cultures. *Current anthropology*, 39(5), 591–614.
- Butler, Christopher. S. 2003. *Structure and Function - A Guide to Three Major Structural-Functional Theories, Part 2: From clause to discourse and beyond*. Amsterdam/Philadelphia: John Benjamins.

- Butler, Christopher. S. 2008. Cognitive adequacy in structural-functional theories of language. *Language Sciences*, 30(1), 1–30.
- Butler, Christopher. S., and Francisco Gonzálvez-García. 2014. *Exploring Functional-Cognitive Space*. Amsterdam/Philadelphia: John Benjamins.
- Butt, Mirjam, and Tracy Holloway King. 2015 Lexical-Functional Grammar. In Tibor Kiss, and Artemis Alexiadou (eds.) *Syntax – Theory and Analysis: An International Handbook*. Berlin/Boston: De Gruyter Mouton.
- Bybee, Joan. 1999. Usage-based phonology. In Darnell et al. (eds.) *Functionalism and Formalism in Linguistics*, Vol. 1, 211–242.
- Bybee, Joan. 2014. Cognitive processes in grammaticalization. In M. Tomasello (ed.) *The New Psychology of Language*, 159–182. Hove: Psychology Press.
- Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. The Hague: Mouton.
- Chomsky, Noam. 1980. *Rules and representations*. New York: Columbia University Press.
- Chomsky, Noam. 2002. *On nature and language*. Cambridge University Press.
- Chomsky, Noam. 2009. Opening remarks. In M. Piattelli-Palmarini, J. Uriagereka, and P. Salaburu (eds.) *Of minds and language: A dialogue with Noam Chomsky in the Basque Country*. Oxford: OUP, 13–43.
- Comrie, Bernard, and Tania Kuteva. 2013. Relativization on Subjects. In: Dryer, Matthew S., and Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/122>, Accessed on 2019-08-31.)
- Couch, Mark B. 2011. Causal role theories of functional explanation. In James Fieser, and Bradley Dowden (eds.), *Internet Encyclopedia of Philosophy*. <<https://www.iep.utm.edu/explanat/>>
- Croft, William, and D. Alan Cruse. 2004. *Cognitive linguistics*. Cambridge University Press.
- Croft, William. 1993. Functional-typological theory in its historical and intellectual context. *STUF – Language Typology and Universals*, 46(1), 15–26.
- Croft, William. 1995. Autonomy and functionalist linguistics. *Language*, 71(3), 490–532.

- Croft, William. 2000. *Explaining Language Change: An Evolutionary Approach*. New York: Pearson.
- Croft, William. 2001. Grammar: functional approaches. *International Encyclopedia of the Social and Behavioral Sciences*. 1st Edition.
- Croft, William. 2003. *Typology and Universals*. Cambridge University Press.
- Croft, William. 2006. The relevance of an evolutionary model to historical linguistics. In O. N. Thomsen (ed.) *Competing Models of Linguistic Change: Evolution and Beyond*. Amsterdam/Philadelphia: John Benjamins.
- Croft, William. 2015. Grammar: functional approaches. In Neil J. Smelser and Paul B. Baltes (eds) *International Encyclopedia of the Social and Behavioral Sciences*. 2nd Edition, vol. 9.
- Culbertson, Jennifer, and Simon Kirby. 2016. Simplicity and specificity in language: Domain-general biases have domain-specific effects. *Frontiers in Psychology*, 6, 1964.
- Dahl, Östen. 2004. *The Growth and Maintenance of Linguistic Complexity*. Amsterdam/Philadelphia: John Benjamins.
- Daneš, F. 1987. On Prague School functionalism in linguistics. In R. Dirven, and V. Fried (eds.) *Functionalism in linguistics*. Amsterdam/Philadelphia: John Benjamins.
- Darnell, M., E. A. Moravcsik, M. Noonan, F. J. Newmeyer, and K. Wheatley (eds.). 1999. *Functionalism and Formalism in Linguistics: Volume I: General papers*. Amsterdam/Philadelphia: John Benjamins.
- Darwin, Charles. 1859. *On the Origin of Species by Means of Natural Selection, or Preservation of Favoured Races in the Struggle for Life*. London: John Murray.
- Darwin, Charles. 1871. *The Descent of Man and Selection in Relation to Sex*. 1981 reprint. Princeton: University Press.
- Dawkins, R. 1976. *The Selfish Gene*. Oxford: OUP.
- De Bot, K. 2015. *A History of Applied Linguistics: From 1980 to the Present*. London: Routledge.
- Dik, Simon. 1989. *The Theory of Functional Grammar, Part I: The Structure of the Clause*. Dordrecht: Foris.
- Dryer, M. S. 1992. The Greenbergian word order correlations. *Language*, 68(1), 81–138.

- Dryer, Matthew S. 2013a. Expression of Pronominal Subjects. In: Dryer, Matthew S., and Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/101>, Accessed on 2019-08-29.)
- Dryer, Matthew S. 2013b. Order of Subject, Object and Verb. In: Dryer, Matthew S., and 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 2019-09-04.)
- Dryer, Matthew S. 2013c. Relationship between the Order of Object and Verb and the Order of Adposition and Noun Phrase. In: Dryer, Matthew S., and Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/95>, Accessed on 2019-08-29.)
- Dryer, Matthew S. 2013d. Relationship between the Order of Object and Verb and the Order of Relative Clause and Noun. In: Dryer, Matthew S., and Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/96>, Accessed on 2019-08-29.)
- Dryer, Matthew S. 2013e. Relationship between the Order of Object and Verb and the Order of Adjective and Noun. In: Dryer, Matthew S., and Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/97>, Accessed on 2019-08-31.)
- EC 2014. *Organic Farming - A Guide on Support Opportunities for Organic Producers in Europe*. Luxembourg: Publication Office of the European Union.
<<https://publications.europa.eu/en/publication-detail/-/publication/3b428e25-aa3b-4e9d-a7d4-9c640d6f5a30/language-en/format-PDF>>
- Egré, Paul. 2015. Explanation in linguistics. *Philosophy Compass*, 10(7), 451–462.
- Encyclopaedia Britannica 2015. Neurolinguistics.
<<https://www.britannica.com/science/neurolinguistics>>
- Everaert, M. B., M. A. Huybregts, N. Chomsky, R. C. Berwick, and J. J. Bolhuis. 2015. Structures, not strings: linguistics as part of the cognitive sciences. *Trends in Cognitive Sciences*, 19(12), 729–743.

- Eysenck, Michael. W. 2012. *Simply psychology*. Hove: Psychology Press.
- Fillmore, Charles. J., and Collin Baker. 2015. A frames approach to semantic analysis. In B. Heine, and H. Narrog (eds.) *The Oxford handbook of Linguistic Analysis*. Oxford: OUP.
- Fillmore, Charles. J. 2013. Berkeley Construction Grammar. In Hoffmann, T., and G. Trousdale (eds.) *The Oxford Handbook of Construction Grammar*. Oxford: OUP.
- Firbas, Jan. 1966. On defining the theme in functional sentence perspective. *Travaux linguistiques de Prague*, 2.
- Firbas, Jan. 1971. On the concept of communicative dynamism in the theory of functional sentence perspective. *Linguistica* A 19, Brno.
- Firbas, Jan. 1974. Some aspects of the Czechoslovak approach to problems of functional sentence perspective. In F. Daneš (ed.) *Papers on Functional Sentence Perspective, Janua Linguarum*, sm 147. The Hague: Mouton.
- Firbas, Jan. 1985. Thoughts on functional sentence perspective, intonation and emotiveness. *Brno studies in English*, 16(1), K7.
- Firbas, Jan. 1987. On the delimitation of the theme in functional sentence perspective. In R. Dirven, and V. Fried (eds.) *Functionalism in Linguistics*, 137–156. Amsterdam/Philadelphia: John Benjamins.
- François, Jacques. 2018. The stance of Systemic Functional Linguistics amongst functional(ist) theories of language and its 'systemic' purpose. In A. Sellami-Baklouti, and L. Fontaine (eds.) *Perspectives from Systemic Functional Linguistics*, 1–5. London: Routledge.
- Frank, Roslyn. M. 2008. The Language-organism-species analogy: A complex adaptive systems approach to shifting perspectives on "language". In Roslyn M. Frank, et al. (eds.) *Sociocultural Situatedness*, Vol. 2. Berlin/Boston: De Gruyter Mouton.
- Frey, Werner. 2015. Word order. In Tibor Kiss, and Artemis Alexiadou (eds.) *Syntax – Theory and Analysis: An International Handbook*. Berlin/Boston: De Gruyter Mouton.
- Fried, Mirjam. 2015. Construction Grammar. In Tibor Kiss, and Artemis Alexiadou (eds.) *Syntax – Theory and Analysis: An International Handbook*. Berlin/Boston: De Gruyter Mouton.

- Gell-Mann, Murray. 1994. Complex adaptive systems. In G. Cowan, D. Pines, and D. Meltzer (eds.) *SFI Studies in the Sciences of Complexity, Proc. Vol. XIX*. Boston: Addison-Wesley.
- Givón, Talmy. 1979. *On understanding grammar*. New York: Academic Press.
- Givón, Talmy. 1988. The pragmatics of word-order: predictability, importance and attention. In M. Hammond, E. A. Moravcsik, and J. Wirth (eds.) *Studies in Syntactic Typology*. Amsterdam/Philadelphia: John Benjamins.
- Givón, Talmy. 1995. *Functionalism and Grammar*. Amsterdam/Philadelphia: John Benjamins.
- Givón, Talmy. 2002. *Bio-linguistics: The Santa Barbara Lectures*. Amsterdam/Philadelphia: John Benjamins.
- Goldberg, Adele E. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford: OUP.
- Greenberg, Joseph H. 1979. Rethinking linguistics diachronically. *Language* 55(2), 275–290.
- Gundel, Jeanette K. 1988. Universals of topic-comment structure. *Studies in Syntactic Typology*, 17(1), 209–239.
- Haberland, Hartmut, and Lars Heltoft. 1992. Universals, explanations and pragmatics. In J. van der Auwera & M. Kefer (eds.) *Meaning and Grammar: Cross-Linguistic Perspectives*, 17–27. Berlin/Boston: De Gruyter Mouton.
- Haider, Hubert. 2015. "Intelligent design" of grammars – a result of cognitive evolution. In Adli, A., García, M. G., and Kaufmann, G. (eds.). *Variation in Language: System- and Usage-Based Approaches*, 205–240. Berlin/Boston: De Gruyter Mouton.
- Hajičová, Eva. 1994. Topic/focus and related research. In Philip A. Luelsdorff (ed.) *The Prague School of Structural and Functional linguistics*. Amsterdam/Philadelphia: John Benjamins, 245–275.
- Hakulinen, Auli, and Fred Karlsson. 1979. *Nykysuomen lauseoppi*. Helsinki: SKS.
- Hakulinen, Auli, Maria Vilkkuna, Riitta Korhonen, Vesa Koivisto, Tarja Riitta Heinonen, and Irja Alho. 2004. § 531 Muille kuuluva työ vs. työ joka kuuluu ~ kuului muille. In (eds.) *Iso suomen kielioppi*. Helsinki: SKS.
<<http://scripta.kotus.fi/visk/sisallys.php?p=531>>

- Harder, Peter. 2005. Dansk Funktionel Lingvistik. En Introduktion. In Lars Heltoft (ed.) *Dansk Funktionel Lingvistik: En helhedsforståelse af forholdet mellem sprogstruktur, sprogbrug og kognition*. København og Roskilde: Roskilde Universitet, 1-38.
- Harris, Alice. C., and Lyle Campbell. 1995. *Historical Syntax in Cross-Linguistic Perspective*. Cambridge University Press.
- Haspelmath, Martin. 1999. Optimality and diachronic adaptation. *Zeitschrift für Sprachwissenschaft*, 18(2), 180–205.
- Hawkins, John. A. 1994. *A Performance Theory of Order and Constituency*. Cambridge University Press.
- Hawkins, John. A. 2004. *Efficiency and Complexity in Grammars*. Oxford: OUP.
- Hawkins, John. A. 2014. *Cross-Linguistic Variation and Efficiency*. Oxford: OUP.
- Hawkins, John. A., and Murray Gell-Mann. 1992. *The Evolution of Human Languages. Proceedings of the Workshop on the Evolution of Human Languages*, held August, 1989 in Santa Fe. Boston: Addison-Wesley.
- Hjelmslev, Louis. 1943. *Omkring Sprogteoriens Grundlæggelse*. Copenhagen: The Linguistic Circle of Copenhagen.
- Hockett, Charles. F. 1955. How to learn Martian. *Astounding Science Fiction*, October 1955. Available at <http://gotomars.free.fr/How_to_Learn_Martian_1955.pdf>
- Hopper, Paul. 1987. Emergent grammar. *Annual Meeting of the Berkeley Linguistics Society* 13, 139–157).
- Hull, David L.: 1988. *Science as a Process*. University of Chicago Press.
- Itkonen, Esa. 1999. Functionalism yes, biologism no. *Zeitschrift für Sprachwissenschaft*, 18(2), 219–221.
- Itkonen, Esa. 2011. On Coseriu's legacy. *Energeia* III, 1-29.
- Jakobson, Roman. 1969. Linguistics and adjacent sciences. In *Actes du Xe Congrès International des Linguistes*, Vol.1, 75–11
- Karlsson, Fred. 2007. Constraints on multiple center-embedding of clauses. *Journal of Linguistics*, 43(2), 365–392.
- Keenan, Edward. L., and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry*, 8(1), 63–99.
- Kefer, Michel 1986. On functional explanations of language change, universals and rules. *Belgian Journal of Linguistics* 1: 53–66

- Keller, Rudi. 1994. *On Language Change: The Invisible Hand in Language*. London: Routledge. Translation and expansion of the original German *Sprachwandel: Von der unsichtbaren Hand in der Sprache*.
- Keller, Rudi. 1997. In what sense can explanations of language change be functional?. *Trends in Linguistics Studies and Monographs*, 98, 9–20. Berlin/Boston: De Gruyter Mouton.
- Kim, Tae 2012. *Japanese Grammar Guide*.
<http://www.guidetojapanese.org/grammar_guide.pdf>
- Kirby, Simon. 1999. *Function, Selection, and Innateness: The Emergence of Language Universals*. OUP Oxford.
- Kirby, Simon. 2011. *The Language Organism: evolution, culture, and what it means to be human*. Inaugural lecture, <<https://www.ed.ac.uk/arts-humanities-soc-sci/news-events/lectures/inaugural-lectures/archive/inaugural-lectures-2010-2011/simon-kirby>>
- Kiss, Tibor, and Artemis Alexiadou (eds.) 2015. *Syntax – Theory and Analysis: An International Handbook*. Berlin/Boston: De Gruyter Mouton.
- Kluender, Robert, and Marta Kutas. 1993. Subjacency as a processing phenomenon. *Language and Cognitive Processes*, 8(4), 573–633.
- Lai, Vicky. T, Olivia Howerton, and Rutvik H. Desai. 2019. Concrete processing of action metaphors: Evidence from ERP. *Brain Research*, 1714, 202–209.
- Lakoff, George, and Rafael Núñez. 2000. Where mathematics comes from: How the embodied mind brings mathematics into being. *AMC*, 10, 12.
- Lakoff, George, and Mark Johnson. 1999. *Philosophy in the Flesh*. New York: Basic Books.
- Langacker, Ronald. W. 2010. *Grammar and Conceptualization*. Berlin/Boston: De Gruyter Mouton.
- Lehmann, Christian. 1986. On the typology of relative clauses. *Linguistics*, 24(4), 663–680.
- Li, Charles N., and Sandra A. Thompson. 1974a. An explanation of word order change SVO→SOV. *Foundations of Language*, 12(2), 201–214.
- Li, Charles N., and Sandra A. Thompson. 1974b. Co-verbs in Mandarin Chinese: verbs or prepositions? *Journal of Chinese Linguistics*, 2(3), 257–278.

- Lyell, Charles. 1863. *The Geological Evidences of the Antiquity of Man: with remarks on theories of the origin of species by variation*. London: John Murray.
- Martinet, André. 1955. *Economie des changements phonétiques: Traité de phonologie diachronique*. Berne: A. Francke.
- Matsumura, Kazuto. 1982. Two Types of Relative Clauses in Finnish. *GENGO KENKYU* (Journal of the Linguistic Society of Japan), 81, 60–82.
- Müller, Gereon 2015. Optimality-Theoretic syntax. In Tibor Kiss, and Artemis Alexiadou (eds.) *Syntax – Theory and Analysis: An International Handbook*. Berlin/Boston: De Gruyter Mouton.
- Müller, Max. 1870. Darwinism tested by the science of language (review). *Nature* 1, 256–259.
- Müller, Stefan 2015. HPSG - a synopsis. In Tibor Kiss, and Artemis Alexiadou (eds.) *Syntax – Theory and Analysis: An International Handbook*. Berlin/Boston: De Gruyter Mouton.
- Nagel, Ernest. 1961. *The Structure of Science: Problems in the Logic of Scientific Explanation*. Reprint. Indianapolis/Cambridge: Hackett.
- Nakagawa, Natsuko. 2016. *Information Structure in Spoken Japanese: Particles, Word Order, and Intonation*. PhD diss., University of Kyoto.
<<https://repository.kulib.kyoto-u.ac.jp/dspace/bitstream/2433/215634/2/dnink00779.pdf>>
- Nichols, J. 1984. Functional theories of grammar. *Annual Review of Anthropology*, 13, 97–117.
- Noonan, Michael. 1999. Non-structuralist syntax. In M. Darnell et al. (eds.) *Functionalism and Formalism in Linguistics*, Vol. 1, 11–31.
- Norde, Muriel. 2009. *Degrammaticalization*. Oxford University Press.
- Osborne, Timothy. 2015. Dependency Grammar. In Tibor Kiss, and Artemis Alexiadou (eds.) *Syntax – Theory and Analysis: An International Handbook*. Berlin/Boston: De Gruyter Mouton.
- Özge, Duygu, Theodoros Marinis, and Deniz Zeyrek. 2010. Production of relative clauses in monolingual Turkish children. *BUCLD34 Proceedings*.
- Papagno, Costanza, and Carlo Cecchetto. 2019. Is STM involved in sentence comprehension?. *Cortex* 112, 80–90

- Payne, Doris. L. 1987. Information structuring in Papago narrative discourse. *Language* 63(4), 783–04.
- Peeters, Bert. 1998. Cognitive musings. *Word*, 49(2), 225–237.
- Peirce, Charles. S. 1998. *The categories defended*. In *The Essential Peirce*.
Bloomington: Indiana University Press, 160–178.
- Pike, Kenneth. L. 1943. Taxemes and immediate constituents. *Language*, 19(2), 65–82.
- Polichak James W. 2002. Memes as pseudoscience. In M. Shermer, and P. Linse. (eds.)
The Skeptic Encyclopedia of Pseudoscience, Vol. 1. Santa Barbara: ABC Clío.
- Richards, Marc. 2015. Minimalist Program. In Tibor Kiss, and Artemis Alexiadou (eds.)
Syntax – Theory and Analysis: An International Handbook. Berlin/Boston: De
Gruyter Mouton.
- Robbins, Philip. 2017. Modularity of mind. In Edward n. Zalta (ed.) *The Stanford
Encyclopedia of Philosophy*.
<<https://plato.stanford.edu/archives/win2017/entries/modularity-mind/>>.
- Saeed, John. 1999. *Somali*. Amsterdam/Philadelphia: John Benjamins.
- Saussure, Ferdinand de. 1916. *Cours de linguistique générale*. Paris: Payot.
- Saussure, Ferdinand de. 1959. *Course in general linguistics*. Translation of Saussure
(1916) by Wade Baskin. New York: Philosophical Library.
- Schleicher, August. 1869. *Darwinism tested by the science of language*. Translated by
Alex V. W. Bickers from Schleicher 1863 "Die Darwinsche Theorie und die
Sprachwissenschaft". London: John Camden Hoten.
- Schlesewsky, Matthias, and Ina Bornkessel-Schlesewsky. 2013. Computational
primitives in syntax and possible brain correlates. In Boeckx, C., and Grohmann,
K. K. (eds.). 2013. *The Cambridge Handbook of Biolinguistics*. Cambridge
University Press.
- Schmidtke-Bode, Karsten. 2009. *A Typology of Purpose Clauses*.
Amsterdam/Philadelphia: John Benjamins.
- Schwarz-Friesel, Monika. 2012. On the status of external evidence in the theories of
cognitive linguistics: compatibility problems or signs of stagnation in the field?
Or: why do some linguists behave like Fodor's input systems?. *Language
Sciences* 34, 656–664.

- Sériot, Patrick. 1999. The impact of Czech and Russian biology on the linguistic thought of the Prague Linguistic Circle. *Prague Linguistic Circle Papers* 3, 15–24.
- Seuren, Pieter A. 2015. Prestructuralist and structuralist approaches to syntax. In Tibor Kiss, and Artemis Alexiadou (eds.) *Syntax – Theory and Analysis: An International Handbook*, 134–157. Berlin/Boston: De Gruyter Mouton.
- Song, Jae Jung. 2010. Word order typology. In J. J. Song (ed.) *The Oxford Handbook of Linguistic Typology*. Oxford: OUP.
- Tesnière, Lucien. 1959. *Éléments de Syntaxe Structurale*. Paris: Klincksieck.
- Van Valin, Robert. D., Jr. 1995. Role and Reference Grammar. In Jef Verschueren, Jan-Ola Östman, and Jan Blommaert (eds.) *Handbook of Pragmatics: Manual*. Amsterdam/Philadelphia: John Benjamins, 461–469.
- Vincentini, Alessandra. 2003. The Economy Principle in Language. Notes and Observations from Early Modern English Grammars. *Mots, Palabras, Words* 3.
- Wells, Rulon S. 1947. Immediate constituents. *Language*, 23(2), 81–117.
- Zhou, X., Li, M., Li, L., Zhang, Y., Cui, J., Liu, J., and Chen, C. 2018. The semantic system is involved in mathematical problem solving. *Neuroimage* 166, 360–370.