Explanation in typology

Diachronic sources, functional motivations and the nature of the evidence

Edited by

Karsten Schmidtke-Bode

Natalia Levshina

Susanne Maria Michaelis

Ilja Seržant

Conceptual Foundations of Language Science

language science press

Conceptual Foundations of Language Science

No scientific work proceeds without conceptual foundations. In language science, our concepts about language determine our assumptions, direct our attention, and guide our hypotheses and our reasoning. Only with clarity about conceptual foundations can we pose coherent research questions, design critical experiments, and collect crucial data. This series publishes short and accessible books that explore well-defined topics in the conceptual foundations of language science. The series provides a venue for conceptual arguments and explorations that do not require the traditional book-length treatment, yet that demand more space than a typical journal article allows.

In this series:

1. N. J. Enfield. *Natural causes of language*.

ISSN: 2363-877X

Explanation in typology

Diachronic sources, functional motivations and the nature of the evidence

Edited by

Karsten Schmidtke-Bode

Natalia Levshina

Susanne Maria Michaelis

Ilja Seržant

language
science
press

Karsten Schmidtke-Bode , Natalia Levshina , Susanne Maria Michaelis & Ilja Seržant (eds.). 2018. *Explanation in typology: Diachronic sources, functional motivations and the nature of the evidence* (Conceptual Foundations of Language Science). Berlin: Language Science Press.

This title can be downloaded at:
http://langsci-press.org/catalog/book/000

ISBN: no digital ISBN
no print ISBNs!
ISSN: 2363-877X
no DOI
ID not assigned!

Cover and concept of design: Ulrike Harbort
Fonts: Linux Libertine, Libertinus Math, Arimo, DejaVu Sans Mono
Typesetting software: XeLaTeX

Language Science Press
Unter den Linden 6
10099 Berlin, Germany
langsci-press.org

Storage and cataloguing done by FU Berlin

no logo

Contents

Introduction

Karsten Schmidtke-Bode
Leipzig University

The present volume addresses a foundational issue in linguistic typology and language science more generally. It concerns the kinds of explanation that typologists provide for the cross-linguistic generalizations they uncover, i.e. for so-called universals of language. The universals at issue here are usually probabilistic statements about the distribution of specific structures, such as the classic Greenbergian generalizations about word order and morphological markedness patterns. Some examples are given in (1)–(4) below:

(1) With overwhelmingly greater than chance frequency, languages with normal SOV order are postpositional. (**Greenberg1963**)

(2) A language never has more gender categories in nonsingular numbers than in the singular. (**Greenberg1963**)

(3) If a language uses an overt inflection for the singular, then it also uses an overt inflection for the plural. (**Croft2003** 89, based on **Greenberg1966** 28)

(4) In their historical evolution, languages are more likely to maintain and develop non-ergative case-marking systems (treating S and A alike) than ergative case-marking systems (splitting S and A). (**BickelEtAl2015** 5)

As can be seen from these examples, cross-linguistic generalizations of this kind may be formulated in terms of preferred types in synchronic samples or in terms of higher transition probabilities for these types in diachronic change (see also **Greenberg1978**; **Maslova2000**; **Cysouw2011**; **Bickel2013** for discussion of the latter approach). But this is, strictly speaking, independent of the question we are primarily concerned with here, namely how to best account for such generalizations once they have been established.

Karsten Schmidtke-Bode. Introduction. In Karsten Schmidtke-Bode, Natalia Levshina, Susanne Maria Michaelis & Ilja Seržant (eds.), *Explanation in typology: Diachronic sources, functional motivations and the nature of the evidence*, iii–ix. Berlin: Language Science Press. DOI:??

The most widespread typological approach to explanation is grounded in functional properties of the preferred structural types: For example, typical correlations in the ordering of different types of phrases (e.g. object–verb and NP–postposition) have been argued to allow efficient online processing (e.g. **Hawkins1994** 2004). Markedness patterns in morphology (e.g. the distribution of zero expression in case, number or person systems) have been attributed to economy, i.e. the desire to leave the most frequent and hence most predictable constellations unexpressed, or rather to a competition between economy and the motivation to code all semantic distinctions explicitly (e.g. **Haiman1983**; **Comrie1989**; **Aissen2003**; **Croft2003**; **Haspelmath2008** among many others). The general idea behind this approach is thus that speech communities around the world are subject to the same kinds of cognitive and communicative pressures, and that the languages they speak tend to develop structures that respond to these pressures accordingly, or, as **Bickel2014** puts it, "in such a way as to fit into the natural and social eco-system of speakers: that they are easy to process, that they map easily to patterns in nonlinguistic cognition, and that they match the social and communicative needs of speakers."

There is a clear parallel to evolutionary biology here, in that languages are said to *converge* on similar structural solutions under the same functional pressures, just like unrelated species tend to develop similar morphological shapes in order to be optimally adapted to the specific environment they co-inhabit (**Deacon1997**; **Caldwell2008**; **EvansLevinson2009**; **Givón2010**). When applied to language, this line of explanation at least implicitly invokes what is known as "attractor states", i.e. patterns of structural organization that languages are drawn into in their course of development.[1] For this reason, one could also speak of a **result-oriented** approach to explanation.

There is, however, another way of looking at the same patterns, one that redirects attention from the functional properties to the diachronic origins of the linguistic structures in question. On this view, many universal tendencies of order and coding are seen as by-products, as it were, of recurrent processes of morphosyntactic change, notably grammaticalization, but without being adaptive in the above sense: There is no principled convergence on similar structural traits because these traits might be beneficial from the perspective of processing, iconicity or economical communicative behaviour. Instead, the current

[1] The term attractor state (or basin of attraction) is adopted from the theory of complex dynamic systems (e.g. **Cooper1999**; **HoweLewis2005**; **Holland2006**), which has become increasingly popular as a way of viewing linguistic systems as well (see **BecknerEtAl2009** and **Port2009** for general overviews, and **Haig2018** or **Nichols2018** for very recent applications to typological data).

synchronic distributions are argued to be long-term reflections of individual diachronic trajectories, in particular the diachronic sources from which the structures in question originate. **Givón1984** and **Aristar1991** for example, suggested that certain word-order correlations may simply be a consequence of a given ordering pair (e.g. Gen–N & Rel–N, or V–O & Aux–V) being directly related diachronically: Auxiliaries normally grammaticalize from main verbs that take other verbs as complements, and since these complements follow the verb in VO languages, they also follow the auxiliary in the resulting Aux–V construction; the mirror-image pattern holds for OV languages (see also **Lehmann1986** 12–13). If this line of reasoning extends to most other word-order pairs, there is no need to motivate the synchronic correlations in functional-adaptive terms, e.g. by saying that the correlations arise *in order to* facilitate efficient sentence processing.

In the domain of morphology, **Garrett1990** argued that patterns in case marking, specifically of differential ergative marking, are exhaustively explained by the properties of the source of the ergative marker: When ergative case arises from the reanalysis of instrumental case, the original characteristics of the latter, such as a restriction to inanimate referents, are directly bequeathed to the former. The result is a pattern in which animate A-arguments are left unmarked, but since this is a direct "persistence effect" (**Hopper1991**) of the history of the ergative marker, there is again no need for an additional functional-adaptive explanation in terms of other principles, such as a drive for economical coding patterns. Rather than being result-oriented, then, this way of explaining universals can be characterized as **source-oriented**.

Such source-oriented explanations thus move away from attractor states of grammatical organization and often emphasize the importance of "attractor trajectories" instead (**BybeeBeckner2015** 185): In some domains of grammar, the patterns of reanalysis and ensuing grammaticalization are so strikingly similar across the world's languages that it is not surprising that they yield similar outcomes, such as strong correlations between V–O & Aux–V or V–O & P–NP ordering. In other cases, it is argued that many individual, and partly very different, diachronies are capable of producing a uniform result, but without any consistent functional force driving these trajectories. **Cristofaro2017** for instance, claims that this is the case for plural markers: An initial system without number marking can develop an overt plural morpheme from many different sources – usually by contextual reanalysis – and thus ultimately come to contrast a zero singular with an overt plural, but these developments are neither triggered nor further orchestrated by a need for economical coding: They do not happen to keep the (generally more frequent) singular unmarked and the (generally less frequent)

plural overtly signalled.

In other words, whether the individual diachronic trajectories are highly similar or rather diverse, the premise of the source-oriented approach is that they can scale up to produce a predominant structural pattern in synchronic samples. Hence they obviate the need for highly general functional principles tying these patterns together.

While the source-oriented approach was still a more marginal position in previous volumes on explaining language universals (e.g. **Hawkins1988a**; **Good2008**), it has gained considerable ground over the last decade, notably in a series of articles by Cristofaro (e.g. **Cristofaro2012** 2014; 2017) but also in other publications (e.g. **Anderson2016**; **Creissels2008**; **GildeaZúñiga2016**). Moreover, while the basic thrust of the two explanatory approaches is straightforward, clarification is needed on a number of – equally fundamental – details. After all, both approaches are functionalist in nature, as they rely on domain-general mechanisms (**Bybee2010**) to explain the emergence of language structure and linguistic universals; and in both approaches, these mechanisms constrain how languages "evolve into the variation states to which implicational and distributional universals refer" (**Hawkins1988b**). But as **Plank2007** notes, "what is supposed to be the essence and force of diachronic constraints would merit livelier discussion." It is the goal of the present book to offer precisely a discussion of this kind.

The volume begins with a programmatic paper by **Martin Haspelmath** on what it means to explain a universal in diachronic terms. He aims to clarify how diachrony is involved in result-oriented and source-oriented accounts, respectively, and thus lays out a general conceptual framework for the explanation of universals. At the same time, Haspelmath opens the floor for debating the strengths and weaknesses of the two explanatory accounts at issue here. His own position is that, in many cases, current source-oriented explanations are ill-equipped to truly explain the phenomena they intend to account for, and hence cannot replace result-oriented motivations. Haspelmath's arguments for this position, as well as his terminological proposals, provide a frame of reference to which all other contributions respond in one way or another.

The lead article is followed by two endorsements of source-oriented explanations, articulated by **Sonia Cristofaro** and **Jeremy Collins**, respectively. They both describe the approach in widely accessible terms, allowing also readers outside of linguistic typology to appreciate the general argument as well as the specific examples discussed. The phenomena themselves involve domains that are particularly well-known for being explained in functional-adaptive terms, namely differential argument marking, number marking and word-order corre-

lations, and these are all argued to be best captured by persistence effects from their respective diachronic origins.

We then proceed to papers that allow for progressively more room for functional-adaptive motivations and, importantly, for methodological discussions on how to obtain evidence for such pressures. Accordingly, all of these papers adduce novel empirical data and discuss them in light of the present debate.

Matthew Dryer's paper is an immediate follow-up on Collins' discussion of word-order correlations. On the one hand, Dryer argues that the various correlates of adposition–noun ordering (e.g. OV and NP–P, and Gen–N and NP–P) are, indeed, best accounted for in source-oriented terms. In particular, only this approach proves capable of explaining the occurrence (and the individual semantic types) of both prepositions and postpositions in SVO languages. On the other hand, however, Dryer contends that there are some significant correlations for which a source-based account either fails to offer an explanation or else makes the opposite prediction of the patterns we find synchronically. Dryer concludes, therefore, that neither a purely source-based nor a purely result-based explanation is sufficient to deal with word-order correlations.

In a similar fashion as Dryer's paper, **Holger Diessel**'s article demonstrates that different aspects of the same grammatical domain – in this case adverbial clause combinations – are amenable to different types of explanation. Diessel focuses specifically on the structure and development of preposed adverbial clauses and argues that some of their typological characteristics, notably the properties of their subordinating morphemes, receive a satisfactory explanation in terms of the respective source construction(s), thereby supplanting earlier processing-based explanations. On the other hand, he proposes that the position of adverbial constructions (in general) is clearly subject to a number of functional-adaptive pressures, and that these may already have affected the diachronic sources from which the current preposed adverbial clauses have grammaticalized.

Karsten Schmidtke-Bode offers a review of Hawkins' (2004, 2014) research programme of "processing typology", examining the plausibility of Hawkins' functional-adaptive ideas in diachronic perspective. On a theoretical level, it is argued that a predilection for efficient information processing is operative mostly at the diffusion stage of language change, regardless of the source from which the respective constructions originate. On a methodological level, the paper proposes that the cross-linguistic predictions of Hawkins' programme can be tested more rigorously than hitherto by combining static and dynamic statistical models of large typological data sets; this is demonstrated in a case study on the distribution of article morphemes in VO- and OV-languages, respectively.

An important methodological point is also made by **Ilja A. Seržant**, who claims that certain functional-adaptive pressures may not actually surface in standard typological analysis because they are weak forces, clearly at work but also easily overridden by other, language-specific factors. Because of their weak nature, they may not be directly visible any-more in a synchronic type, but they can be detected in qualitative data from transition phases. Based on diachronic data from Russian, Seržant shows how the development of differential object marking was crucially influenced by considerations of ambiguity avoidance (and hence a classic functional-adaptive motivation), over and above the constraints inherited from the source construction. In the absence of such longitudinal data, transition phases can be identified on the basis of syn-chronic variability, and Seržant shows that a wide variety of languages currently exhibit variation in differential object marking that mirrors the diachronic findings from Russian, and that is not predictable from the source meaning of the marker in question.

Susanne Maria Michaelis adds another source of data to the debate at hand. She argues that creole languages provide a unique window onto the relationship between synchronic grammatical patterns and their diachronic trajectories, as the latter are often relatively recent and also accelerated when compared to normal rates of grammatical change. The developments are, consequently, more directly accessible and less opaque than in many other cases. By inspecting creole data on possessive forms in attributive and referential function (e.g. *your* versus *yours*), Michaelis finds evidence for the development of the same kinds of coding asymmetries that this domain offers in non-contact languages around the world. She proposes that the data are indicative of result-oriented forces that drive diverse diachronic pathways towards the same synchronic outcome. This stance contrasts most explicitly with Cristofaro's, who interprets such situations in exactly the opposite way (i.e. as providing evidence *against* a unifying functional explanation).

Natalia Levshina, finally, adopts an entirely different methodological approach to illuminate the present discussion: In her paper, she showcases the paradigm of artificial language learning, which can be employed to inspect whether users of such newly acquired languages develop performance biases that are in keeping with hypothesized functional principles, such as an increasingly efficient distribution of morphological marking. Her case study clearly demonstrates such biases and discusses where they may ultimately come from, i.e. how they fit into the new conceptual framework of constraints offered by Haspelmath's position paper.

The volume is rounded off by a brief **epilogue** in which **Karsten Schmidtke-**

Bode and **Eitan Grossman** summarize and further contextualize the arguments put forward by the contributors.

Overall, the purpose of the present book is to provide a state-of-the-art overview of the general tension between source- and result-oriented explanations in linguistic typology, and specifically of the kinds of arguments and data sources that are (or can be) brought to bear on the issue. It should be made clear from the outset that the two types of explanation are framed as antagonistic here even though in most cases, an element of both will be needed in order to fully account for a given grammatical domain. As we emphasize in the epilogue, the diachronic source of a grammatical construction certainly constrains its further development, but the major issue at stake here is the extent to which result-oriented, functional-adaptive motivations enter these developments as well. By the end of the day, universals of language structure will thus differ in the *degree* to which they are shaped by such adaptive pressures.

Acknowledgements

The present volume originated in the context of the project *Form-frequency correspondences in grammar* at Leipzig University. The support of the European Research Council (ERC Advanced Grant 670985, Grammatical Universals) is gratefully acknowledged. An oral precursor to this volume was a workshop on the topic at the 49th Annual Conference of the Societas Linguistica Europaea in Naples in 2016, co-organized by the editors of this book. We would like to thank the participants and the audience of that workshop for insightful contributions and discussion. We would also like to thank Eitan Grossman and Mark Dingemanse for extensive feedback on all papers in the present volume. Finally, we are grateful to Sebastian Nordhoff and his team at Language Science Press as well as the participants of Language Science Press's community proofreading.

Chapter 1

Grammaticalization accounts of word order correlations

Matthew S. Dryer
University at Buffalo

This paper examines the role that grammaticalization plays in explaining word order correlations. It presents some data that only grammaticalization accounts for, but also argues that there are correlations that grammaticalization does not account for. The conclusion is that accounts entirely in terms of grammaticalization or accounts that make no reference to grammaticalization are both inadequate.

1 Introduction

There is extensive literature both on identifying word order correlations (**Greenberg1963 Hawkins1983 Dryer1992**) and on possible explanations for these correlations. Proposed explanations can be grouped loosely into three types. First, it is proposed that some correlations exist because of some sort of similarity or shared property of the pairs that correlate. An example of this is the hypothesis that the order of object and verb correlates with the order of adposition and noun phrase because both involve a pair of head and dependent. A second type of explanation is in terms of sentence processing (**Kuno1974 Dryer1992** 2009; **Hawkins1994** 2004, 2014), under which the types that do not conform to the correlations are less frequent because structures containing the two inconsistent types are more difficult to parse. This would be what Haspelmath (2018 [this volume]) calls a functional-adaptive type of explanation. A third line of explanation is in terms of grammaticalization (**Givón1979 HeineReh1984** 241–244; **Bybee1988 Aristar1991** De**Lancey1994 Collins2012** 2018 [this volume]). For example it is hypothesized that the reason (or a reason) why the order of adposition and noun phrase correlates with the order of verb and object is that one grammaticalization source for

Matthew S. Dryer. Grammaticalization accounts of word order correlations. In Karsten Schmidtke-Bode, Natalia Levshina, Susanne Maria Michaelis & Ilja Seržant (eds.), *Explanation in typology: Diachronic sources, functional motivations and the nature of the evidence*, 1–29. Berlin: Language Science Press. DOI:??

adpositions is verbs and the order or verb and object remains the same when the verb grammaticalizes as an adposition. This line of explanation is thus crucially based on the diachronic sources of adpositions and hence a type of source-based explanation (**Cristofaro2018** [this volume]).

Despite these competing hypotheses for explaining word order correlations, there is surprisingly little attempt by proponents of an explanation in terms of grammaticalization to argue against other approaches or by proponents of other approaches to argue against grammaticalization. In fact, proponents of other approaches rarely even mention the possible role of grammaticalization. The goal of this paper is to argue that both explanations in terms of grammaticalization and explanations in terms of shared features or processing are needed in explaining word order correlations. I will focus on the pros and cons of grammaticalization explanations, largely ignoring the difference between accounts in terms of shared features and accounts in terms of sentence processing.[1]

In §2, I discuss explanations for correlations involving order of adposition and noun phrase and discuss evidence that only a grammaticalization approach can account for. Namely I examine SVO& GenN languages that have both prepositions and postpositions and show that not only does an approach involving grammaticalization predict languages with both prepositions and postpositions but also correctly predicts the semantics associated with each type of adposition. In §3, I give reasons why grammaticalization cannot account for all word order correlations, concluding that grammaticalization and other factors conspire to account for some correlations. And in §4, I discuss data involving word order properties of definiteness markers where grammaticalization seems to make the wrong prediction.

2 A grammaticalization account of the correlations with the order of adposition and noun phrase

In this section, I present evidence for a grammaticalization account for correlations involving the order of adposition and noun phrase that only grammaticalization can account for. In §2.1, I discuss evidence of adpositions grammaticalizing from verbs. In §2.2, I discuss evidence of a second grammaticalization source for adpositions, namely head nouns in genitive constructions. The next section, §2.3, is, in my view, the most important section of this paper. In that section, I dis-

[1] In **Dryer1992** I argue for a processing account over an account in terms of heads and dependents for the various correlations between pairs of elements and the order of verb and object.

cuss SVO languages which employ GenN word order. Grammaticalization theory predicts that in such languages, if adpositions arise from both grammaticalization sources discussed in §2.1 and §2.2, the language will have both prepositions and postpositions, the former arising from verbs, the latter from head nouns in genitive constructions, with particular semantics associated with each. I present evidence from a number of languages showing that this prediction is borne out.

2.1 Adpositions that grammaticalize from verbs

Let me turn now to one of the best-known word order correlations, between the order of object and verb and the order of adposition and noun phrase, where VO languages tend to have prepositions while OV languages tend to have postpositions (**Greenberg1963 Dryer1992**). Evidence for this correlation is given in Tables 1 and 2. The data for VO languages is given in Table 1. The numbers in Table 1 denote numbers of genera containing languages of the given sort, divided into five large continental areas (**Dryer1989**). The more frequent type in each area is enclosed in square brackets.

Table 1: Order of adposition and noun phrase in VO languages

	Africa	**Euras**	**Oceania**	**N.Amer**	**S.Amer**	**TOTAL**
VO& Po	10	6	3	5	15	39
VO& Pr	[37]	[25]	[48]	[27]	15	152

Table 1 shows that prepositions outnumber postpositions in VO languages by a wide margin in four of the five areas, with the fifth area (South America) having an equal number of genera containing languages with prepositions and those containing languages with postpositions. Overall, VO& Pr outnumbers VO& Po by 152 to 39 genera.

The corresponding data for OV languages is given in Table 2.

Table 2: Order of adposition and noun phrase in OV languages

	Africa	**Euras**	**Oceania**	**N.Amer**	**S.Amer**	**TOTAL**
OV& Po	[25]	[45]	[82]	[28]	[41]	221
OV& Pr	3	2	5	0	1	11

Table 2 shows a stronger preference for postpositions in OV languages than the preference for prepositions in VO languages shown in Table 1, in that Table 2 shows only 11 genera containing OV& Pr languages while Table 1 shows 39 genera containing VO& Po languages. I discuss an explanation for this difference in terms of grammaticalization in §3 below.

An explanation of this correlation in terms of grammaticalization appeals to the fact that verbs are a common source for adpositions, so that when a verb is grammaticalized as an adposition, the order with the verb followed by object is retained as preposition followed by noun phrase, while the order with the object followed by verb is retained as noun phrase plus postposition.

Two examples of this process of grammaticalization in English are the prepositions *including* and *concerning*, as in (1).

(1) English

a. Four men, including John, arrived.

b. I will talk to you later concerning your thesis.

Both of these prepositions retain the present participle form ending in *-ing*, coming from the verbs *include* and *concern*.

Grammaticalization of adpositions from verbs is common in many other languages and widely described in the literature. The examples in (2) to (7) illustrate apparent examples of grammaticalization of particular semantic types of adpositions from verbs with particular meanings.

should these be headings of some sort?

give → for

(2) Efik (Niger-Congo: Nigeria; **Givón2001** 163)
nam utom emi **ni mi**.
do work this **give me**
'Do this work **for me**.'

give → to (marking addressee)

(3) Yoruba (Niger-Congo; Nigeria; **Givón2001** 163)
mo sọ fún ọ.
I said give you
'I said **to you**.'

go → to (marking goal of motion)

(4) Nupe (Niger-Congo: Nigeria; **Givón1979** 221)
ū bīcī lō dzūká.
he ran go market
'He ran to the market.'

follow → with (comitative)

(5) Mandarin Chinese (Sinitic: China; **LiThompson1981** 423)
tā bu **gēn** **wǒ** jiǎng-hua.
3SG NEG **follow 1SG** speak-speech
'He talks **with me**.'

take → object case marker

(6) Yatye (Niger-Congo: Nigeria; **Givón2001** 163)
ìywi **awá** **utsì** ikù.
boy took door shut
'The boy shut **the door**.'

be at → at

(7) Mandarin Chinese (Sino-Tibetan, Sinitic: China; Yu Li, p.c.)
tā **zài** **guō-li** chǎo fàn.
3SG **be.at pot-in** fry rice
'He is frying rice **in the pot**.'

The grammaticalization of adpositions from verbs provides a possible basis for an explanation of the correlation between the order of verb and object and the order of adposition and noun phrase.

2.2 Adpositions that grammaticalize from head nouns in nominal possessive constructions

Another common grammaticalization source for adpositions (and probably the more common source) is head nouns in genitive constructions. English has a number of prepositions that have arisen from head nouns in genitive constructions, including those in (8).

(8) English
a. in the side of NP → inside NP

b. by the side of NP ⟶ beside NP

c. by the cause of NP ⟶ because of NP

Because these adpositions arose from head nouns in a genitive construction with NGen order, they ended up as prepositions rather than postpositions. The opposite situation arose in Amharic, where the examples in (9) illustrate two postpositions arising from head nouns in a GenN construction.

(9) Amharic (Afro-Asiatic, Semitic: Ethiopia; **Givón1971** 399)

a. NP + bottom ⟶ NP + under
kä-bet tač allä.
at-house bottom is
'He is **under the house**.'

b. NP + reason ⟶ NP + because of
bä-issu mikniyat näw.
at-he reason is
'It is **because of him**.'

This type of grammaticalization of adpositions from head nouns in genitive constructions would explain the correlation between the order of noun and genitive and the order of adposition and noun phrase. The data in Tables 3 and 4 provides evidence for this correlation. The data in Table 3 shows GenN languages being overwhelmingly postpositional, while the data in Table 4 shows NGen languages being overwhelmingly prepositional.

Table 3: Order of adposition and noun phrase in GenN languages

	Africa	**Euras**	**Oceania**	**N.Amer**	**S.Amer**	**TOTAL**
GenN& Po	[31]	[50]	[73]	[33]	[54]	241
GenN& Pr	2	7	13	3	4	29

2.3 An interesting prediction of grammaticalization accounts for adpositions

Sections §2.1 and §2.2 illustrate two grammaticalization sources for adpositions, one from verbs, the other from head nouns in genitive constructions. In languages which are VO and NGen, both sources will lead to prepositions rather

Table 4: Order of adposition and noun phrase in NGen languages

	Africa	Euras	Oceania	N.Amer	S.Amer	TOTAL
NGen& Po	7	0	6	0	1	14
NGen& Pr	[34]	[19]	[29]	[19]	[10]	111

than postpositions. Conversely, in languages which are OV and GenN, both sources will lead to postpositions. But there are many languages which are VO but GenN. Dryer (1997, 2013) shows that although OV languages tend to be GenN and verb-initial languages tend to be NGen, both orders of noun and genitive are common among SVO languages, as shown in Table 5.

Table 5: Order of genitive and noun in SVO languages

	Africa	Euras	Oceania	N.Amer	S.Amer	TOTAL
SVO& GenN	11	11	13	2	[11]	48
SVO& NGen	[41]	[16]	13	[5]	3	78

Table 5 shows that NGen is more common overall than GenN among SVO languages by 78 genera to 48. However, the higher number of genera containing SVO& NGen languages turns out to be entirely due to languages in Africa. Outside Africa, SVO& NGen and SVO& GenN are both found in exactly 37 genera. The general conclusion is that there is no evidence of any preference for NGen order over GenN order among SVO languages.

Because of the two grammaticalization sources for adpositions described in the two preceding sections, grammaticalization theory makes an interesting prediction about SVO& GenN languages. Namely, if adpositions arise in any such languages from both grammaticalization sources, the language should have both prepositions and postpositions, those arising from verbs being prepositions and those arising from nouns being postpositions. The evidence in this section shows that this prediction is borne out.

In fact, grammaticalization theory makes more specific predictions about what meanings will be associated with prepositions and what meanings will be associated with postpositions in such languages. The lefthand column in Table 6 summarizes typical meanings associated adpositions that arise from verbs, while the righthand column summarizes typical meanings associated with adpositions that

arise from head nouns in genitive constructions.

Table 6: Typical meanings associated with adpositions

do we need 'a.' and 'b.' (all tables)?

a. **Typical meanings associated with adpositions that come from verbs:**	b. **Typical meanings associated with adpositions that come from nouns:**
benefactive ('for') instrumental ('with') comitative ('with') similative ('like') allative ('to, toward') general locations ('at') adpositions marking direct objects	specific locations like 'under', 'behind', 'in front of' 'because of'

Note that it is typically adpositions denoting specific locations that arise from nouns. Adpositions denoting general locations (meaning something like 'at') often arise from verbs. Similarly adpositions associated with motion away from or towards a location also more often arise from verbs. Grammaticalization theory predicts that in an SVO& GenN language with both prepositions and postpositions, the prepositions will tend to have meanings like those in the lefthand column in Table 6, while the postpositions will tend to have meanings like those in the righthand column. This section shows that these predictions are also borne out.

The first language illustrating how these predictions are borne out is N|uuki. The SVO order of N|uuki is illustrated in (10), GenN order in (11).

(10) N|uuki (Tuu: South Africa; **CollinsNamaseb2011** 10)
ǂharuxu ke ãi ʘoe.
Haruxu DECL eat meat
'Haruxu is eating meat.'

(11) N|uuki (Tuu: South Africa; **CollinsNamaseb2011** 37)
siso ŋǂona
Siso knife
'Siso's knife'

N|uuki has both prepositions and postpositions. Examples illustrating a preposition *ŋ|a* are given in (12) and (13), (12) illustrating an instrumental use, (13) a comitative use.

(12) N|uuki (Tuu: South Africa; **CollinsNamaseb2011** 25)
n-a si |aa ʘoe **ŋ|a ŋǂona**.
1SG-DECL IRR cut meat **with knife**
'I will cut the meat **with a knife**.'

(13) N|uuki (Tuu: South Africa; **CollinsNamaseb2011** 25)
|a|aʕe ke sĩĩsen **ŋ|a ŋ|aŋgusi**.
|a|aʕe DECL work **with N|aŋgusi**
'|a|aˀe works **with N|aŋgusi**.'

In contrast, example (14) illustrates a postposition *xuu* 'in front of'.

(14) N|uuki (Tuu: South Africa; **CollinsNamaseb2011** 80)
|x'esi |ʔaa sũi **|oβa xuu**
necklace go sit.down **child front**
'The necklace fell **in front of the child**.'

The prepositions and postpositions in N|uuki (**CollinsNamaseb2011** 24–25) are listed in Table 7.

Table 7: Prepositions and postpositions in N|uuki

	Prepositions			Postpositions	
a.	ŋ|a	'instrumental, comitative'	b.	||ãʔẽ	'in'
	||a	'like'		xuu	'in front of'
	ŋ	'linker'		tsʔĩi	'behind'
				|qhaa	'next to'

Two of the three prepositions, *ŋ|a* 'instrumental, comitative' and *||a* 'like', conform to the semantic types of adpositions arising from verbs and the fact that they are prepositions rather than postpositions can be explained if they have arisen from verbs in a VO language. And all four of the postpositions represent specific locations, conforming to what we expect semantically of adpositions arising from head nouns in genitive constructions; the fact that they are postpositions rather

than prepositions can be explained in they have arisen from head nouns in a genitive construction in a GenN language.

A second example is provided by Logba. Like N|uuki, Logba is SVO, as illustrated in (15), and GenN, as in (16).

(15) Logba (Niger-Congo: Kwa: Ghana; **Dorvlo2008** 105)
Setor ó-kpe i-gbeḍi=é.
Setor SG-peel NC-cassava=DET
'Setor peeled the cassava.'

(16) Logba (Niger-Congo: Kwa: Ghana; **Dorvlo2008** 71)
Kɔdzo a-klɔ=a
Kɔdzo NC-goat=DET
'Kɔdzo's goat'

Also like N|uuki, Logba has both prepositions and postpositions. A preposition *kpɛ* with instrumental or comitative meaning is illustrated in (17).

(17) Logba (Niger-Congo: Kwa: Ghana; **Dorvlo2008** 96)
Udzi=é ó-glɛ uzugbo **kpɛ a-futa**.
woman=DET SG-tie head **with NC-cloth**
'The lady tied her head **with a cloth**.'

In contrast, an example illustrating a postposition *etsi* 'under' is given in (18).

(18) Logba (Niger-Congo: Kwa: Ghana; **Dorvlo2008** 98)
i-datɔ=a í-tsi a-fúta=á etsi.
NC-spoon=DET SG-be.in **NC-cloth=DET under**
'The spoon is **under the cloth**.'

In Table 8 is a list of the prepositions and postpositions of Logba (**Dorvlo2008** 95, 98).

While one of the prepositions has a meaning more commonly associated with adpositions that arise from nouns (*na* 'on'), the other prepositions all have meanings that grammaticalization theory predicts for adpositions arising from verbs and all the postpositions have meanings involving specific locations, the types of meanings that grammaticalization predicts for adpositions that arise from nouns.

The third SVO& GenN language with both prepositions and postpositions is Eastern Kayah Li, a Karenic language in the Sino-Tibetan family spoken in Myanmar and Thailand. The prepositions and postpositions of Eastern Kayah Li are listed in Table 9 (**Solnit1997** 209–214).

Table 8: Prepositions and postpositions of Logba

	Prepositions			Postpositions	
a.	fɛ	'at'	b.	nu	'inside'
	na	'on'		etsi	'under'
	kpɛ	'instrumental, comitative'		tsú	'on'
	gu	'about'		ité	'in front of'
	dzígu	'from'		zugbó	'on'
				yó	'surface contact'(e.g. on a wall)
				anú	'at tip of, at edge of'
				otsoe	'on the side of'
				amá	'behind'

Table 9: Prepositions and postpositions of Eastern Kayah Li

	Prepositions			Postpositions	
a.	dɤ	'at'	b.	kū	'inside'
	mú	'at'		klɔ	'outside'
	bɤ	'at'		khu	'on, above'
	bá	'as much as'		kɛ ~ kɛdē	'down inside'
	tí	'as big as'		khʌ	'at apex of'
	tɤ ~ thɤ	'as long as'		lē	'under, downhill from'
	phú ~ hú	'like'		chá	'near'
				ŋē ~ béseŋē	'in front of'
				khjā ~ békhjā	'behind'
				lo	'on non-horizontal surface'
				klē	'in (an area)'
				rɔklē	'beside'
				ple ~ ple kū	'in narrow space between'
				cɔkū	'in middle of, between'
				thɯ	'on edge of'
				təkjā	'in the direction of'

Apart from three prepositions with unusual meanings ('as much as', 'as big as', 'as long as'), the rest of the prepositions and all of the postpositions have meanings conforming to the semantics typically associated with adpositions arising from verbs and adpositions arising from head nouns in genitive constructions respectively.

The fourth SVO& GenN language exhibiting a similar pattern is Jabem, an Oceanic language in the Austronesian family spoken in Papua New Guinea. In Table 10 is a list of the prepositions and postpositions of Jabem (**Dempwolff1939** **BradshawCzobor2005** 42–44; **Ross2002** 291).

Table 10: Prepositions and postpositions of Jabem

	Prepositions			**Postpositions**	
a.	tamiŋ	'next to, onto'	b.	lêlôm	'inside'
	baŋ	'close to'		lôlôc	'on top of'
	paŋ	'close to'		làbu	'under'
	ŋa	'instrumental'		sawa	'between'
	aᵑga	'from'		lùŋ	'in middle of'
				nêm	'in front of'
				mu	'behind'
				gala	'near'
				tali	'at edge of'

While all the postpositions again have meanings denoting specific locations, as we would expect of adpositions arising from head nouns in genitive constructions, three of the prepositions also have meanings of that sort ('next to', 'close to'). In fact, Dempwolff specifically suggests that these prepositions arose from verbs (suggesting, for example, that *tamiŋ* 'next to' comes from a verb meaning 'to be close upon').[2]

In Table 11 to 16 are lists of prepositions and postpositions from six other SVO& GenN languages that have both. All show patterns similar to those in the four languages describing above in this section, with the prepositions having meanings associated with adpositions arising from verbs and the postpositions with meanings associated with adpositions arising from nouns.

[2]I base this on Bradshaw & **Czobor2005**'s (**Czobor2005**) English translation of **Dempwolff1939**

Table 11: ǂHoã (Kxa: Botswana, **CollinsGruber2014** 101–105)

	Prepositions			Postpositions	
a.	kì	‘linker’	b.	na	‘in’
	ke	‘comitative’		za	‘by, beside’
				ǁq’am	‘above’
				ǂkà	‘below’
				ǂ’hàã	‘in front of’
				kya“m	‘near’

Table 12: Koromfe (Niger-Congo, Gur: Burkina Faso, Mali, **Rennison1997** 2017)

	Prepositions			Postpositions	
a.	la	‘instrumental, comitative’	b.	nɛ	‘benefactive, purpose, about’
	hal	‘until’		kana	‘like’
				dɔba	‘on top of’
				hɛrəga	‘beside, near’
				hogo	‘under’
				jɪka nɛ	‘in front of’
				joro	‘in, inside’
				bɛllɛ	‘behind’
				tʊllɛ	‘in the middle of, between’

Table 13: Mandarin Chinese (Sino-Tibetan, Sinitic: China, **LiThompson1981**)

	Prepositions (or coverbs)			Postpositions (or locative particles)	
a.	gēn	'with (comitative)'	b.	shàng	'on top of, above'
	gěi	'for' (benefactive)		xià	'below'
	bǎ	object marker		lǐ	'in, inside'
	duì	'toward'		wài	'outside'
	cóng	'from'		qián	'in front of'
	zài	'at'		hòu	'behind'
	tì	'instead of'		páng	'beside'
	bèi	'by'		dōngbu	'east of'
	àn	'according to'		zhèr	'this side of'
	dào	'to'		qián	'in front of'
				hòu	'behind'
				páng	'beside'
				zhōngjian	'in the centre of'

Table 14: Koyra Chiini (Songhay: Mali, **Heath1999** 104–109)

	Prepositions			Postpositions	
a.	nda	'comitative, instrumental'	b.	se	'dative'
	bilaa	'without'		ra	'locative'
	hal	'until'		ga	'beside, from'
	jaa	'since'		doo	'at the place of'
	bara	'except'		banda	'behind'
	kala	'except'		beene	'above'
				čire	'under'
				kuna	'in'
				jere	'beside'
				jine	'in front of'
				maasu	'inside'
				tenje	'facing'

Table 15: Taba (Austronesian; Indonesia, **Bowden2001** 109–111)

	Prepositions			Postposition	
a.	ada	'comitative, instrumental'	b.	li	'on, in, at'[a]
	pake	'instrumental'			
	untuk	'benefactive'			
	lo	'like'			

[a]The fact that the one postposition in Taba has general locative meaning does not fit the expectations for a postposition in a GenN language. But the fact that it is locative while the prepositions are not does fit loosely. It is possible that it originally had a narrower locative meaning that has become bleached.

Table 16: Dagbani (Niger-Congo, Gur: Ghana, **Olawsky1999**)

	Prepositions			Postpositions	
a.	ni	'comitative, instrumental'	b.	nyaaŋa	'behind'
	jɛndi	'about, concerning'		zuɣu	'on top of'
				gbinni	'under'
				sani	'towards'
				sunsuuni	'in the middle of'
				ni	'in, at, to'
				puuni	'inside'
				polo	'in the direction of'
				lɔŋni	'under'

check uniformity. Maybe Table 7?

The languages illustrated in (10) to Table 16 above are instances of SVO languages with GenN order and both prepositions and postpositions. Though less common, there are also languages of the opposite sort, OV languages with NGen order and both prepositions and postpositions, where the semantics associated with prepositions and postpositions respectively is the opposite of that found in SVO& GenN languages. An example is Iraqw. Example (19) illustrates a preposition *daandú* 'behind'. That it has grammaticalized from the head noun in a genitive construction is clear from the fact that it occurs in construct state, the morphological form that head nouns take in genitive constructions.

(19) Iraqw (Cushitic: Tanzania; **Mous1993** 97)
loo'a i **daandú** **hunkáy**.
sun 3SBJ **behind.CONSTR cloud**
'The sun is **behind the cloud**.'

In contrast, example (20) illustrates a postpositional clitic *=i* 'directional' that attaches to the last word in the noun phrase. In (20) it attaches to the noun *do'* 'house', the possessor of *afkú* 'mouth' ('door'), but it is marking the entire noun phrase *afkú do'* 'mouth (door) of the house' as the goal of the motion denoted by the verb *qaas* 'put'.

(20) Iraqw (Cushitic: Tanzania; **Mous1993** 252)
famfe'amo u-n af-kú do'=i qaas-áan.
snake MASC.OBJ-EXPEC **mouth-CONSTR.MASC house=DIR** put-1PL
'Let us put a snake **on the door of the house**.'

In Table 17 is a list of prepositions and postpositions in Iraqw (**Mous1993** 95–107).

Setting aside momentarily the first three prepositions in Table 17, the semantics associated with the prepositions and postpositions in Iraqw is the reverse of what we found in (10) to (16) for SVO& GenN languages. Namely, in Table 17, it is the prepositions which denote specific locations, while the postpositions have meanings that are generally associated with adpositions arising from verbs.

The first three prepositions in Table 17 have the same meanings as the first three postpositions in the table. Their meanings are thus ones that we might have expected to be associated with postpositions in an OV language. These prepositions take the form of /a/ plus the corresponding postpositional clitics. **Mous1993** speculates that the /a/ in these forms may have originally been the copula *a*. It is possible that these prepositions have arisen by analogy to other prepositions in the language.

Table 17: Prepositions and postpositions in Iraqw

	Prepositions			Postpositions	
a.	ar	‘instrumental’	b.	=(a)r	‘instrumental, comitative’
	as	‘because of’		=sa	‘because of’
	ay	‘to’		=i	‘to’
	dír	‘to		=wa	‘from’
	amór	‘at’			
	daandú	‘on’			
	alá	‘behind’			
	gurúu	‘inside’			
	gamú	‘under’			
	bihháa	‘beside’			
	tlaʿá(ng)	‘between’			
	tseeʿá	‘outside’			
	afíqoomár	‘until’			
	gawá	‘on’			
	geerá	‘before’			
	afá	‘at the edge of’			
	bará	‘in’			

A second instance of an OV& NGen language with both prepositions and postpositions is Kanuri. Example (21) illustrates the locative-instrumental postpositional clitic *=lan* attaching to a postnominal modifier *Musa=be* ‘Musa’s’, marking *fər Musa=be* ‘Musa’s horse’ as an instrumental.[3]

(21) Kanuri (Saharan: Nigeria, Niger; **Hutchinson1976** 5)
[fər Musa=be]=lan kadio.
[horse Musa=GEN]=INS come.PST.3SG
‘He came **on/by Musa’s horse**.’

Kanuri also has prepositions, like *suro* ‘inside’ in (22).

[3] There are thus two postpositional clitics in the phonological word *Musa=be=lan* in Table 17, the *=be* marking Musa as possessor of *fər* ‘horse’ and the *=lan* marking *fər Musa=be* ‘Musa’s horse’ as an instrumental.

(22) Kanuri (Saharan: Nigeria, Niger; **Hutchinson1976** 80)
suro fato=be=ro kargawo.
inside house=GEN=to enter.PST.3SG
'He went **into the house**.'

Note that *suro* retains its nominal nature in (22), in that its complement *fato* 'house' is marked as a possessor, with the genitive postpositional clitic *=be*, and the entire phrase marked with the postpositional clitic *=ro* 'to', so that (22) could be glossed as 'He went to the inside of the house'. To what extent these locational nouns have grammaticalized as prepositions is not clear. Even if they have not grammaticalized much yet, they illustrate how an OV& NGen language could acquire prepositions.

In Table 18 is a list of prepositions and postpositions of Kanuri (**Hutchison1981** 257–263).

Table 18: Prepositions and postpositions of Kanuri

	Prepositions			**Postpositions**	
a.	bótówò	'next to'	b.	=(là)n	'locative, instrumental'
	cî	'at edge of'		=rò	'benefactive, indirect object, to'
	dàryé	'at the end of'		=mbèn	'through, towards'
	dáwù	'in middle of'			
	fúwù	'in front of'			
	fərtə	'at base of'			
	gəré	'next to'			
	kátè	'between'			
	kəlâ	'on top of'			
	ngáwò	'behind, after'			
	sədíà ~ cídíà	'under'			
	súró	'inside, during'			

The meanings associated with the prepositions in Kanuri are similar to those of the prepositions in Iraqw, but are also similar to the meanings of the postpositions in the various SVO & GenN languages discussed above. Conversely, the meanings associated with the postpositions in Kanuri are similar to those of the postpositions in Iraqw and also similar to the meanings of the prepositions in the various SVO& GenN languages discussed above

There is another instance of a language with both prepositions and postpositions that provides an interesting variation of the argument in this section,

namely English. While English is predominantly a prepositional language, it has at least two postpositions, *ago* and *notwithstanding*, as in (23).[4]

(23) (English)

a. I saw him three weeks ago.

b. I went to the concert, the doctor's advice notwithstanding.

What is unusual about these two postpositions in English is that although both are apparently grammaticalizations of verbs, they are ones where what is now the object of that postposition was originally the subject of the verb (rather than the object, the more common situation with grammaticalizations from verbs). According to the Merriam Webster online dictionary,[5] *ago* comes from an obsolete verb meaning 'pass' so that *three weeks ago* derives from *three weeks have passed*, where *three weeks* was originally the subject of this verb. And *notwithstanding* comes from *not* plus a form of the verb *meaning 'withstand'* in the sense of *providing an obstacle for*; again, what is now the object of the postposition *notwithstanding* was originally the subject of the verbal expression. The fact that these two words arose as postpositions rather than as prepositions reflects the fact that subjects normally preceded the verb, even in earlier varieties of English when word order was more flexible. Again, only a grammaticalization account explains these.

The evidence in this section involves data that only grammaticalization can explain. An explanation in terms of grammaticalization for the correlation between the order of verb and object and order of adposition and noun phrase as well as the correlation between the order of noun and genitive and order of adposition and noun phrase predicts that we should find both prepositions and postpositions in the same language where the former derive from verbs and the latter head nouns in genitive constructions, as well as predicting the semantic differences between the two types of adposition. The evidence in this section shows how these predictions are borne out. There is no obvious way by which accounts in terms of processing or similarity could explain this data.

[4] *Notwithstanding* also occurs as a preposition. The postpositional use is apparently the original use. I suspect that the use as a preposition arose due to its semantic similarity to another preposition *despite*.

[5] https://www.merriam-webster.com/dictionary

3 What grammaticalization does not explain

The preceding section provides evidence that grammaticalization explains, at least partly, the correlation between the order of verb and object and order of adposition and noun phrase as well as the correlation between the order of noun and genitive and order of adposition and noun phrase. In this section, I discuss the question whether grammaticalization fully explains word order correlations and argue that it does not. I first discuss word order correlations for which there does not seem to be any good explanation in terms of grammaticalization. Table 19 provides a list of pairs of elements that are shown by **Dryer1992** to correlate with the order of verb and object, where the verb patterner refers to elements that occur first in these pairs more often among VO languages than among OV languages (and where the object patterner refers to the other member of the pair).

Table 19: Pairs of elements that correlate with the order of verb and object

verb patterner	object patterner	example
verb	adpositional phrase	*slept + on the floor*
verb	manner adverb	*ran + slowly*
copula verb	predicate	*is + a teacher*
'want'	VP	*wants + to see Mary*
noun	relative clause	*movies + that we saw*
adjective	standard of comparison	*taller + than Bob*
complementizer	clause	*that + John is sick*
question particle	sentence	intentionally empty??
adverbial subordinator	clause	*because + Bob has left*

For none of these pairs of elements that correlate with the order of verb and object is there a convincing explanation in terms of grammaticalization. For example, the order of verb and adpositional phrase most likely correlates with the order of verb and object because of semantic similarities between these two pairs of elements or because of processing factors. It is hard to imagine an explanation in terms of grammaticalization for this correlation.

I devote the remainder of this section to discussing the correlation between the order of verb and object and the order of noun and genitive. While there have been attempts to explain this correlation in terms of grammaticalization, I claim here that such attempts fall short of providing a plausible explanation. A good

summary of this approach is provided by Collins (this volume). However, most of the cases discussed by Collins are highly speculative, especially compared to the evidence for adpositions deriving from verbs or nouns. The arguments involve cases where the constructions now used for main clauses are claimed to have originated from nominalizations (where a construction like *John's seeing Peter* is claimed to have replaced an existing finite construction like *John saw Peter*).[6] Assuming that the word order in nominalizations reflects the order of noun and genitive (an assumption that is probably valid), the new construction will employ an order of verb and object that reflects the order of noun and genitive.[7]

While there probably have been some instances in which a nominalization construction came to be used as the primary construction for main clauses, there is little evidence of this in most families and the correlation between the order of verb and object and the order of noun and genitive seems far too strong to be explained purely in this way. Consider the data in Table 20 on the relative frequency of the different orders of noun and genitive in OV languages.

Table 20: Order of noun and genitive in OV languages

	Africa	**Euras**	**Oceania**	**N.Amer**	**S.Amer**	**TOTAL**
OV& GenN	[26]	[46]	[87]	[34]	[54]	247
OV& NGen	13	1	10	0	1	25

[6]Some of Collins' arguments are particularly unconvincing. He cites data from Angas showing nominalizations being used for complements of the verb meaning 'want'. But this only shows that some languages express such complements using nominalizations; it provides no evidence of nominalizations coming to be used as main clauses. He also cites the large number of Austronesian languages as evidence for the frequency by which nominalizations become main clauses. But quite apart from the fact that Collins provides no evidence to support his claim that it is generally accepted that nominalizations came to be used as main clauses in Austronesian, the size of the family is not relevant; what is relevant is the number of instances of changes of this sort. A number of proposals that main clause constructions originated as nominalizations are based largely on the fact that the same case marker is used for both possessors and subjects (or transitive subjects). But there are many ways by which this can arise without nominalizations coming to be used as main clauses.

[7]It will also determine the order of verb and subject, especially for intransitive verbs. There are issues arising here that are beyond the scope of this paper. And while I find the evidence that grammaticalization explains the correlation between the order of verb and object and the order of noun and genitive unconvincing, I must concede that it would account for the large number of SVO& GenN languages. In other words, it would account for the fact that the order of noun and genitive is one of the few orders that correlates not only with the order of verb and object but also with the order of verb and subject (**Dryer2013**).

Table 20 shows that GenN order outnumbers NGen by 247 to 25 genera, a ratio of almost 10-to-1. The evidence for nominalizations coming to be used as main clauses is far too meagre to account for such a strong correlation.

It should be noted that the order of noun and genitive correlates with the order of verb and object less strongly than the order of adposition and noun phrase correlates with either the order of verb and object or the order of noun and genitive: Tables 1 and 2 above show a particularly strong correlation between the order of verb and object and the order adposition and noun phrase; Tables 3 and 4 show an even stronger correlation between the order of noun and genitive and the order of adposition and noun phrase. But the large number of SVO& GenN languages shows that the correlation between the order of verb and object and the order of noun and genitive is less strong.

One possible explanation for why the correlation between the order of verb and object and the order of noun and genitive is weaker is that all three of these correlations are due in part to factors other than grammaticalization (such as the processing explanations of **Dryer1992** and **Hawkins1994**; **Hawkins2004**; **Hawkins2014**), but that grammaticalization augments the correlation between the order of verb and object and the order of adposition and noun phrase as well as the correlation between the order of noun and genitive and the order of adposition and noun phrase. In other words, it may be a mistake to try to choose between grammaticalization and other factors in explaining word order correlations; they may conspire to lead to these stronger correlations.

In fact, data presented by Dryer (1992, 2013) suggests that the correlation between the order of verb and object and the order of adposition and noun phrase as well as the correlation between the order or noun and genitive and the order of adposition and noun phrase are stronger than most of the correlations in Table 19 above. Since there do not appear to be promising explanations for those correlations in terms of grammaticalization, the fact that the two correlations involving adpositions are particularly strong suggests again that both grammaticalization and other factors play a role in explaining those correlations.

Note also that grammaticalization explains the fact mentioned above in §2.1 that the preference for postpositions among OV languages is stronger than the preference for prepositions among VO languages. Namely, OV languages are overwhelmingly GenN so that both sources for adpositions lead to postpositions in OV languages. In contrast there are many SVO languages with GenN order. In such languages the adpositions derived from head nouns will be postpositions, so that (assuming some such languages lack adpositions derived from verbs) we expect to find SVO languages with postpositions.

4 Order of noun and definiteness marker

In this section, I discuss a different type of problem for grammaticalization accounts of word order correlations. In the cases discussed in §3, grammaticalization simply fails to predict a word order correlation which can be shown to be real. In the case discussed in this section, grammaticalization makes a prediction that turns out not to hold, involving the order of definiteness marker and noun.

The most common grammaticalization source for markers of definiteness appears to be demonstratives. In fact my database contains 102 instances of languages that use demonstratives as markers of definiteness, compared to 274 languages with markers of definiteness that are distinct from demonstratives. Both the order of definiteness marker and noun and the order of demonstrative and noun exhibit weak correlations with the order of verb and object, but what is surprising from the perspective of grammaticalization is that they exhibit opposite correlations. Namely, definiteness markers *precede* the noun more often in VO languages than in OV languages, while demonstratives *follow* the noun more often in VO languages than in OV languages.

Consider first definiteness markers in VO languages. Table 21 provides data on the order of definiteness marker and noun in VO languages. The last line in Table 21 gives the proportion of the number on the first line as a proportion of the sum of the number on the first line and the number on the second line. For example, the .21 on the third line in Table 21 under Africa represents 8 as a proportion of 39 (the sum of 8 and 31). I use these proportions in the discussion below.

Table 21: Order of noun and definiteness marker in VO languages

	Africa	**Euras**	**Oceania**	**N.Amer**	**S.Amer**	**TOTAL**
VO& DefN	8	[11]	[16]	[17]	[7]	59
VO& NDef	[31]	3	13	8	0	55
Proportion DefN	.21	.79	.55	.68	1.00	Mean=.64

can we use ∅ or $\bar{x}$ instead of 'mean' (all tables)?

Table 21 shows the two orders of definiteness marker and noun to be about equally common among VO languages, with DefN order found in languages in 59 genera and NDef order found in languages in 55 genera. This is a case, however, where the total numbers of genera are somewhat misleading, since one area,

Africa, exhibits a very different pattern from what we find in the other four areas. In Africa, genera containing VO languages in which the definiteness marker follows the noun outnumber genera containing VO languages in which the definiteness marker precedes the noun by 31 to 8. In the other four areas, in contrast, it is more common among VO languages for the definiteness marker to precede the noun; in fact, in three of the areas (Eurasia, North America, and South America), DefN order is more than twice as common as NDef order. The mean of the proportions over the five areas, namely .64, also reflects a preference for DefN order among VO languages. Another way to see this is that if we exclude Africa, DefN outnumbers NDef among VO languages by 51 to 24.[8]

Table 22 provides comparable data on the order of definiteness marker and noun among OV languages. We again find only a small difference, though it is NDef that outnumbers DefN among OV languages, by 53 genera to 38.

Table 22: Order of noun and definiteness marker in OV languages

	Africa	**Euras**	**Oceania**	**N.Amer**	**S.Amer**	**TOTAL**
OV& DefN	3	[9]	15	4	[7]	38
OV& NDef	[12]	5	[23]	[9]	4	53
Proportion DefN	.20	.64	.39	.31	.64	Mean=.44

But what is revealing is to compare the proportions from the last lines of Tables 21 and 22, given in Table 23.

Table 23: Proportion of genera containing DefN languages among VO vs. OV languages

	Africa	**Euras**	**Oceania**	**N.Amer**	**S.Amer**	**TOTAL**
VO	[.21]	[.79]	[.55]	[.68]	[1.00]	Mean=.64
OV	.20	.64	.39	.31	.64	Mean=.44

last column header total->mean?

[8] The higher preference for NDef order among VO languages in Africa reflects a general difference between Africa and the rest of the world in that postnominal modifiers are more common in Africa than elsewhere (**Dryer2010**). Table 20 above shows a similar difference between Africa and the rest of the world: while GenN outnumbers NGen among OV languages overall by almost 10-to-1, the ratio in Africa is only 2-to-1 and over half (13 out of 25) of the genera containing OV& NGen languages are in Africa.

Here we find that although the margin of difference in Africa is very small, it is still the case that the proportion of genera containing DefN languages is greater among VO languages in all five areas. This gives us reason to conclude that there is a correlation, albeit a weak one, between the order of verb and object and the order of definiteness marker and noun, with the definiteness marker preceding the noun more often among VO languages than among OV languages.

Given the fact that the most common grammaticalization source for definiteness markers appears to be demonstratives, we might expect to find a similar correlation between the order of verb and object and the order of demonstrative and noun. We do find a clear trend, but it is the opposite correlation. Namely while definiteness markers precede the noun more often among VO languages compared to OV languages, demonstratives tend to follow the noun more often among VO languages compared to OV languages.

Tables 24 to 24 provide data supporting this. Table 24 provides relevant data for VO languages. It shows that although NDem order is slightly more common than DemN order, by 118 genera to 92, this order is more common in only three of the five areas (and in fact, if we exclude Africa, it is DemN order that is more common among VO languages, by 84 genera to 66)

Table 24: Order of noun and demonstrative in VO languages

	Africa	**Euras**	**Oceania**	**N.Amer**	**S.Amer**	**TOTAL**
VO& DemN	8	12	24	[24]	[24]	92
VO& NDem	[52]	[16]	[31]	12	7	118
Proportion DemN	.13	.43	.44	.67	.77	Mean=.49

However, Table 25 shows that among OV languages, DemN order is about twice as common as NDem order, by 181 genera to 95, although there are two areas where NDem is more common among OV languages.

Table 25: Order of noun and demonstrative in OV languages

	Africa	**Euras**	**Oceania**	**N.Amer**	**S.Amer**	**TOTAL**
OV& DemN	16	[44]	45	[30]	[46]	181
OV& NDem	[18]	6	[57]	6	8	95
Proportion DemN	.47	.88	.44	.83	.85	Mean=.70

Again, it is useful to compare the proportions from the last lines of Tables 24 and 25, shown in Table 26.

Table 26: Proportion of genera containing DemN languages among VO vs. OV languages

	Africa	**Euras**	**Oceania**	**N.Amer**	**S.Amer**	**TOTAL**
VO	.13	.43	.44	.67	.77	Mean=.49
OV	[.43]	[.88]	.44	[.83]	[.85]	Mean=.70

last column header total->mean?

Table 26 shows that the proportion of genera containing DemN languages is higher among OV languages in four areas while the proportion is the same in the fifth area (Oceania).[9] There is thus a clear trend in the opposite direction from what we found for the order of definiteness marker and noun. Given that the most common grammaticalization source for definiteness markers appears to be demonstratives, this contrast is quite surprising.

I have no explanation for the source of this difference between definiteness markers and demonstratives. But I will share some interesting data from particular languages that conforms to this difference. First, there are a few languages in which the same form is used as a demonstrative and as a marker of definiteness, but this form occurs on different sides of the noun, depending on its function. In Swahili, the forms that are used as distal demonstratives when following the noun function as markers of definiteness when they precede the noun, as shown in (24). Since Swahili is SVO, this difference conforms to the contrast in the crosslinguistic data shown above.

(24) Swahili (Niger-Congo, Bantoid; **Ashton1947** 59)

a. m-tu **yu-le**
NC$_1$-man **NC$_1$-that**
'that man'

b. **yu-le** m-tu
NC$_1$-DEF NC$_1$-man
'the man'

[9] If we compute the proportions to three decimal places, DemN is also higher among OV languages compared to VO languages in Oceania (by .441. to .434). However, this difference is too small to base any conclusion on.

In Abui, we find the opposite situation: the form *do* functions as a demonstrative when it precedes the noun, as in (25a), but as a marker of definiteness when it follows the noun, as in (25b).

(25) Abui (Timor-Alor-Pantar: Indonesia; **Kratochvil2007** 111, 114)

a. **do** sura
this book
'this book (near me)'

b. kaai **do** 'the dog (I just talked about)'
dog DEF

Significantly, Abui is an OV language, so the fact that Abui exhibits the opposite pattern from what we saw in Swahili again conforms to the crosslinguistic pattern described above.

The situation in Ute is similar to that in Abui. Namely Ute is OV and the word *'u* functions as a demonstrative when it precedes the noun, as in (39a), but as a marker of definiteness when it follows the noun, as in (39b).

(26) Ute (Uto-Aztecan: United States; **Givón2011** 50, 38)

a. **'ú** kava sá-gha-rʉ-mʉ qhárʉ-kwa-pʉga.
that.S horse.SBJ white-have-NMLZ-ANIM.SBJ run-go-REM
'That white horse ran away.'

b. ta'wa-chi **'u** sivaatu-chi paqha-qa.
man-ANIM.SBJ **DEF.SBJ** goat-ANIM.OBJ kill-ANT
'The man killed a goat.'

The situation in Loniu is somewhat different. In Loniu, the definiteness marker and demonstrative are similar in form, though not identical, with *iy* as the definiteness marker and *iyɔ* as the demonstrative. The two in fact can co-occur as in (27), with the definiteness marker preceding the noun, and the demonstrative following the noun.

(27) Loniu (Austronesian: Papua New Guinea; **Hamel1994** 100)
iy amat iyɔ
DEF man this
'this man'

Again, since Loniu is VO, this order difference conforms to the crosslinguistic pattern described above.

And we find similar phenomena in cases where the definiteness marker and demonstrative are completely different in form but can co-occur, with one preceding the noun and one following. In Kana, the definiteness marker precedes the noun while the demonstrative follows, as in (28).

(28) Kana (Niger-Congo: Nigeria; **Ikoro1996** 70)
ló bárí āmā
DEF fish this
'this fish'

Since Kana is VO, this conforms to the crosslinguistic pattern. Contrast this with the situation in Kwoma (Washkuk), which is OV, and in this case it is the demonstrative that precedes the noun and the definiteness marker that follows, as in (29).

(29) Kwoma (Sepik: Papua New Guinea; **Kooyers1974** 49)
kata ma rii
that man DEF
'that man'

These differences between demonstratives and definiteness markers are a puzzle if demonstratives are the primary grammaticalization source for definiteness markers. It should be emphasized, however, that although definiteness markers and demonstratives exhibit very different patterns in terms of how they correlate with the order of verb and object, it is still the case that they correlate with each other, that the order of definiteness marker and noun and the order of demonstrative and noun correlate. This is shown in Tables 27 and 28, excluding languages where the definiteness marker is the same as the demonstrative. Table 27 shows that among DefN languages with definiteness markers that are distinct from demonstratives, it is approximately twice as common for the demonstrative to precede the noun as well, by 41 genera to 20.

Table 27: Order of noun and demonstrative in DefN languages

	Africa	**Euras**	**Oceania**	**N.Amer**	**S.Amer**	**TOTAL**
DefN& DemN	3	[7]	[12]	[11]	[8]	41
DefN& NDem	[4]	3	7	3	3	20

Conversely, Table 28 shows that among NDef languages with definiteness markers that are distinct from demonstratives, it is much more common for the demonstrative to follow the noun as well, by 67 genera to 11.

Table 28: Order of noun and demonstrative in NDef languages

	Africa	Euras	Oceania	N.Amer	S.Amer	TOTAL
NDef& DemN	4	3	2	1	1	11
NDef& NDem	[33]	[6]	[19]	[8]	1	67

While grammaticalization probably plays some role in explaining this correlation, it seems likely that the clear semantic similarity between definiteness markers and demonstratives plays a role as well. There is also a correlation between the order of definiteness marker and noun and the order of indefinite marker and noun, a correlation that is presumably due to semantic similarity or processing, not grammaticalization.

5 Conclusion

I have argued that there is evidence that any approach to explaining word order correlations that ignores the role of grammaticalization is inadequate. At the same time, I have argued that while grammaticalization may play a role in explaining some correlations, a purely grammaticalization approach fails as well.

Although I have focused my discussion of SVO& GenN languages on those with both prepositions and postpositions, further research is needed on SVO& GenN languages with prepositions as the only or dominant type or with postpositions as the only or dominant type. Grammaticalization theory would predict that SVO& GenN languages with prepositions will be ones where the primary source of adpositions is verbs, while SVO& GenN languages with postpositions will be ones where the primary source of adpositions is head nouns in genitive constructions. I suspect that this is true and if so, would further bolster the argument that grammaticalization plays an important role in explaining correlations involving adpositions. One reason to suspect it is true is the geographical distribution of the two types of languages. My database includes 21 genera containing SVO& GenN languages with prepositions and 13 of these genera (almost two thirds of them) are in an area stretching from China and southeast Asia through Austronesian. The fact that so many of the SVO& GenN languages are

in this region is significant since my impression is that the grammaticalization of adpositions from verbs is especially common in this region. Conversely, my database includes 19 genera containing SVO& GenN languages with postpositions and only two of these genera are in the region mentioned above stretching from China through Austronesian where SVO& GenN& Pr languages are common. I suspect that this is because outside that region, it is more common for adpositions to grammaticalize from nouns. However, this is a matter for future research.

6 Abbreviations

The paper abides by the Leipzig Glossing Rules. Additional glosses include the following ones:

ANIM	animate	EXPEC	expectational
ANT	anterior	NC	noun class
CONSTR	construct state	REM	remote

7 Acknowledgements

I am indebted to Lea Brown, Karsten Schmidtke-Bode and members of the audience at the 2015 meeting of the DGfS (the German Linguistic Society) for comments on an earlier version of this paper. I also acknowledge funding from The Social Sciences and Humanities Research Council of Canada, the National Science Foundation (in the United States), the Max Planck Institute for Evolutionary Anthropology (in Leipzig, Germany) and the Humboldt Foundation (in Germany).

Explanation in typology