The definitive solution of dominant word order

Since the 1960s, various models have been proposed to explain cross-linguistic variation of the transitive clause. A persistent problem has been the difficulty of making correct generalisations to account for asymmetries in the empirical data. In the absence of a clearly formulated null hypothesis, higher-level theorising has lacked a scientific grounding. This paper examines word-order research to uncover the null hypothesis. The simplest explanation of the attested distribution pattern is that the conceptual similarity of subject (S) and object (O), together with their statistical difference, gives rise to the preference of highlighting the secondariness of object in the basic order. As an outcome of primacy highlighting, fronting SO is preferred cross-linguistically. The inverse, fronted OS, is the most infrequent type.

Keywords: transitive; markedness; topicalization; word order; dominant; philosophy of science

1. Introduction

The basic or dominant word order of the transitive clause is perhaps the single most frequently cited feature in typological literature. Yet, it is still to be established why some orderings of subject (S), object (O), and verb (V) are utterly common as a dominant order cross-linguistically and others strikingly rare. It has been suggested that statistical biases may depend on logical and psychological universals, on all or most languages having a common ancestor, or on the way languages evolve and spread.

When a new language arises from a proto-language, the most reliable predictor is inheritance: in 78 percent of the cases, the new language will have the same order as its parent language (Hammarström 2015, 2016)[1]. We find examples in the British Isles where Scots, deriving from Middle English, has the same SVO order as English. Celtic

[1] Hammarström's presentation slides (Hammarström 2015) offer a more comprehensive analysis of his database research than the published paper (Hammarström 2016).

languages, including Cornish, Irish, Manx, Scottish Gaelic, and Welsh, derive from a common proto-language with VSO, maintaining its order. This is not always the case. Modern Romance languages such as French, Italian, and Spanish have SVO despite deriving from Latin, which has free word order with a high frequency of SOV.

The effect of language contact is estimated at eight percent (Hammarström 2015, 2016). Many Uralic languages, for example, Estonian and Finnish, have prehistorically shifted from SOV to SVO, the word order of their neighbouring languages including Latvian, Russian, and Swedish. In contrast, Basque (SOV, isolate) has not adopted SVO order from its Romance neighbours.

There appears to be an unidentified factor accounting for the remaining 14 percent of transitive dominance. Hammarström's (2015, 2016) conclusion is that a universal force overrides genealogical and contact effects in the long run giving rise to the attested distribution pattern (figure 2). He bases his conclusion on a large corpus research, representing the "classical" universalist view, most famously advocated by Joseph Greenberg (1915–2001). With this, Hammarström goes against a current stream: there is a trend away from universal explanation in favour of evolutionary explanations (section 3).

2. Universals, the data, and the problem

An example of a transitive expression in English is the SVO sentence '*The girl kicks the ball*'. A transitive expression is a predicate–argument structure with two arguments as in $P(x,y)$. The three members are traditionally called verb, subject, and object although it would be more correct to call them event, agent, and patient because typological generalisations are made on a semantic basis. All languages have a means to express transitive events or actions, but their surface forms do not always bend to an analysis in terms of subject and object.

Typological examples of different dominant transitive orders include the following.

a. Hindi (SOV)

उमा कमीज़ लाई।

uma: *kami:z* *la:i:*

Uma-NOM shirt brought

'Uma brought a shirt.' (Koul 2008, 38)

b. Mandarin (SVO)

爸爸爱妈妈

Bàba *ài* *māmā*

dad love mom

'Dad loves mom.'

c. Standard Arabic (VSO)

كتب الطلاب الدرس

kataba T-Tulaab-u d-dars-a

wrote.3MS the-students-NOM the-lesson-ACC

‘The students wrote the lesson.’ (Jouini 2018, 234)

d. Malagasy [VOS]

misotro ronono izahay

drink milk we

‘We are drinking milk.’ (Keenan and Ochs 1979, 119)

e. Panare [OVS]

piʔ kokampö unkïʔ

child washes woman

‘The woman washes the child.’ (Song 2012, 16)

f. Nadëb [OSV]

awad kalapéé hapʉ́h

jaguar child see.IND

'The child sees the jaguar.' (Weir 1994, 309)

Languages also make use of a variety of orderings, but a language is considered to have a dominant or basic word order if one is attested in texts at least twice as frequently than any other (Dryer 2013a). On such grounds, a number of languages are labelled as having no dominant order. Well-known European examples include Dutch, German and Hungarian (SOV with SVO), and modern Greek (SVO with VSO), each with two dominant orders. Based on Dryer, Belarusian has none based on the 2:1 ratio. All cases lacking a single dominant order ("NO-DOM"), making up less than ten percent of all languages, are often put aside when looking for an explanation for the universals, which are more correctly statistical biases (Dryer 2013b).

From a purely logical point, any ordering should be just as good as the other. Modern predicate logic uses VSO ordering out of notational convenience, but it is not purported to represent any logical reality. What is more, even though there are free word-order languages, each attested transitive clause must have materialised with one of the six logically available orderings: SOV, SVO, VSO, VOS, OVS, or OSV, with V representing the predicate or finite verb. Since there is no reason from a logical point why any of the orderings should be superior or inferior, their distribution is expected to

be random by default, each landing at around 16.7 percent or 1/6. The actual statistics is far removed from that, necessitating an explanation of the attested distribution (Comrie 1981, 19; see figure 1 and 2).

2.1 Adjusting for a colonial effect

Since a majority of linguists speak English and other European languages with SVO, it may strike as odd that SOV should be the most common order cross-linguistically. It is nonetheless a well-established fact. The distribution hierarchy, from the most to the least frequent, is SOV > SVO > VSO > VOS > OVS > OSV. Additionally, the frequencies of SOV and SVO are relatively close. These are clearly the two most common orderings and, following them, values drop more systematically.

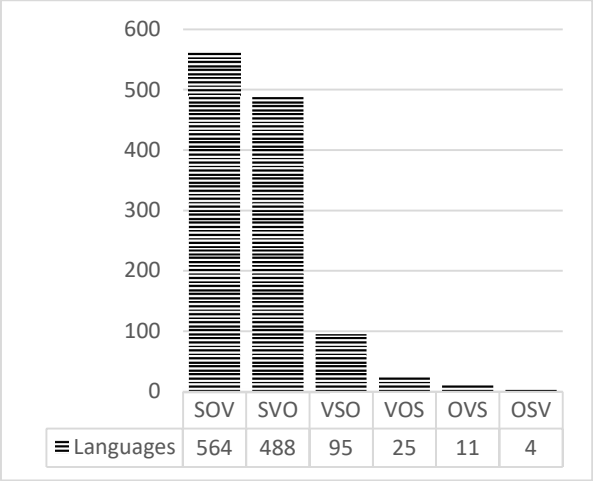

Figure 1. A dataset (Dryer 2013b) demonstrating the transitive distribution when counting languages.

Hammarström (2015, 2016) explains this to be an effect of colonialism. Hundreds of languages have vanished in the Americas, especially on the Brazilian East coast, and in Australia. These are known to have been non-SVO hotspots. The effect is much weaker in Africa and Asia, which form a landmass with Europe.

It is well-established that diseases contributed to the downfall of hundreds of native nations in the new continents (Cook 1998), so it is not so surprising that we find the effect of colonialism to have been far more moderate in Africa and Asia. For example, it is suggested that the bubonic plague, once wreaking havoc in Europe, originated in or around China, spreading via trade routes (Morelli et al. 2010). Along these routes, we find today different types of word-order hotspots that have withstood the Colonial Period. Namely, SOV hotspots on New Guinea and the Indian subcontinent, and SVO hotspots in Sub-Saharan Africa and in mainland South East Asia (Dryer 2013b).

At the same time, there are conspicuously empty spots in the Caribbean, continental America, Australia, Tasmania, and New Zealand. Based on areal linguistics and what is known from historical sources, these have previously included further non-SVO hotspots. According to Hammarström's (2015, 2016) analysis, it is their disappearance that explains the apparent relatively high frequency of SVO in comparison to all other types. To be clear, colonialism did not significantly increase the number of SVO languages as the spread of English, Spanish, Portuguese, and others does not add to the number of SVO languages. Rather, it is a historical accident that they would mostly come to wipe out other types.

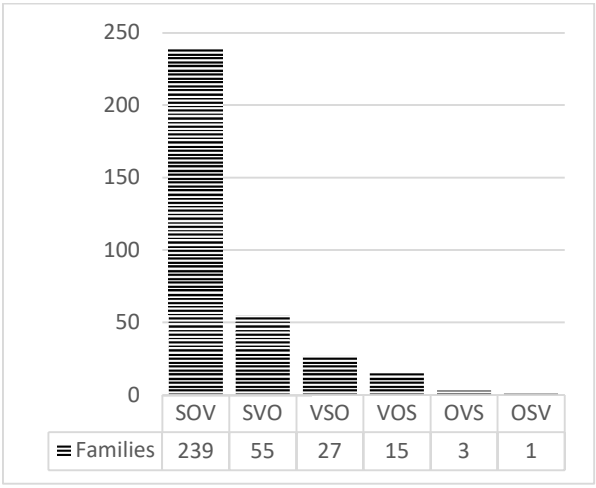

Figure 2. A dataset (Hammarström 2016) demonstrating that when language families and isolates are counted, SOV is clearly the most frequent type.

For this reason, Hammarström (2015, 2016; cf. figure 2) makes a stratified analysis to uncover a pre-colonial distribution, which turns out to be heavily dominated by SOV, showing no sign of a relatively high tendency for SVO.

One may argue that the current distribution is nonetheless significant because it is the natural outcome of the cultural evolution of mankind. It may be, but bear in mind that there is no evidence that Western European nations were the best at colonialism *because* they spoke SVO languages. There is also no evidence that word order affects resistance to diseases. As such, it seems correct to conclude that the relatively high frequency of SVO is a by-effect of colonialism, not a universal feature.

3. Previous attempts

Greenberg (1963) presented a set of transitive data as part of his ground-breaking research on linguistic universals. He based his cross-linguistic analyses on markedness theory as advocated by Roman Jakobson (Battistella 2015). Jakobson's (1963) explanation of the basic word order of the transitive clause is that object, when fronted, is perceived as emphatic. Thus, SO is preferred over OS as the basic, unmarked order. Greenberg found markedness theory instrumental for uncovering a number of universals, building his implicational hierarchies on the concept, but it left the explanation of the basic transitive order incomplete. Although Jakobson's pragmatic-communicative principle correctly predicts the SO > OS preference, it has not been extended to an account of the full tripartite pattern.

Instead, what Greenberg's (1963) analyses provided were further insights into the old observation that the morphosyntactic structures of SOV languages, such as

Turkish and Japanese, mirror those of SVO languages. Consequently, Lehmann (1973) and Vennemann (1974) removed the subject from consideration, giving rise to OV/VO typology, which is today considered to have been a more fruitful enterprise than seeking a solution to the problem of the transitive clause (Song 2012).

Nonetheless, the quest never ended, and since the 1970s there has been a steady flow of attempts to explain the full transitive distribution. Classical universalist explanations (section 3.2) dominate until the early 2000s when they become challenged by evolutionary explanations (section 3.3).

It remains a common problem that the proposed models fail to make the right prediction. Among all the proposals presented in the rest of this section, there is just one model, by Manning and Parker (1989), making the correct one. It will be selected for further discussion in section 4. Before moving to the explanatory models, section 3.1 will discuss terminology relating to "explanation"—hypothesis, theory, model, prediction, generalisation, data etc.—to help make sense of the different levels or stages of theorising.

3.1 Levels of explanation

McCawley (1982) calculated that a handful of variables gives potential to thirty-million linguistic theories. He points to a large number of linguistic theories proposed in the 1970s. These were made within the frameworks of generative grammar and generative semantics, which themselves are theories, so there must be theories within theories or explanations within explanations owing to a link between the two concepts. The same terminological ambiguity pertains to typology. There is no universal taxonomy of theory types for the likely reason that science progresses stage by stage. In the ideal case, yesterday's theory is today either a rigorously established fact or an outdated notion. What follows below is a hierarchy of theories/explanations made for present

purposes.

Level 1: distribution. The data pattern is presented in figure 1 and 2. The question of which statistical pattern is right remains relevant for section 3.2 and 3.3 because many of the models were designed to fit datasets that are now considered outdated. Some of the later models still use old data, suggesting a lack of consensus. In any case, it is vital to bear in mind that, when discussing "evidence" for a given claim, such evidence could relate to Level 1 rather than to the higher levels which are the main focus of the present discussion.

Level 2: generalisation. The next task is to make a generalisation, or a mathematical model, to explain or "predict" the Level-1 distribution pattern. Note that as the data have already been made available by linguists making mass-comparisons, work on the next stage, Level 2, is inductive. Predictions made by the models do not chronologically predict the data but, rather, explain it on an a posteriori basis. Thus, 'prediction' may seem like a mere fancy term, but its use is justified by pointing to ambiguities concerning the word 'explanation'. Instead, a difference between 'data' and 'prediction' seems intuitive.

Level 3: causation. The next step is to propose a theory or explanation for the Level-2 generalisation: Why is it that a given model correctly predicts the distribution pattern? Theorists operating at this level often take an interdisciplinary approach proposing explanatory principles from other sciences including history, psychology, and biology. Others stay closer to general linguistics with principles from semiotics or communication theory.

Level 4: theoretical framework. Level-3 theorists often advocate a given view of linguistics representing an established school of linguistic thought. Author (year) sees a main divide between humanistic and sociobiological theories, and these are further

divided into camps favouring a historical or a systemic mode of explanation. Others (e.g. Newmeyer 2017) suggest that the field is divided into "formalists and functionalists": proponents and opponents of linguistic innatism.

Further up the ladder, there are levels of explanation linking theories of language to different fields of science, to philosophy of science and, ultimately, to an explanation of the universe. At the lower end, there is likewise a level zero concerned with analytical tools and concepts including word classes, parts of speech, etc. There are also important questions concerning criteria for the data analyses in question, but these are taken for granted in this paper. For a discussion of such problematics, see e.g. Frey (2015).

3.2 Classical universalist explanations

Inspired by Greenberg's work on linguistic universals, there were several attempts to explain dominance data in the 1970s. For a long time, the existence of OVS and OSV languages was controversial (Tomlin 1986), so it may have been thought that these orderings are unnatural as dominant ones. Such a notion is echoed in the 1980 science-fiction film *The Empire Strikes Back*, which introduces an extra-terrestrial character called Yoda, who speaks "Yoda-speak", that is, English with strange syntax including OSV sentences, creating an alien atmosphere (LaFrance 2015).

Greenberg's sample included only three types: SOV, SVO, and VSO languages although well-known examples of VOS include Malagasy, the main language of Madagascar with 25 million speakers. It is added by Ultan (1969).

Bach (1974) suggests that all languages have VSO as an underlying constituent order, generating their surface form with a maximum of two transformation rules and materialising as one of the three possible SO varieties. He suspects that OVS languages

are fundamentally VSO languages with an O-thematisation rule. On such a basis, one would expect VSO to be the commonest.

Emonds (1980), proposing markedness rules, arrives at the false prediction SVO > SOV = VOS > OVS, with the equation mark indicating similar frequency (Manning and Parker 1989).

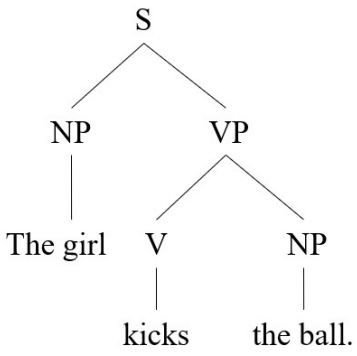

Figure 3. A parse tree for '*The girl kicks the ball*' based on classical grammar[2] arises from two operations: (i) the sentence is divided into subject and predicate (NP, VP), and (ii) the predicate is divided into verb and object (V, NP). Operation (i) is compatible with SV/VS typology and verb-object bonding, but prioritising smallest changes from optimal order incorrectly predicts SVO > SOV > VOS > OVS > VSO > OSV.

Like many others, Culicover and Wexler (1974) take the generative sentence split to subject and verb phrase to be a cognitive universal (cf. figure 3), making conclusions concerning nominative-accusative and ergative languages, with their model apparently suggesting that SVO and OVS should be the most optimal types (Tomlin 1986, 12). It makes the further incorrect predictions that VSO languages should be ergative, but not SOV languages. This idea is rejected by Schwartz (1974).

---

[2] According to Seuren (1998), classical grammar arises from the writings of Plato, persists through history, and is currently defended by generative grammar.

Diehl (1975) proposes generative constraints limiting the possible combinations to SOV, SVO, VSO, and VOS, thus predicting the absence of O-initial languages. Additionally, Pullum (1977) accounts for the relative rarity of VOS in his model. However, Derbyshire (1977) and Derbyshire and Pullum (1981) report the discovery of OVS and OSV, effectively making previous proposals obsolete.

The 1980s see attempts to solve the problem by combining three pragmatic principles into a unified explanation. One of these is provided by Mallinson and Blake (1981), who equate S-initial with topic-initial, predicting the prevalence of S-initial languages. But Manning and Parker (1989) point out that object can frequently be construed as the topic, suggesting Mallinson and Blake's set cannot account for the rarity of O-initial. Krupa's (1982) set can, but its predictions are nonetheless incorrect, asserting SVO > SOV and OVS = OSV.

Tomlin (1986) improves on three-way models in conjunction with his proposal for a six-way hierarchy on Level 1: SOV = SVO > VSO > VOS = OVS > OSV. Asserting a lack of statistical significance (table 1), he conflates the frequencies of SOV and SVO, and the frequencies of VOS and OVS. Tomlin's sample does not include any OSV languages. He constructs a weighted model with two principles reinforcing SO > OS to explain the primacy of SOV, SVO, and VSO over the remaining types, and just one principle reinforcing verb–object bonding, arriving precisely at the proposed hierarchy. The more principles out of the three a given ordering satisfies, the more frequently it will appear as a dominant order (table 1).

	Frequency	TFP	AFP	VOB
SOV	44.78 %	+	+	+
SVO	41.79 %	+	+	+
VSO	9.20 %	+	+	-
VOS	2.99 %	-	-	+
OVS	1.24 %	-	-	+
OSV	0.00 %	-	-	-

Table 1. Based on his data (cf. the frequency column) Tomlin (1986) proposes the hierarchy SOV = SVO, VSO, VOS = OVS, OSV. The more principles an ordering satisfies, the more frequently it will appear as a basic order cross-linguistically.

Tomlin's two principles reinforcing SO > OS include TFP (Thematic First Principle) and AFP (Animated First Principle). TFP agrees with the classical notion that subject represents the theme or topic of the sentence, and object represents the rheme: a comment about the topic. Thus, it is natural for the topic to appear before the comment, and the subject is more likely to contain topical information.

AFP, on the other hand, builds on the notion that subjects are statistically more often animate than objects, as in our example sentence "The girl kicks the ball". In some languages, such as Navajo, word order is based on animacy, and some animacy effects are probably attested in all languages. Therefore, the Animate First Principle should be part of the explanation of the biases.

Song (2012, 28) is critical of Tomlin's solution based on AFP and TFP together weighing more than VOB, and suggests conflating the two whereby Tomlin's model would collapse. On the other hand, Hawkins (1994, 279) points out that Tomlin gives no proper motivation why the verb should bond with the object. On Level 3, VOB is little more than the reiteration that, in the data, O is adjacent to V more often than it is not.

Dryer (1989) dismisses Tomlin's claim of no statistical significance between SOV and SVO, pointing out that it was created by an over-representation of a couple of large SVO families. Dryer demonstrates a clear SOV > SVO preference outside the Afro-Eurasian landmass (table 2).

	Africa	Eurasia	Aus-NG	NAmerica	SAmerica	families
SOV	**22**	**26**	**19**	**26**	**18**	**111**

SVO	21	19	6	6	5	57

Table 2. An areally and genealogically balanced dataset (Dryer 1989) demonstrates the overall preference for SOV over SVO (Aus-NG = Australia-New-Guinea, NAmerica = North America, SAmerica = South America).

Yet another way to approach the problem is via semantics. According to Maxwell (1984), Edward Keenan asserted in his lectures in 1980 that syntactic order is directly influenced by semantic relationships concluding that SOV and VOS should be the two preferred orderings (see Manning and Parker 1989). Hawkins (1983) makes the same incorrect prediction, but argues that it arises from formal symmetry issues relating to OV versus VO languages. Maxwell (1984) makes his attempt at semantic explanation arguing for a preference for V-medial on the basis that verb should link subject and object semantically, arriving at the likewise incorrect prediction that SVO and OVS should be preferred.

Manning and Parker (1989) construct a model making the exceptional correct prediction SOV > SVO > VSO > VOS > OVS > OSV. As a whole, their proposal falls into near-oblivion following a dismissal by Peeters (1991) and Song (1991), but since it at least appears to be the only model predicting the correct hierarchy, it is selected here for discussion in section 4.

Dryer (1991, 1997) proposes approaching the problem of the transitive clause indirectly, suggesting that the three-partite pattern could arise as a by-effect of related universals. He prefers focusing on the OV/VO and SV/VS variables (cf. figure 3) and to draw conclusions from them. Though first suggested by Greenberg (1963), the idea lacks typological appeal because both OV and SV are attested at the high and low end of the transitive distribution, suggesting a lack of explanatory value. What is more, OV/VO typology, despite its higher overall success than transitive typology, was

specifically postulated because the place of the subject does not correlate well with other kinds of syntactic universals (Song 2012). Dryer (2013c) maintains his position but admits a lack of further progress on this front.

Unlike Dryer, Hawkins (1994) accepts Tomlin's hierarchy approaching the problem from a different angle. He proposes rejecting the three-principle explanation altogether in favour of his parsing model which he later names the Performance–Grammar Correspondence Hypothesis (PGCH, Hawkins 2004). Hawkins justifies using Tomlin's data citing Polinskaja (1989), who disputes the claim that OSV languages truly exist, to support the VOS = OVS > OSV ranking.

According to Hawkins, constituencies are ordered from shortest to longest in VO languages and longest to shortest in OV languages. Focusing on the most common type, SOV, he suggests that subject is the longest constituency, object is second longest, and verb is the shortest of the three, initially arriving at his (Hawkins 1983) earlier false conclusion SOV = VOS. However, Hawkins (1994) points to further structural differences between OV and VO languages, arguing that SVO optimally organises constituents from shortest to longest because they have heavy objects, unlike SOV languages, which have heavy subjects. Song (2012) is sceptical of Hawkins's solution, arguing it is based on a too small sample of languages to support the claim. He also reiterates that Dryer (1989) has already demonstrated Tomlin's hierarchy to be mistaken. All in all, Song sees the virtue of Hawkins's model in its explanation of OV/VO rather than in transitive typology.

If languages are capable of making fine-tuned adjustments to sustain a given basic order, then why do such adjustments not take place throughout the hierarchy giving all languages equal parsing efficiency? Hawkins argues that some word orders are efficient in terms of language processing, while others cause processing difficulty.

However, a slowly growing body of brain imaging research into the transitive clause suggests that language users are cross-linguistically best at processing the dominant order of their language, which often requires fewer marking particles than other acceptable orderings (Koizumi and Kim 2016).

Research in the third millennium makes increasing use of computational techniques to solve the problem of the transitive clause. Many of these will be discussed in the context of evolutionary linguistics in section 3.3, but there are also classical universalist attempts.

Since languages appear to be rather evenly split between OV and VO, Cysouw (2008) proposes focusing on OS/SO and SV/VS. The two parameters combined, he arrives at the prediction SOV = SVO, VSO, OSV, VOS = OVS which is incorrect, most notably because OSV should be last.

Maurits, Perfors, and Navarro (2010) propose just one pragmatic principle: even distribution of information throughout the sentence. They purport that SVO is preferred to SOV because the import of a sentence is anticipated from the combination of subject and object. For example, when we hear the words *The girl* and *the ball*, we anticipate the verb *kicks*. Thus, to keep up suspense until the end, verb is placed before object. This proposal has a number of problems, most notably that it makes the false prediction SVO > SOV. The researchers suggest that an information distribution principle could nonetheless be part of an explanation, but they fail to consider that the object could also be anticipated from the combination of subject and verb: *The girl kicks…* (*the ball*) whereby SVO should also be unpreferred.

We will continue to see in the next sub-section that computational methods are increasingly made use of, but the above example suggests that programming provides

some impressive tools that do not always add to the traditional paper-and-pen method. There is no know attempt to use artificial intelligence to create an explanatory model.

3.2 Evolutionary explanations

Universalist linguistics is challenged by Evans and Levinson (2009) as the twenty-first century sees a new rise of evolutionary linguistics. Rather than assuming *a priori* conditions underlying linguistic structures, evolutionary linguists propose a fully empirical approach to the study of language. However, due to lack of sufficient historical data, reconstructive methods are employed in an attempt to explain the transitive distribution on a genealogical basis. Computational techniques are also used by scholars aiming to explain the data in terms of language contact and adaptation.

Gell-Mann and Ruhlen (2011) suggest that the transitive distribution pattern depends on all or most languages deriving from a common ancestor with SOV, followed by a switch to SVO. It is purported to be the current trend, while some languages have already taken the next evolutionary step by turning V-initial. The idea is not new, it borrows from Givón ([1979] 2018) and Bickerton (1981), whose concepts share similarities with the nineteenth-century Neo-grammarians, who proposed laws of linguistic change to uncover computational patterns that, when reversed, would make it possible to reconstruct lost ancient languages (cf. Norde 2009).

Based on an interpretation of their own dataset, Gell-Mann and Ruhlen (2011) demonstrate that SOV languages are likely to evolve to SVO. This claim, too, is based on Givón ([1979] 2018), who stipulates that the reverse, when attested, depends on language contact. Thus, the raw data do not support Gell-Mann and Ruhlen's (2011) claim, and the positive results are produced by ignoring real-life changes to SOV. In contrast, Hammarström (2015, 2016), treating any change equally, finds the diachronic change pattern to be the same as the synchronic pattern: when a language switches to a

new dominant order, the likelihood pattern is likewise SOV > SVO > VSO > VOS > OVS > OSV, giving rise to a stable synchronic pattern.

Based on his review, Sinnemäki (2014) nonetheless claims that there is evidence showing that the way languages change gives rise to syntactic universals, supporting Christiansen and Chater's (2008) view of language as an organism which adapts to the human brain or the speech community (see also Sinnemäki and Di Garbo 2018, citing Beckner et al. 2009). Research cited by Sinnemäki (2014) suggests that Chomsky's notion of a genetic grammar is incompatible with experimental findings. However, there is no direct link to the claim that universals arise diachronically, i.e., as linguistic units struggle for life by adapting to their changing environment. In fact, Christiansen and Devlin ultimately support *a priori* learning and processing conditions as causes of the transitive distribution. More precisely, Christiansen and Devlin (1997) propose a branching approach to transitive universals citing Hawkins (1994, cf. section 3.2).

The idea of an SOV Proto-World language has some intuitive value. Indo-Aryan languages, including Sanskrit, had SOV, as does Persian. Following the Indo-European expansion westward, Latin has free word order with an inclination for SOV, but the subsequent vernaculars (Italian, French, Spanish etc.) have SVO, and at the West end of the European map (Dryer 2013b), we find Celtic languages with VSO. Hammarström's (2015, 2016) conclusion from this is different, suggesting there were a larger number of Celtic VSO languages in Western Europe until they were replaced by SVO languages. Thus, there is no indication of Europe now turning VSO.

Moving eastward from SOV-dominant India, we find a mainland South-East Asian hotspot of SVO languages, followed by a human expansion into the Pacific with many V-initial languages including Rapanui (Easter Islands, VSO) at the East end of the

Austronesian language map (Dryer 2013b). Australian data does not support the pattern equally well.

Historical evidence for an SOV Proto-World is controversial and has, for instance, been resisted by Chinese historical linguistics (Chappell, Ming, and Peyraube 2007). The estimated time span is also not favourable. Based on archaeological evidence, human language may have emerged as long as 50,000 years ago (Gell-Mann and Ruhlen 2011), though some estimate it to be up to seven times older (Perreault and Mathew 2012). The oldest surviving "texts", written in proto-cuneiform Sumerian, are less than six-thousand years old (Chrisomalis 2009). This means that the amplitude of language change lending itself for observation and reliable reconstruction could have occurred manyfold in the prehistory. There is no guarantee that any relics from a Proto-World survive in the authentic material, or that we can identify them as being such.

Even assuming the Proto-World hypothesis, the explanatory value of the enterprise is called into question. Gell-Mann and Ruhlen (2011) appear to be proposing a kind of *historicism*: Rapanui has the order it has because it first had SOV. Then it changed to SVO, and then it changed to VSO. For lack of preserved data, we do not know whether this pattern is historically correct. If we assume it is, for the sake of argument, even then the claim does not seem to fully address the question "Why does Rapanui have VSO rather than some other order?"

For such reasons, Gell-Mann and Ruhlen (2011) are criticised by Givón ([1979] 2018, 207f), who based his original idea on language learning issues. He speculates that childhood language learning patterns and primitive communication needs favour SV and OV, concluding that proto-human had SOV. However, it should not be ignored that the same pairs are also compatible with OSV. As was pointed out in (3.2), the

combination of SV > VS and OV > VO does not appear to have predictive value with respect to the transitive distribution pattern.

Bickerton (1981), building on the principle of natural selection, argues that Proto-World must have had SVO, which is better for disambiguation in the absence of highly evolved grammar including subject or object markers. Maurits et al. (2010), likewise, argue for the betterness of V-medial ordering (3.2), as does Ferrer-i-Cancho (2008). Ferrer-i-Cancho (2014) reviews a number of twenty-first-century historical and cognitive studies on the topic. He concludes that languages first had SOV, then turned SVO, and are now turning back to SOV—but admits that experimental findings do not shed sufficient light on the matter.

Ferrer-i-Cancho (2015) proposes that, despite its functional value, the low frequency of OVS is due to its long permutational distance from the original SOV order. But what if Proto-World was OVS? As it turns out, Maurits and Griffiths's (2014) analysis points to an O-initial Proto-World, but they nonetheless reconstruct it as SOV, guided by "expert opinion".

Whatever the preferences and non-preferences, Trudgill (2011, 100–101) suggests that non-preferred types only thrive in small speaker communities because it is difficult to find wide acceptance for non-optimal orders. Trudgill borrows from Nettle (1999, 139), who argues that the rare O-initial languages have a small median speaker number.

However, S-initial languages, being the commonest ones, are by default expected to constitute a majority of all languages, large and small alike. For example, most of the endangered and moribund languages of Europe have SVO because it is typical in the region. Hammarström (2015) tested Nettle's claim, finding that when

languages are sampled randomly, the median number of speakers of O-initial languages is actually slightly higher than that of SVO languages.

Finally, as already mentioned, Hammarström (2015) finds the effect of language contact, in terms of adopting the order of a neighbouring language, to explain eight percent of the observed dominant order variation. If it was much higher, obviously, languages would end up with the same dominant order in the long run, but that is not what is suggested by the diachronic and synchronic data.

4. How to explain the distribution

Summarising section 3, there have been many attempts to explain the typological transitivity distribution, but no clear breakthrough. Of all the models, only that of Manning and Parker (1989) predict the correct pattern at level 2 and is selected for further discussion in this section.

A fundamental issue underlying linguistic theorising concerns the problem of evidence (Song 1991). Authors of the models discussed in section 3 have been eager to present research evidence for their proposals to convince the reader of their usefulness. But why do models supported by solid scientific evidence fail to make the correct prediction? This issue is most evident in the context of the 1970s and early 1980s models that were abandoned by their creators (3.1).

The bulk of purported evidence for later models cannot be discussed here comprehensively, but to pick an example, Christiansen and Devlin (1997) suggest that word-order universals "may emerge" from non-linguistic constraints on learning. These constraints are corroborated by "typological language evidence". Thus, their proposal is based on circular reasoning. The authors do discuss findings from language processing research, as well as from typology, but fail to establish a link between the two, leaving their discussion to a speculative level. They advocate a parsing approach, but as

discussed in (3.2), these have failed to make a credible case on Level 2, that is, to make the correct prediction for the data.

What linguistics should have, instead, is a grounding for Level-2 and higher in empirical evidence from adjacent sciences such as psychology or neurobiology. To be perfectly clear, no such evidence is available, not in the cited literature or likely anywhere. Thus, the best anyone can do is to examine whatever material is available and give the most fitting, simplest explanation for it, staying as close to the null hypothesis as possible. This way it will be possible to compare the proposals and select the provisionally correct one, as is standard practice in science. We will see in the following sub-sections that principles of science are applicable to theoretical linguistics, too, as a sustainable explanation for the transitivity distribution is extracted from the material, step by step.

4.1 Selecting the right hierarchy

The Level-1 question is what exactly needs to be predicted. As seen in section 2, the prima facie hierarchy has been rather consistently SOV > SVO > VSO > VOS > OVS > OSV and is selected by default. Some have questioned it, most notably Tomlin (1986), but Song (2012) considers his alternative to have been rejected by Dryer (1989). While it remains a possibility that the prima facie hierarchy is not good, there is no obvious justification to replace it with a different one.

4.2 Selecting the right prediction

As stated in section 3, the only Level-2 model in the literature predicting the correct hierarchy as stipulated in (4.1) is that of Manning and Parker (1989).

A second model is added here. We make the generalisation that languages most frequently front SO and most infrequently OS. This notion is expanded into a prediction

of the whole hierarchy when proposing SO > OS, and that the front–back difference can be expressed with decreasing mathematical value, for example, by giving S or O three points in the initial position, two points in the middle position, and one point in the final position. However, if the order is OS, the constituents will receive negative points. These operations generate the following table.

Order	Points for position			Total points
SO	+2	+1	+0	
SOV	S	O	(V)	+3
SVO	S	(V)	O	+2
VSO	(V)	S	O	+1
OS	-2	-1	-0	
VOS	(V)	O	S	-1
OVS	O	(V)	S	-2
OSV	O	S	(V)	-3

Table 3. The correct frequency hierarchy is generated by granting decreasingly positive points (left to right) to subject and object in SO languages, and the corresponding negative points in OS languages.

There are now two models making the correct prediction SOV > SVO > VSO > VOS > OVS > OSV. These are selected to discuss why they do so (4.3) after a complexity comparison here. The novel prediction (table 3) is generated by four statements[3]:

i. Each position in the transitive clause gives a mathematical value to S or O.

ii. This value systematically decreases from left to right.

[3] The very simplest way to put it is (i) if SO, the value of V increases from left to right, and (ii) if OS, the value of V decreases from left to right. However, the place of V itself is not significant because V-final is most and least preferred, V-medial is second-most-and-least preferred, and V-initial is third-most-and-last preferred. Thus, the place of V appears to be a by-effect of the orientation of S and O in the clause.

iii. If SO, the value is positive.

iv. If OS, the value is negative.

Manning and Parker's (1989) model is more complex requiring the following statements:

i. Constituents are ideally ordered from smallest to largest.

ii. S is small.

iii. O is medium.

iv. V is large.

v. The first constituent from the left is assessed first.

Thus, S-initial (small first) is preferred over V-initial (large first),

vi. However, O is not independent but bound to the next constituent.

vii. O-bonding makes the O-initial orderings (OV)S and (OS)V front-larger than other types regardless of what is bonded with O.

viii. The rest of the sentence is assessed second.

Favouring small before large in step (viii), as ordered in (i), we arrive at the same six-way distribution as in table 3. Manning and Parker's model is broken down to eight statements, four more than the new one above. Much of the difference depends on the three stipulations defining the constituency sizes (S=small, O=medium, V=large), which are coincidentally the reverse of Hawkins's (1994) opinion regarding SOV. If these were proved to be a typological fact, they could be taken for granted, thus reducing the full account to five statements. But Song (1991) specifically points out that Manning and Parker's stipulations are not supported by typological evidence. Since there are no further candidates left, however, their model is selected for further comparison in the next sub-section.

4.3 Selecting the correct theory

Manning and Parker's (1989) explanation for their Level-2 prediction ventures quickly away from simple answers, attracting a dedicated critique from Peeters (1991). Manning and Parker first argue that their opponents represent the false doctrines of creationism and Lamarckism while they themselves advocate a correct view. The authors then substantiate their Level-2 prediction by an appeal to gestalt psychology and Peircean semiotics, concluding the journey with a discussion of "the acquisition and survival of grammars […] as analogous to the evolution and survival of rare biological species." More specifically, Manning and Parker base their explanation on the visual perception of figure and ground, speculating that a small constituency (S) is easier to perceive against a large background (OV).

Perception is undoubtedly relevant to visual and auditory processing alike, but as regards the auditory background, it is more readily understood as consisting of things like water burble, traffic hum, wind, radio, etc. It does not seem to be the case that verb functions as the background for subject in communicative situations. This does not prove that Manning and Parker's theory is wrong, only that it makes a series of alternative hypotheses, that is, claims which should be substantiated by empirical evidence. Thirty years on, the virtue of their model appears to have been limited to the fact that it makes the correct prediction, but why? The obvious reason is that, whatever the distribution, it is always possible to construct an explanation of some sort with a series of disconnected ad hoc solutions.

Still, the transitive distribution remains too heavily biased to be arbitrary, and after sixty years of research into it, there has been little progress. Maybe this is because the null hypothesis was not sufficiently clearly formulated, and it should be the task of

the present paper. The hope is that, once the null explanation is induced from the empirical evidence, it can be used as a foundation for further research.

Looking for a Level-3 explanation for the model proposed in (4.2), we notice that it cycles the SO > OS preference by adding that SO fronting is preferred, too.

Jakobson (1963), basing his idea on Russian word-order variation, argues that all OS orderings are emphatic. The implication is that people have a noticeable cross-linguistic tendency to perceive O-fronting as emphatic rather than unmarked. This could be extended to the correct prediction S-initial > V-initial > O-initial. Thus, the verb is neutral in this respect, as we see from the full distribution where V-initial, V-medial, and V-final orderings are equally preferred and unpreferred. Rather than divide the sentence into subject and predicate (figure 3), the position of verb is more likely a by-product of attention competition between subject and object.

Regarding content words, categorisation research has shown the distinction between noun and verb to be most persistent cross-linguistically (Sasse 2015). It causes a pragmatic problem in the transitive clause relating specifically to subject and object, which are both nominal phrases. We see that, in classical and generative grammar, the problem is solved by placing the second noun phrase (i.e. object) into the verb phrase (figure 3). It is thought that this structure represents a cognitive reality (Wundt 1901; Chomsky 1965).

Following the success of generative grammar, Chomsky urged brain scientists to study how such purported structures manifest themselves in brain imaging experiments. The researchers (Kluender and Kutas 1993) found no specific indication that the object is processed as if it was inside the verb phrase, and suggested that language processing is based on the interaction of syntax and semantics, rather than operating along fixed routes. It appears that the practice of placing object into a verb phrase works on paper

but not in real communication where it must be solved in a different way, that is, by marking the difference between subject and object by word order or particles.

We find the data to be compatible with the idea that the object tends to be perceived as emphatic. What is more, the effect is intensified when the relationship of O and S is highlighted by clausal fronting. OSV, highlighting OS the most, should be the most emphatic ordering. Such tension gradually eases when the verb moves to the front, changing to a positive tension—think of it in terms of rightful discrimination—as the order is flipped to SO.

So, what brings about the SO > OS hierarchy in the first place? Tomlin (1986) proposes two reasons: theme first and animacy first. These are both valid hypotheses but cannot be taken as granted.

As pointed out by Manning and Parker 1989 (contra Mallinson and Blake 1981, see section 3.2) the definition of theme or topic is not fully clear. It is nonetheless related to markedness. Alternatively, it has also been linked with the information packaging of new and given information, although this is contested by Givón (1988). It is also possible that SO ordering is particularly useful with deixis or anaphora in mind, but dedicated research is needed to allow firm conclusions.

Animacy first is also interesting, but note that first and second person markers, highly animate, are often suffixed into the verb, and this causes them to frequently appear last in sentences cross-linguistically. Such cases are omitted from transitive data, but research demonstrating a clear link is needed.

The simplest explanation for SO > OS arises directly from the logical premises. In communication consisting of intransitive and transitive clauses, subject is more frequent. It was Beauvoir (1949) who famously noted that the other is the second. Statistics gives rise to the otherness of object, and it is considered a general organising

principle of human thinking that the other comes second (Battistella 2015). This can be seen as arising from a first-person perspective. For example, a quick Internet search shows "*English and French*" to be the preferred ordering in English, but "*français et anglais*" in French. It is also important to see that this is far from being a rule. Both orderings are well attested, it is only that the one is clearly more frequent than the other.

The effect can be illustrated by imagining visiting a bookstore in New York to inquire in what languages they have Gurnah's latest novel. If they tell you it is available in French and English, the answer is fine, but it can be interpreted as bringing attention to the fact that they have it in French. If they say it is available in English and French, it seems more normal. As we see from the French example (français > anglais), such normality is not based on alphabetic order.

We have found a sustainable Level-3 explanation in markedness theory. The next section will discuss what kind of larger theoretical framework it represents.

4.4 Selecting the right theoretical framework

In the previous sub-section, an elaboration of markedness theory is proposed as closest to a null hypothesis. The task is now to find the simplest Level-4 explanation for the question what might give rise to a mechanism favouring highlighting the secondariness of object in relation to subject.

At this level, explanations are commonly sought in biology. Based on Jakobson's writings, Chomsky draws the conclusion that unmarked syntax represents a genetically determined initial language state from where the surface form is derived by transformational operations. This way of thinking is continued in Optimality Theory which likewise bases its concept on unmarked forms. (Battistella 2015)

Any claim of a grammar gene should be substantiated with evidence from genetics, but there has never been any. A recent research review (Mountford and

Newbury 2019) finds that genes do affect language learning and maintenance, but empirical evidence is limited to confirming the role of general intelligence and the lack of specific speech pathologies. There is no specific link between word order and genes.

Such lack of evidence poses a problem for Chomsky (2000, 4) who suggests grammar may have emerged from a random mutation caused by a "cosmic ray shower". To avoid Chomsky's problem, the conclusion is increasingly often made that languages are indirectly biological, and that their evolution is comparable to that of species, being based on the principle of natural selection. For example, Christiansen and Chater (2008) assert that markedness issues indicate "differential fitness across languages" which eventually results in the survival or perish of languages and linguistic structures. This approach, reminiscent of Richard Dawkins's memetics (Blackmore 2008), has come to be seen as representing a sociological perspective of language (Beckner et al. 2009) though Scott gives little consideration to it in his comprehensive review of social theories (Scott 2006).

The idea of cultural units as virus-like replicators, whether these are called memes, linguemes, or constructions, should be removed per Occam's razor. It is possible to suggest that language communities favour structures that are easier to learn and remember, but this does not indicate that languages have life-cycles like living organisms, except in a purely metaphorical sense, and metaphors are best left for educational purposes. Scientific theories should strive for accuracy, not approximation. Of course, it is also the vital task of theorising to venture beyond what is already established, but this can only be done in relation to a properly defined null hypothesis, allowing for alternative hypotheses that require support from further evidence. It is a mechanism that generates scientific progress. Being able to remove all unnecessary elements is the basic test of a scientific theory. If we allow adding just one

unsubstantiated element to any theory, we arrive at a new theory whereby we are allowed to add one more ad infinitum.

To follow scientific procedure, it suffices to say that people are known to communicate for a variety of reasons, and as they do, markedness effects emerge from the statistics of communication. It is a well-documented fact that people are capable of creating languages (Okrent 2009), and there is no indication that this would not be possible without grammar genes or selfish memes. It is of course true that if people did not have organs for receiving, processing, and producing language, we would unlikely be able to create it. However, there does not appear to be anything in these organs that causes specific grammatical structures to appear. In order for us to be able to express whatever we need to, a language must logically have the means to express predicate-argument structures, and this is the unnegotiable foundation of transitive typology.

Sustainable, non-biologistic approaches to language are found in phenomenology (Husserl [1920] 2013), building on the tradition of Western philosophy from Plato onward, and in the structuralism arising from Saussure's work in historical–comparative linguistics (Daneš 1987). Ideally, these start from the simple realisation that language appears to be a man-made solution to the communication problem, and whichever facts of it are agreed to depend on biology must be firmly established in relevant research.

5. Conclusion

The simplest explanation for transitive variation is that there is a cross-linguistic tendency to perceive objects as emphatic if they appear (i) before subject or, (ii) sentence-initially. Both cases are usually avoided in the basic transitive order. Conversely, if subject precedes object, the fronting of the two is preferred because it highlights the primacy subject over object. The place of verb is a by-effect of the need

to mark subject and object which, unlike verb, are nominal phrases. These principles give rise to the hierarchy SOV > SVO > VSO > VOS > OVS > OSV.

At the level of the distribution pattern, the attested frequency of SOV over SVO languages has been proved to be significant albeit not superior by a great deal. This is most likely to be a historical accident. It has been demonstrated that SOV is clearly the most common ordering across language families.

The universal explanation proposed in the present paper exploits markedness theory, especially as presented by Roman Jakobson (1896–1982) at the seminal Dobbs Ferry conference on linguistic universals in 1961. Based on Russian grammar, Jakobson proposed markedness as an explanation of the SO > OS preference.

The present paper makes dual use of this one principle, proposing the full explanation as a null hypothesis for transitive typology. There is no room for biological speculations in the null hypothesis because markedness can simply arise from the logical premises which materialise in the communicative situations, where subjects are statistically more frequent than objects. Object is second to subject in terms of frequency and, as a solution to the linearisation problem, it is marked as such by the basic word order. Deviation from it is most often used for topicalisation or for different semantic functions such as interrogation and relativisation.

Bibliography

Bach, Emmon W. 1974. *Syntactic Theory*. New York, NY: Holt, Rinehart and Winston.

Battistella, Edwin. 2015. “Markedness in Linguistics.” *International Encyclopedia of the Social & Behavioral Sciences* 533–37.

de Beauvoir, Simone. 1949. *Le Deuxième Sexe, I: Les Faits et Les Mythes*. Paris: Gallimard.

Beckner, Clay, Richard Blythe, Joan Bybee, Morten H. Christiansen, William Croft,

Nick C. Ellis, John Holland, Jinyun Ke, Diane Larsen-Freeman, and Tom Schoenemann. 2009. “Language Is a Complex Adaptive System: Position Paper.” *Language Learning* 59:1–26.

Bickerton, Derek. 1981. *The Roots of Language*. Ann Arbor, MI: Karoma.

Blackmore, Susan. 2008. “Memes Shape Brains Shape Memes.” *Behavioral and Brain Sciences* 31(5):513–513.

Chappell, Hilary, Li Ming, and Alain Peyraube. 2007. “Chinese Linguistics and Typology: The State of the Art.” *Linguistic Typology* 11(1).

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. *The Architecture of Language*. Oxford: Oxford University Press.

Chrisomalis, Stephen. 2009. “The Origins and Co-Evolution of Literacy and Numeracy.” Pp. 59–74 in *The Cambridge Handbook of Literacy*, edited by D. R. Olson and N. Torrance. Cambridge: Cambridge University Press.

Christiansen, Morten H., and Nick Chater. 2008. “Language as Shaped by the Brain.” *Behavioral and Brain Sciences* 31(5):489–509.

Christiansen, Morten H., and Joseph T. Devlin. 1997. “Recursive Inconsistencies Are Hard to Learn: A Connectionist Perspective on Universal Word Order Correlations.” in *Proceedings of the 19th Annual Cognitive Science Society Conference*, edited by M. Shafto and P. Langley.

Comrie, Bernard. 1981. *Language Universals and Linguistic Typology*. 1st ed. Chicago: University of Chicago Press.

Cook, Noble David. 1998. *Born to Die: Disease in the New World Conquest, 1492-1650*. Cambridge: Cambridge University Press.

Culicover, Peter W., and Kenneth Wexler. 1974. *The Invariance Principle*. 55. Irvine,

CA: University of California.

Cysouw, Michael. 2008. “Linear Order as a Predictor of Word Order Regularities.” *Advances in Complex Systems* 11(03):415–20.

Daneš, František. 1987. “On Prague School Functionalism in Linguistics.” in *Functionalism in Linguistics*, edited by R. Dirven. Amsterdam: John Benjamins.

Derbyshire, Desmond C. 1977. “Word Order Universals and the Existence of OVS Languages.” *Linguistic Inquiry* 8(3):590–99.

Derbyshire, Desmond C., and Geoffrey K. Pullum. 1981. “Object-Initial Languages.” *International Journal of American Linguistics* 47(3):192–214.

Diehl, Lon. 1975. *Space Case: Some Principles and Their Implications Concerning Linear Order in Natural Language*. 19. Bismarck, ND.

Dryer, Matthew S. 1989. “Large Linguistic Areas and Language Sampling.” *Studies in Language* 13(2):257–92.

Dryer, Matthew S. 1991. “SVO Languages and the OV:VO Typology.” *Journal of Linguistics* 27(2).

Dryer, Matthew S. 1997. “On the Six-Way Order Typology.” *Studies in Language* 21(1):69–103.

Dryer, Matthew S. 2013a. “Determining Dominant Word Order.” *The World Atlas of Language Structures Online*.

Dryer, Matthew S. 2013b. “Order of Subject, Object and Verb.” *The World Atlas of Language Structures Online*.

Dryer, Matthew S. 2013c. “Order of Subject and Verb.” *The World Atlas of Language Structures Online*.

Emonds, Joseph. 1980. “Word Order in Generative Grammar.” *Journal of Linguistic Research* 1:33–54.

Evans, Nicholas, and Stephen C. Levinson. 2009. “The Myth of Language Universals: Language Diversity and Its Importance for Cognitive Science.” *Behavioral and Brain Sciences* 32(5):429–48.

Ferrer-i-Cancho, Ramon. 2008. “Some Word Order Biases from Limited Brain Resources: A Mathematical Approach.” *Advances in Complex Systems* 11(03):393–414.

Ferrer-i-Cancho, Ramon. 2014. “Why Might SOV Be Initially Preferred and Then Lost or Recovered? A Theoretical Framework.” Pp. 66–73 in *The Evolution of Language*. WORLD SCIENTIFIC.

Ferrer-i-Cancho, Ramon. 2015. “The Placement of the Head That Minimizes Online Memory.” *Language Dynamics and Change* 5(1):114–37.

Frey, Werner. 2015. “Word Order.” Pp. 514–62 in *Syntax - Theory and Analysis, Vol. 1*, edited by T. Kiss and A. Alexiadou. Berlin, München, Boston: De Gruyter.

Gell-Mann, M., and M. Ruhlen. 2011. “The Origin and Evolution of Word Order.” *Proceedings of the National Academy of Sciences* 108(42):17290–95.

Givón, Talmy. 1988. “The Pragmatics of Word-Order: Predictability, Importance and Attention.” in *Studies in Syntactic Typology, Vol. 17*, edited by M. Hammond, E. A. Moravcsik, and J. Wirth. Amsterdam: John Benjamins.

Givón, Talmy. 2018. *On Understanding Grammar*. Revised. Amsterdam/Philadelphia: John Benjamins.

Greenberg, Joseph H. 1963. “Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements.” Pp. 73–113 in *Universals of Language*, edited by J. H. Greenberg. Cambridge, MA: MIT Press.

Hammarström, Harald. 2015. *The Basic Word Order Typology: An Exhaustive Study*.

Hammarström, Harald. 2016. “Linguistic Diversity and Language Evolution.” *Journal*

of Language Evolution 1(1):19–29.

Hawkins, John A. 1983. *Word Order Universals*. New York, NY: Academic Press.

Hawkins, John A. 1994. *A Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press.

Hawkins, John A. 2004. *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press.

Husserl, Edmund. 2013. *Logical Investigations*. translated. Abingdon-on-Thames: Routledge.

Jakobson, Roman. 1963. "Implications of Language Universals for Linguistics, Vol. II." Pp. 263–78 in *Universals of Language*, edited by J. H. Greenberg. Cambridge, MA: MIT Press.

Jouini, Kamel. 2018. "A Feature-Based Analysis of the Derivation of Word Order and Subject-Verb Agreement in Arabic Varieties." *Arab World English Journal* 9(1):233–55.

Keenan, Edward Louis, and Elinor Ochs. 1979. "Becoming a Competent Speaker of Malagasy." Pp. 113–58 in *Languages and Their Speakers*, edited by T. Shopen. Cambridge, MA: Winthroop.

Kluender, Robert, and Marta Kutas. 1993. "Subjacency as a Processing Phenomenon." *Language and Cognitive Processes* 8(4):573–633.

Koizumi, Masatoshi, and Jungho Kim. 2016. "Greater Left Inferior Frontal Activation for SVO than VOS during Sentence Comprehension in Kaqchikel." *Frontiers in Psychology* 7.

Koul, Omkar N. 2008. *Modern Hindi Grammar*. Springfield, VA: Dunwoody Press.

Krupa, Viktor. 1982. "Syntactic Typology and Linearization." *Language* 58(639–645).

LaFrance, Adrienne. 2015. "An Unusual Way of Speaking, Yoda Has." *The Atlantic*,

December 18.

Lehmann, Winfred P. 1973. “A Structural Principle of Language and Its Implications.” *Language* 49(1):47–66.

Mallinson, Graham, and Barry J. Blake. 1981. *Language Typology: Cross-Linguistic Studies in Syntax*. Amsterdam: North-Holland.

Manning, Alan D., and Frank Parker. 1989. “The SOV ... OSV Frequency Hierarchy.” *Language Sciences* 11(1):43–65.

Maurits, Luke, and Thomas L. Griffiths. 2014. “Tracing the Roots of Syntax with Bayesian Phylogenetics.” *Proceedings of the National Academy of Sciences* 111(37):13576–81.

Maurits, Luke, Amy Perfors, and Daniel Navarro. 2010. “Why Are Some Word Orders More Common than Others? A Uniform Information Density Account.” in *Advances in Neural Information Processing Systems*, edited by J. Lafferty, C. Williams, J. Shawe-Taylor, R. Zemel, and A. Culotta.

Maxwell, Daniel N. 1984. “A Typologically Based Principle of Linearization.” *Language* 60(2):251.

McCawley, James D. 1982. *Thirty Million Theories of Grammar*. London: University of Chicago Press.

Morelli, Giovanna, Yajun Song, Camila J. Mazzoni, Mark Eppinger, Philippe Roumagnac, David M. Wagner, Mirjam Feldkamp, Barica Kusecek, Amy J. Vogler, Yanjun Li, Yujun Cui, Nicholas R. Thomson, Thibaut Jombart, Raphael Leblois, Peter Lichtner, Lila Rahalison, Jeannine M. Petersen, Francois Balloux, Paul Keim, Thierry Wirth, Jacques Ravel, Ruifu Yang, Elisabeth Carniel, and Mark Achtman. 2010. “Yersinia Pestis Genome Sequencing Identifies Patterns of Global Phylogenetic Diversity.” *Nature Genetics* 42(12):1140–43.

Mountford, Hayley S., and Dianne F. Newbury. 2019. “The Genetics of Language Acquisition.” Pp. 33–50 in *International Handbook of Language Acquisition*. New York, NY: Routledge, [2019] | Series: Routledge: Routledge.

Nettle, Daniel. 1999. *Linguistic Diversity*. Oxford: Oxford University Press.

Newmeyer, Frederick J. 2017. “Form and Function in the Evolution of Grammar.” *Cognitive Science* 41:259–76.

Norde, Muriel. 2009. “Unidirectionality.” Pp. 48–105 in *Degrammaticalization*. Oxford University Press.

Okrent, Arika. 2009. *In the Land of Invented Languages*. New York, NY: Random House.

Peeters, Bert. 1991. “Basic Word Order Frequencies or Manning/Parker Contra Tomlin.” *Language Sciences* 13(1):79–88.

Perreault, Charles, and Sarah Mathew. 2012. “Dating the Origin of Language Using Phonemic Diversity” edited by M. D. Petraglia. *PLoS ONE* 7(4):e35289.

Polinskaja, Maria S. 1989. “Object Initiality: OSV.” *Linguistics* 27(2).

Pullum, Geoffrey K. 1977. “Word Order Universals and Grammatical Relations.” Pp. 249–77 in *Syntax and Semantics, Vol. 8*, edited by C. P. and S. J. New York, NY: Academic Press.

Sasse, Hans-Jürgen. 2015. “Syntactic Categories and Subcategories.” Pp. 158–217 in *Syntax - Theory and Analysis. An International Handbook Vol. 1*. Berlin, München, Boston: DE GRUYTER.

Schwartz, Arthur. 1974. “The VP-Constituent of SVO Languages.” Pp. 619–37 in *Proceedings of the 11th International Congress of Linguistics*, edited by L. Heilmann. Bologna: Il Mulino.

Scott, John. 2006. *Social Theory. Central Issues in Sociology*. London: Sage.

Seuren, Pieter A. M. 1998. “Western Linguistics: An Historical Introduction.” Hoboken, NJ: Blackwell.

Sinnemäki, Kaius. 2014. “Cognitive Processing, Language Typology, and Variation.” *Wiley Interdisciplinary Reviews: Cognitive Science* 5(4):477–87.

Sinnemäki, Kaius, and Francesca Di Garbo. 2018. “Language Structures May Adapt to the Sociolinguistic Environment, but It Matters What and How You Count: A Typological Study of Verbal and Nominal Complexity.” *Frontiers in Psychology* 9.

Song, Jae Jung. 1991. “On Tomlin, and Manning and Parker on Basic Word Order.” *Language Sciences* 13(1):89–97.

Song, Jae Jung. 2012. *Word Order*. Cambridge University Press.

Tomlin, Russell S. 1986. *Basic Word Order: Functional Principles*. London: Croom Helm.

Trudgill, Peter. 2011. *Sociolinguistic Typology*. Oxford: Oxford University Press.

Ultan, Russell. 1969. “Some General Characteristics of Interrogative Systems.”

Vennemann, Theo. 1974. “Topics, Subjects, and Word Order: From SXV to SVX via TXV.” Pp. 339–76 in *Historical Linguistics I: Syntax, Morphology, Internal and Comparative Reconstruction*. Amsterdam: North-Holland Publishing.

Weir, E. M. Helen. 1994. “Nadëb.” Pp. 291–323 in *Typological Studies in Negation*, edited by P. Kahrel and R. van den Berg. Amsterdam: John Benjamins.

Wundt, Wilhelm. 1901. *Sprachgeschichte Und Sprachpsychologie*. Leipzig: W. Engelmann.