(Accepted for publication in *Journal of Historical Linguistics*)

Argument marking reflects audience-design effects: Evidence from diachrony and typology

Ilja A. Seržant
University of Potsdam

Abstract

This paper contrasts typological and diachronic evidence on argument marking with theoretical proposals made in psycholinguistic research about the ways various types of audience design strategies affect production, with the focus on ambiguity avoidance in argument roles. The aim is to find convergent evidence from psycholinguistic as well as from typological and diachronic research. This evidence suggests that argument marking in functional constructions is shaped by generic audience-design effects, while cross-linguistic and diachronic support for utterance-specific audience-design is very scarce. Diachronically utterance-specific audience-design tends to be abandoned in favor of functionally related generic audience-design strategies. The paper reconciles earlier claims about ambiguity avoidance affecting argument marking and the view of ambiguity avoidance as having no effects on argument marking. It is suggested that ambiguity avoidance is a strong pressure in argument marking. I show that a high ambiguity potential of role identification of various constructions correlates positively with the degree of non-differential argument marking; and, vice versa, differential marking is more likely in those constructions which provide reliable cues for role identification.

1. Introduction

Audience design means "tailoring utterances to addressee needs", as Wardlow Lane & Ferreira (2008: 1) put it. It generally implies that the speaker not only economizes on their production costs due to the limits of available resources (working memory, articulation effort, etc.) but also bears in mind the needs of the interlocutor in order to achieve successful information transfer and, accordingly, tries to adapt their speech to these needs. For example, speaking louder is an audience-design strategy that increases chances to be understood in a noisy situation. Grammatical markers may also be used to increase the chances of being correctly understood, for example, in order to avoid role ambiguity in two-participants events such as *Peter, John, sees* (see Kurumada & Jaeger 2015). Some of these means may then become conventionalized and thus part of the grammar of a language, as I argue below.

Ambiguity is a universal property of human languages. Speakers and hearers share a substantial amount of information from world knowledge and previous discourse, which the hearer can exploit for inferring the intended meaning in contexts of ambiguity. This property of human communication makes information transfer more efficient by allowing the speaker to balance production costs while still achieving communicative goals (Ferreira 1996, Wardlow Lane & Ferreira 2008, Piantadosi et al. 2012, Wasow et al. 2005).

One could take *ambiguity* to be anything that does not supply sufficient information. To give an example, consider the utterance *Yesterday I met …*. This utterance can be said to contain

ambiguity as to who was met. In turn, naming the object referent, for example *Peter* in *Yesterday I met Peter*, might also be said to be an instance of ambiguity avoidance because it "disambiguates" the incomplete utterance with one specific alternative, namely, *Peter* out of the entire generic set of beings one can possibly meet. However, such an approach to ambiguity would be far too broad and ambiguity avoidance would be downplayed to just providing any kind of information. Instead, I refer to this kind of phenomenon as *vagueness* and utterance such as *Yesterday I met* as vague.

Thus, I distinguish between *ambiguity* and *vagueness*, which are not always strictly kept apart (cf. Piantadosi et al. 2012). The intuition behind the concept of ambiguity – as opposed to vagueness – is that ambiguity is a very special subtype of vagueness. *Ambiguity* implies that there is a very small set of activated, specific, non-generic alternative interpretations; for example, the ambiguity in the assignment of semantic and/or syntactic roles in an event with two participants. Moreover, I take the prefix *ambi-* literally and assume only two competing, non-generic interpretations, drawing on Wasow et al.'s more general definition of ambiguity (2005: 265):

(1) Ambiguity
Availability of two specific, activated, non-generic meanings / functions, both of which are likely given the specific activation window.

In addition, I define *activation window* as follows:

(2) Activation window
Activation window is any portion of the information that is immediately available to the speaker (and the hearer) within the current discourse situation.

The activation window may extend over the entire discourse, or it may only include the particular constituent of the utterance that is being produced by the speaker at a given moment of time in incremental speech production.

Accordingly, *ambiguity avoidance* is defined as follows:

(3) Ambiguity avoidance
Potential ambiguity is avoided by dedicated means (marker, construction, word order, etc.) that singles out only one interpretation to the exclusion of the other interpretation.

Various types of ambiguity avoidance strategies have been discussed in psycholinguistic research when exploring differentially marked constructions such as the optional complementizer *that* in English, garden-path sentences, lexical ambiguity, etc. (Ferreira 1996, 2019; Piantadosi et al. 2012; Wasow 2002). My focus is on argument marking.

In typology, *ambiguity avoidance* in argument marking was considered as one of the most important constraints by earlier authors (cf., *inter alia*, Comrie 1978: 379-380, 1989: 124-127, Dixon 1994, Silverstein 1976, Kibrik 1997; for differential agent marking, see Dixon 1979: 73). With respect to differential object marking, Bossong (1985a: 117) even claimed that its emergence is primarily due to ambiguity avoidance – a suggestion that I draw on below in a modified way. However, recent studies take the opposite position and state that ambiguity avoidance does not provide a plausible explanation for differential argument marking systems

(Haspelmath 2021b: 164), certainly not as a major factor (*inter alia*, Arkadiev 2008a). Instead, most recently, Haspelmath (2021a, 2021b) has advanced the role-reference association universal in (4) that regulates the presence vs. absence of a (longer) marker in transitive clauses:

(4) The role-reference association universal (Universal 1) (Haspelmath 2021b: 125)
"Deviations from usual associations of role rank and referential prominence tend to be coded by longer grammatical forms if the coding is asymmetric."

This universal draws on earlier suggestions such as Dixon's (1994: 85).[1]

In this paper, I compare and combine the role-reference universal on argument marking in (4) with other argument-marking phenomena observed across various functional constructions (see below) and argue that, eventually, argument marking in functional constructions is primarily driven by ambiguity avoidance. The role-reference universal in (4), in turn, synchronically does hold but it is in itself the diachronic result of the effects of ambiguity avoidance in argument marking as well. This result is motivated by the strive towards automatized, generic audience-design strategies and away from utterance-specific audience-design strategies (§4). In general, both ambiguity avoidance and a (un)expectedness account such as (4) are closely related factors that most of the time yield the same result and are, therefore, quite difficult to disentangle, especially if the activation window (2) is just the very nominal in question (which is likely given that the speech is produced and comprehended incrementally). However, this is not always the case and I will discuss a set of examples in which the activation window is, at least, the entire clause and the properties of the other argument are taken into account.

The main line of argumentation in this paper is based on the argument that there is a positive correlation between the degree of ambiguity potential in the assignment of semantic/syntactic roles, on the one hand, and the likelihood of argument marking, on the other, across different constructions. Moreover, sometimes entirely expected role-reference associations are nevertheless consistently marked. These constructions, however, have inherently a very high degree of ambiguity potential. And, vice versa, monovalent intransitive constructions with no ambiguity in role-reference association consistently prefer no marking of the sole argument even if its role-reference is entirely unexpected. I restrict myself to cross-linguistic types of *functional constructions* (mostly bivalent ones) such as transitive constructions (A and P arguments), ditransitive constructions (with a focus on A and R arguments), comparative constructions (Standard and Comparee arguments) and locational constructions (Figure and Ground arguments). Functional constructions are constructions that convey grammatical functions. Argument roles are primarily assigned by grammatical functions and not lexically, for example, not by specific (bivalent) lexical verbs such as English *to look at.*

I do not discuss monovalent or bivalent *lexical constructions*. With lexical constructions, argument marking is primarily determined by the specific lexical verbs and the particular semantic roles of the arguments and adjuncts (Siewierska & Bakker 2009: 292), compare English *to look at* vs. *to compare with* vs. *to distinguish from*, etc. (the *indexing function* in Arkadiev 2008a, Næss 2006; or the *corresponding function* in De Hoop & Malchukov 2007). One could claim that the markers of oblique objects and adjuncts in lexical

[1] Building on his Nominal Hierarchy, which is a reference hierarchy, Dixon (1994: 85) states: "It is plainly most natural and economical to 'mark' a participant when it is in an unaccustomed role."

constructions, in addition to marking the semantic role, also serve the purpose of ambiguity avoidance as they ensure that both participants (e.g. *the looker* and *the object looked at*) are distinguished (Siewierska & Bakker 2009: 292). However, since one cannot disentangle these two functions and thus unequivocally prove the effect of ambiguity avoidance, the null hypothesis here remains to be only that bivalent lexical constructions do not contradict the idea of the ambiguity-avoidance function of argument marking.

The remainder of the paper consists of five sections. Section 2 discusses synchronic effects of ambiguity avoidance on argument marking. Section 3 provides diachronic evidence of various types. Section 4 discusses the evidence provided in the preceding sections in the light of psycholinguistic research on audience design and suggests a cognitive explanation. Finally, Section 5 summarizes the results and conclusions.

2. Synchronic evidence for ambiguity avoidance in bivalent functional constructions

2.1 Transitive constructions

Differential argument marking (henceforth DAM) (Witzlack-Makarevich & Seržant 2019) is exemplified with the *differential object marking* (henceforth DOM) in (5). In Spanish, only animate and specific Ps are marked by the preposition *a* while other object types generally remain unmarked.

(5) Modern Spanish (von Heusinger & Kaiser 2005: 35)

a. *Vi* **(a)* *la* *mujer.*
see.PST.1SG **DOM** DEF woman
'I saw the woman.'

b. *Vi* *(*a)* *la* *mesa.*
see.PST.1SG **DOM** DEF table
'I saw the table.'

The DAM system of Japanese exemplified (DOM along with the differential subject marking) in (6) has been argued to be conditioned not only by such factors as animacy but also by ambiguity avoidance (Kurumada & Jaeger 2015, see also Fedzechkina et al. 2012). Given only the lexical input in (6) the hearer might not be able to determine who sees whom. Lack of argument marking would therefore lead to ambiguity here (given an activation window no larger than the utterance itself).

(6) Japanese (Kurumada & Jaeger 2015: 155)
Taro-(ga) *Hanako-(o)* *mi-ta.*
Taro-(NOM) Hanako-(ACC) see-PST
'Taro saw Hanako.'

Ambiguity avoidance means that sufficient grammatical marking is provided to avoid even potential confusion in role identification by the hearer. For a clause with two arguments, this requires an explicit coding of the role of at least one of the arguments, e.g. by nominative or

accusative marking in (6). Indeed, Kurumada & Jaeger (2015) provide experimental evidence in favor of ambiguity avoidance as a synchronic pressure conditioning DOM of Japanese.

Generally, ambiguity avoidance has been argued to be the main condition on DAM in earlier works (*inter alia*, Comrie 1978; 1989; Dixon 1994; Silverstein 1976). This stands in contrast with the fact that the cross-linguistic synchronic evidence from DOM systems and DAM systems in general shows only limited support for the role of ambiguity avoidance between A and P (*inter alia*, Aissen 2003, Malchukov 2008, various papers in de Hoop & de Swart 2008, see also McGregor 2010 on differential ergative marking); ambiguity-avoidance effects in particular languages rarely reach statistical significance as independent factors (Schikowski & Iemmolo 2015). DAM systems driven solely by ambiguity avoidance between A and P are typologically extremely rare if not absent at all. Such a DOM system would require an object marker only if both arguments of the verb may be misinterpreted as the subject. One such example may be DOM in Yongren Lolo as described by Gerner (2008):

(7) Yongren Lolo (Tibeto-Burman, Loloish: China; adapted from Gerner 2008: 299)[2]
ŋo *ɕɛmo* *tʰie* *ʈʂɔ* *ʑi.*
1SG snake DOM follow go
'I will follow the snake'

(8) Yongren Lolo (Tibeto-Burman, Loloish: China; adapted from Gerner 2008: 300)
Sɨka *tʰie* *χekʰu* *ti* *na.*
tree DOM house smash broken
'The house smashed the tree.'

The absence of the DOM marker *tʰie* in (7) and (8) would not be ungrammatical but would create ambiguity as to who is following whom in (7) or what is smashing what in (8) (Matisoff 1973: 156, Gerner 2008). However, along with the synchronically primary function of disambiguating P from A (and also R from A, see below (13)), this marker also has the diachronically primary function of coding contrastive focus (Gerner 2008: 298–289). Thus, even in this language, the DOM marker is not exclusively used for ambiguity avoidance.

Another example is Yuhup. Here, as Osipina Bozzi (2002: 145-146) writes, the accusative marking is found if both A and P have the same degree of agentivity (scil. animacy) or if A is lower in agentivity than P (in addition to definiteness and animacy constraints) (see also Stenzel 2008: 168).

(9) Yuhup (Nadahup; Osipina Bozzi 2002: 145)
a. *ɟàbběh* *kəǰhi* *tŏh-dĭh*
dog bite pig-DOM
'The dog bit the pig.'

b. *tŏh kəǰhi* *ɟàbběh-dĭh*
pig bite dog-DOM
'The pig bit the dog.'

[2] I simplified the transliteration and slightly adjusted the glossing of all examples from Gerner (2008).

I provide more evidence on ambiguity avoidance in DOM systems with specific input subtypes below in Section 3.

Very similar to Yuhup is Fore. Here, by contrast, the ergative case marker occurs on the A argument if it is lower in animacy than P:

(10) Fore (Nuclear Trans New Guinea; Foley 1986: 172)

a. *yága: wa a-egú-i-e*
pig man 3SG.P-hit-3SG.A
'the man hits the pig'

b. *yága:-wama wa a-egú-i-e*
pig-ERG man 3SG.P-hit-3SG.A
'the pig attacks the man'

In addition to flagging, indexing may also serve ambiguity avoidance. For example, in Mojeño Trinitario, the third person A is indexed with the general third-person prefix *ty-* only if the P argument is not a third person and, hence, confusion of A and P is not possible (Rose 2011: 477), as in the following example:

(11) Mojeño Trinitario (Arawakan; Bolivia; Rose 2011: 476)
ty-okpo-wokovi
3-meet-1PL
'He/she/it/they meet us.'

However, if both A and P are third person, then the general third-person index *ty-* would not allow the hearer to determine which of the two nominals is A because the third-person P index is always zero (Rose 2011: 474, 477). In scenarios[3] with both A and P being third person, Mojeño Trinitario employs a different set of markers for the third-person A argument which distinguishes gender (in the singular) and number. This, in turn, helps identifying A and P arguments better:

(12) Mojeño Trinitario (Arawakan; Bolivia; Rose 2011: 476)

ma	*'moperu-gra*	*mu-em-'o-po*	*to*	*jani-ono*
ART.M	child-DIM	3M-see-ACT-PERF	ART	bee-PL

'The little boy saw the bees.'

Other instances of ambiguity avoidance via argument marking of A have been reported. For example, Williams (1980: 98) states that, in Yuwaalaraay (Pama-Nyungan), the ergative marking of A may be left out in those contexts in which A and P are sufficiently disambiguated, either semantically (animate A vs. inanimate P) or morphosyntactically (if the P is marked by the dedicated accusative case).

Crucially, the limited synchronic evidence for ambiguity avoidance in transitive constructions correlates with the strong cue reliability of argument input for role identification.

[3] With *scenario* I refer to situations in which the input of both A and P matters for marking and only a specific combination of a specific input in A with a specific input in P – i.e. a specific scenario – would trigger the DOM marker (see Zúñiga 2006).

Animacy and definiteness of the input here very often helps to identify the roles. This evidence suggests that ambiguity-avoidance effects are less pronounced in transitive constructions which have high cue reliability for role identification. It has been repeatedly argued that, for example, inanimate and/or indefinite referents are strongly biased for the P role and thus through their meaning provide a highly reliable cue for role identification (Fauconnier & Verstraete 2014, see recently Levshina 2021). Given their high cue reliability it is expected that transitive constructions will have low potential for generating ambiguous configurations in the first place and, accordingly, should show less evidence for ambiguity avoidance.

2.2 Ditransitive constructions

The situation is entirely different in those constructions in which there is inherent ambiguity, that is, the properties of the input do not provide any cue for the identification of the roles. Thus, the same marker as used in examples (7) and (8) for DOM above is used for ambiguity avoidance between A and R with ditransitives in Yongren Lolo, cf. (13):

(13) Yongren Lolo (Tibeto-Burman, Loloish: China; adapted from Gerner 2008: 300)

ŋo	*subə*	*thi*	*bə ʐɔ*	*thiɛ*	*mo*	*(gə)*
1SG	book	one.NUM	CL 3SG	**DOM**	show	give.PRED

'I show him a book.' not *'He showed me a book.'

A and R are both most frequently definite and animate, leaving the hearer without any reliable cue for disambiguation between the two.[4] This inherent ambiguity and lack of cues for role identification is correlated with consistent, non-zero marking. The cross-linguistic evidence suggests that R tends to be marked differently from A across the board: at least some marking disambiguating between A and R is cross-linguistically the preferred option (see also Siewierska & Bakker 2009: 300). Thus, indirect-object constructions (indirective alignment) always involve some special ("dative-like") marking of R (Malchukov et al. 2011: 3). It is this construction type that is by far the most frequent cross-linguistically with 189 languages (50%) exhibiting this type in Haspelmath's sample (2013). Moreover, other construction types – double-object constructions (22% of the sample) and secondary-object constructions (17%) – most often involve some object marking on R: R=P but R≠A. Even in languages with differential R marking involving zero like English (*I gave Mary the book vs. I gave the book to Mary*) the R argument is unequivocally disambiguated from A via word order and agreement (only A agrees with the verb in English but not R). If differential R marking is found it is, thus, more of a symmetric type in contrast to differential P marking which is mostly asymmetric and involves no marking of P at all as one of the option (see Iemmolo 2013, Witzlack-Makarevich & Seržant 2018: 23-25). Haspelmath (2021b: 139-141) mentions other types of differential R marking – "splits" in his terms – most of which, nevertheless, do not include an option in which A and R would be coded alike but rather an alternation between a shorter coding (incl. the coding just by word order) and a longer coding of R.

To conclude, some R marking (≠A) is universally the preferred option. What is more, the marking of the R argument is only rarely differential with no coding at all as one of the options. There are only few exceptions, e.g., in Spanish, Catalan (Pineda & Royo 2017) or Yongren Lolo (13). In most languages, there is always some unequivocal marking of R (if only by means of word order).

[4] For example, 83% of indirect objects in a sample of written Swedish are human referents (Dahl 2000: 58).

Maintaining the hypothesis that ambiguity avoidance is a major pressure in argument marking can explain why R is so consistently marked differently from A in most languages and that it is rarely marked differentially as opposed to P in monotransitive constructions (where differential marking is the preferred option, see Sinnemäki 2014). By contrast, unexpectedness of the role-reference association in (4) above (Haspelmath 2021b: 125) cannot explain the ways of R marking. Since R is most of the time definite and animate, there should be nothing unexpected in it and, hence, no marking at all should be found as the preferred option if only unexpectedness were at work, as would be suggested by (4).

2.3 Intransitive monovalent constructions

In turn, S is the argument that tends to be morphologically unmarked most of the time and in most languages as per Greenberg's Universal 38 (Greenberg 1966: 59).[5] Yet, S tends to be predominantly definite and animate, as shown in Table 1.

Itzaj Maya	56% animate	Hofling (2003: 407)
Tojolabal	80% given, 79% animate	Unpublished ms.
Kamang	83.9% animate, 82.4% given, topical	Walker et al. (2024: 308, 315)
English	69% human	Everett (2009: 9)
English, conversational data	76% pronominal, 72% human	Fox (1995: 161, 165)
Swedish, conversational data	92% pronominal	Dahl (2000: 43)

Table 1: Frequencies of the animate and/or pronominal input in the S role

If we extended the universal in (4) based on the unexpectedness of the role-reference association, given the frequency distributions of the input properties in Table 1, one might expect also the S argument to prefer differential marking cross-linguistically: the indefinite and/or inanimate S should be marked while the definite and/or animate S should remain unmarked. Given the frequency biases it should behave like a mirror image to P.

However, differential S marking systems are very rare. Moreover, those that do exist do not tend to be primarily based on unexpectedness of the role-reference combination but are mostly found with existential predicates or are based on semantic factors such as (non)volitionality of the S participant (de Hoop & Malchukov 2007, Arkadiev 2008b) as the following example shows.

(14) Batsbi (Tsova-Tush, Nakh-Daghestanian; de Hoop & Malchukov 2007: 1641)

a. *As* *wože*
1SG.ERG fall.PST
'I fell down (it was my fault).'

b. *So* *wože*
1SG.NOM fell
'I fell down.'

[5] Exceptions to this universal are often areally clustered, e.g. in North Africa, see Dixon (1994: 63-65), Handschuh (2014), or determined genealogically, e.g. in the conservative Indo-European languages such as Baltic.

It is evident that the marking of the S argument does not cross-linguistically tend to be based on the (un)expectedness of its role-reference association. Formally, the universal in (4) is not even supposed to apply to the S argument because this universal is restricted to situations in which at least two coding options are available and it explains how these tend to be distributed. Methodologically, however, it is important to note that the universal in (4) does not account for the cross-linguistic trends in the marking of S. Hence, preference should be given to a more powerful explanation that would be able to account for argument coding in all functional constructions: in transitive, ditransitive and monovalent intransitive constructions.

The tendency for the S argument to be morphologically unmarked most of the time and in most of the languages is straightforwardly explained by ambiguity avoidance as has been repeatedly stated in the literature earlier: lack of ambiguity according to (2) – given that there is no alternative role interpretation – is the explanation for the general lack of marking (*inter alia*, Arkadiev 2008a: 152, de Hoop & Lamers 2008).

Now we are in the position to reconcile both, the universal in (4) and ambiguity avoidance. By definition, (4) does not apply to situations in which ambiguity need not (the S role) or cannot (the A/R ambiguity) be resolved by the cues from the input and in which the coding is consistent rather than differential. This is unproblematic. However, it discards ambiguity avoidance as a pressure constraining argument marking in general.

I suggest that the role-reference association in (4) is subordinate to ambiguity avoidance, the latter being the major pressure affecting the way arguments are marked. Specifically, under no ambiguity (the S role), overt coding is generally dispreferred, while under consistent ambiguity (A vs. R), non-differential coding is generally preferred. It is only the configurations that lie between these two extremes that adopt differential marking. Here, the role-reference association in (4) applies to determine the presence vs. absence of marking. I summarize this in Table 2:

No ambiguity potential	no coding	S
Low ambiguity potential, cues from the input available	differential coding	A vs. P
High ambiguity potential, no cues from the input available	consistent coding	A vs. R

Table 2: Ambiguity potential vs. the presence of coding

I claim that ambiguity avoidance is a universal driving force that affects the way arguments are coded in functional constructions, in the way Table 2 describes.

The role-reference association in (4) only applies to functional constructions in which ambiguity may be resolved by cues from the input. In the remainder of the section, I provide more evidence from other functional bivalent constructions in support of the major role of ambiguity avoidance as per Table 2. In Sections 3 and 4, I argue that the role-reference association (4) itself also stems diachronically from ambiguity avoidance.

2.4 Comparative constructions

Comparative constructions are functional constructions because they are required by the grammatical category of comparison. Comparative constructions are bivalent by definition as they involve two participants: the *Standard* (of comparison) and the *Comparee*. Importantly, similarly to A and R participants in ditransitive constructions, comparative constructions do not normally involve configurations in which properties of any of the two arguments may provide a cue for the identification of any of both roles. Speakers mostly compare humans with humans,

cf. *X is taller/cleverer/younger than Y* as in (15), and inanimates with inanimates, cf. *X is longer/cheaper/more robust than Y* as in (16).

(15) Uzbek (Turkic; Stassen 2013)
ɔtam *u* *ɔdam-dan yɔš*
father.my that man-**from** young
'My father is younger than that man.'

(16) Maasai (Nilotic; Stassen 2013)
sapuk *olkondi* *to lkibulekeny*
big hartebeest **to** waterbuck
'The hartebeest is bigger than the waterbuck.'

Like with the R argument of ditransitive constructions and unlike with the P argument of transitive constructions, the Standard argument of comparative constructions is never found in a differential marking system in which one of the options would be no marking at all, involving neither morphological marking nor word order. Stassen (2013) distinguishes four comparative-construction types – *locational, exceed, conjoined* and *particle* comparative constructions – all of which involve dedicated marking that disambiguates the Standard from the Comparee (mostly morphologically but sometimes via word order).

Comparative constructions thus lend support to the universal distribution of argument marking suggested in Table 2 above: high ambiguity potential correlates here strongly with consistent, non-differential argument marking.

2.5 Locational bivalent constructions

By contrast, locational constructions sometimes involve differential Place marking. Differential Place marking follows differential P marking in that a version of (4) applies here too:

(17) Differential Place marking (Haspelmath 2019: 315)
Deviations from usual associations of role meanings and properties of referring expressions tend to be coded by longer grammatical forms.

For example, in French, street names are systematically zero-marked in locational constructions (18b) as opposed to other locations (18a):

(18) French (Haspelmath 2019: 318)
a. *On se rencontre dans le parc.*
1PL REFL meet in the park
'We meet in the park.'

b. *On se rencontre Rue Molière.*
1PL REFL meet Rue Molière
'We meet in Rue Molière.'

In locational constructions, there are also two arguments and accordingly two roles: the Figure and the Ground. Yet, in contrast to comparative constructions, the input may provide reliable

cues for the identification of the Ground role in locational constructions because the Ground is rarely an animate referent and in most cases semantically a location, for example, a street as in (18b). Haspelmath (2019) formulates the scale in (19) which indicates the likelihood of an input of the Ground role to be left unmarked in a locational construction.

(19) Spatial-reference scale (Haspelmath 2019: 323)
place name/topo-noun > common inanimate noun > human noun

Items that are on the left edge are more likely to be locations than items further to the right. The underlying principle in (4) and in (17) is the same: in both instances, the properties of the argument input – such as an inanimate noun as the P argument or a place name as the Ground argument – help to identify the role and thus allow to spare argument marking. By contrast, input properties such as an animate noun as the P argument or as the Ground argument are likely to create ambiguity with the subject and therefore require overt marking. Locational constructions also have another input cue to identify the roles: the Ground referent generally tends to be larger than the Figure referent.

Thus, I suggest that the reason for a fairly frequent trend for differential Place marking of some sort is similar to differential P marking: both participants are disambiguated most of the time via the properties of their input. Since both (4) and (17) are instantiations of the same phenomenon (Haspelmath 2021a), in what follows, I will jointly refer to these as to the *input-role association*.

2.6 Summarizing the synchronic evidence

I summarize the discussion so far in Table 3, drawing on Table 2:

No ambiguity potential	no coding	S
Low ambiguity potential, cues from the input available	differential coding (or even no coding)	A vs. P, Figure vs. Ground
High ambiguity potential, no cues from the input available	consistent coding	A vs. R, Comparee vs. Standard

Table 3: Ambiguity potential vs. the presence of coding

The prevalence of argument marking increases with the rise of ambiguity degree. A correlation like this cannot be accidental and, therefore, ambiguity avoidance must be a major pressure for argument marking in functional constructions.

What is more, the cross-constructional trends summarized in Table 3 may even sometimes explain the variation within a specific construction in a particular language if that construction allows for variability at all. Thus, if specific subtypes do not cues from the input they may behave against the general rules of differential object marking in the language. García García (2007) discusses special instances of the Spanish transitive construction in which the P argument is marked with the DOM marker *a* even if it is inanimate – something that is otherwise non-licit in Spanish (only animate and specific Ps are marked). Yet, most frequently this happens with lexical verbs which – in parallel to comparative constructions – do not allow for any cues for role resolution from the input of A and P and from the action itself, which is likewise symmetrical. These are verbs such as *acompañar* ‘accompany’ (20), *calificar* ‘qualify’ (20), *seguir* ‘follow’ (21), *preceder* ‘precede’ (22), *sustituir* ‘subtitute’ (23), *supercar* ‘outrank’

(24), *reemplazar* ‘replace’, *imitar* ‘imitate’ (García García 2007: 82, Torrego Salcedo 1999: 1788):

(20) *Un adjetivo acompaña / califica **a** un sustantivo.*
an adjective accompanies / qualifies **DOM** a noun
‘An adjective accompanies/qualifies a noun.’

(21) *Los días siguen **a** las noches.*
the days follow-3PL **DOM** the nights
‘The days come after the nights.’

(22) *El uno precede **a**l dos.*
the one precedes **DOM**-DEF two
‘The one precedes the two.’

(23) *En esta receta, la leche puede sustituir **a**l huevo.*
In this recipe the milk can-3SG replace **DOM**-DEF egg
‘In this recipe, egg can be replaced by milk.’

(24) *Esta cuesta supera **a** aquella.*
this slope outranks **DOM** that
‘This slope outranks that one.’

3. Diachronic evidence for ambiguity avoidance and automatization effects

I have claimed above that the input-role associations as per (4) and (17) coherently explain the trends within the transitive and locational constructions which have high cue reliability of the input, however, they cannot account for the lack of argument marking with S and for the consistent argument marking in comparative and ditransitive constructions. Only if ambiguity avoidance is considered the major pressure, are the universal trends in argument marking across all functional constructions accounted for as per Table 3.

However, one problem remains. Differential marking in monotransitive constructions and in locational constructions does not directly correlate with ambiguous contexts but rather is “automatized” along specific input types such as animacy, definiteness, locational noun, etc. (Arkadiev 2008a; Haspelmath 2021b). For example, Kozhanov et al. (forthc.) explored the relationship between ambiguity and DOM marking in a representative corpus of South Macedonian. In total, they annotated 717 sentences with a direct object for contextual ambiguity and found that only 3 of these were contextually ambiguous given the large activation window of the entire discourse. Nevertheless, 235 sentences exhibited the differential object marker *na* (originally meaning ‘to, on’). Thus, a typical DOM system such as that of South Macedonian is not conditioned by contextual ambiguity between A and P given the entire previous discourse as activation window (see also similar argumentation on Nepali in Schikowski & Iemmolo 2015).[6] Haspelmath (2021b) also shows that the input-role associations better explain cross-linguistic preferences of DOM systems than ambiguity avoidance. At the same time, quite a

[6] This DOM system is constrained by such factors as animacy and definiteness and is in this way a typical DOM system.

few examples above (§2.1) and in Seržant (2019) do attest ambiguity avoidance as a synchronic factor in DOM systems at least for specific input types or configurations of A and P. So, what is the relation of ambiguity avoidance and the input-role association in (4) (as per Haspelmath 2019, 2021a, 2021b)?

I suggest that the relation is diachronic, i.e., that ambiguity avoidance leads to emergence of the input-role association (4) for DOM systems (Seržant 2019). This happens via automatization of originally occasional, online marking choices employed to avoid ambiguity in a given context. The reason for the automatization lies in the minimization of production costs (see also below §4). Specifically, the context-dependent, global ambiguity avoidance in (1) with the activation window of the entire discourse is costly in production and perception. It requires whole-utterance planning and online decisions on the part of the speaker whether or not to use the marker in the given utterance. Likewise, it is costly for the hearer since ambiguous NPs (such as German *die Frau* 'DET.NOM=ACC woman') – if placed clause-initially – can only be interpreted by the hearer once enough context has been provided, and not incrementally (Bornkessel-Schlesewsky & Schlesewsky 2014: 107). Seržant (2019) suggests, following previous work (Aissen 2003, Zeevat & Jäger 2002, Jäger 2004, Malchukov 2008: 208, 213), that it is for this reason that ambiguity-avoidance tends to be automatized and thereby marking choices become automatized as well – a development that turns ambiguity avoidance with the activation window of the entire discourse into a local function, in which solely the nominal at issue is the activation window.

Automatization of marking choices makes speech production and comprehension computationally – but not articulatorily – more efficient. Roughly speaking, instead of assessing every single utterance on whether or not it is ambiguous given the entire discourse, those input types that are more often experienced by speakers to occur in ambiguous contexts are disambiguated immediately and automatically and regardless of whether or not there is actually any need for disambiguation. Accordingly, since animate and/or definite Ps are most often deemed confusable by speakers as to whether they represent the A or P role, it is these input types that are affected by emergent DOM systems first. In other words, the input-role association such as (4) is the result of automatization of the effects of ambiguity avoidance as in (1), via narrowing down the activation window to the very nominal in question. The locally-based role marking along (4) is more efficient in production and comprehension because it allows for more reliable incremental processing of the utterance.

3.1 Diachronic evidence for ambiguity avoidance from synchronic variation

I now turn to the evidence that supports this diachronic explanation (drawing on Seržant 2019). DOM systems may evolve via emergence and subsequent expansion of a marker or via differential reduction of a P marker with specific input types. The second developmental path is found with the DOM in English where only personal pronouns (except for *it*) are marked for P whereas all other NP types are not. Historically, the English DOM must have developed from the Proto-Germanic accusative and dative with non-differential, across-the-board object marking, as the historical-comparative evidence unequivocally suggests. For reasons of space, I discuss below only the first developmental path, i.e., the emergence and expansion of a new DOM marker.

Differential P markers emerge from various types of sources. The most prominent one is topic constructions (Dalrymple & Nikolaeva 2011). Topic constructions have obviously their own constraints on their occurrence which are not related to DOM. For example, non-subject

personal pronouns are most likely to be affected by these at the very first, since personal pronouns are most frequently topics (subject personal pronouns are already sufficiently marked as topics by virtue of being subjects and need no additional marking). Typically for any grammatical change, the new grammatical function starts imposing new trends on usage. One phenomenon to be illustrated below is that earlier topic markers start also serving the disambiguating function.

First, there is some synchronic evidence that may be interpreted diachronically. In addition to some very few languages like Yongren Lolo in (7) and (8) above in which the DOM system is primarily driven by ambiguity avoidance, quite a few languages show ambiguity avoidance effects on DOM, restricted to specific input types and/or scenarios whereas the other input types and/or scenarios are either obligatorily marked or obligatorily unmarked along (4).

Here, the observation like the above in Table 3 holds. However, now, it pertains to specific input configurations within transitive constructions (and not to entire constructions as in Table 3): those input configurations that have a higher role ambiguity potential and which provide no cues for role identification (e.g. an animate P) tend to avoid optionality in marking whereas those inputs that do provide cues for role identification (e.g. inanimate P) allow for optionality and the DOM marker is only used to avoid ambiguity here.

Consider the DOM system of the rural variety of Donno Sɔ, as described in Culy (1995). In this language, only animate and definite Ps must be marked (high ambiguity potential). However, this rule may be extended to avoid ambiguity (Culy 1995: 52). Example (25) exhibits the DOM marker even though the nominal is indefinite because otherwise the sentence would become confusing.

(25) Donno Sɔ (Dogon; Culy 1995: 53)

Wɛzɛwɛzɛginɛ	*yaana*	*po-ñ*	*don*	*wo mɔ ni tɛmbɛ.*
crazy.person	woman	large-DOM	place	3SG PS at found

'A crazy person found a large woman at his/her place.'

Another example comes from Telkepe Neo Aramaic (26). In this language, the DOM flag *ta* is primarily conditioned by the (secondary) topic role of the object. However, marginally, it may also be used for ambiguity avoidance as well. So, in (26a), the full NP *gor-a* 'her husband' may be either interpreted as subject indexed as the masculine singular on the verb alongside the object index *-lə* 'him', or, alternatively, *-lə* may be interpreted as indexing the full NP *gor-a* thus interpreted as the object. By contrast, (26b) does not allow the interpretation of *gor-a* as the subject due to the DOM flag; moreover, as Coghill (2015: 354) writes, her "informant insisted on the necessity of using *ta*" in (26).

(26) Telkepe (Aramaic, Semitic; Coghill 2015: 354)

a.

kəm-māxe-lə	*gor-a.*
PST-HIT.3MS-OBJ.3MS	husband-her

?'He hit her husband.' or 'Her husband hit him.'

b.

kəm-māxe-lə	*ta*	*gor-a.*
PST-hit.3MS-3MS.OBJ	DOM	husband-her

'He hit her husband.' but not *'Her husband hit him.'

Another similar DOM system is the one of Khwe (Khoisan). In this language, proper names must obligatorily be marked with *à/-à* in the object position; additionally, this marker encodes contrast and/or focus on the NP (Kilian-Hatz 2006: 82–83). At the same time, the marker may also be used in contexts in which the distinction between A and P would have been impeded, for example, when both arguments are animate and topical (Kilian-Hatz 2006: 82–83):

(27) Khwe (Khoe: Southern Africa; adapted from Kilian-Hatz 2006: 83)

a. *Tcá* *tí* *à* *kx'ő̄á́.*
2SG.M 1SG DOM wait
'You have to wait for me!'

b. *Yàá!* *Cáò* *à* *tí* *kyá-rá-hã!*
yes 2DU.F DOM 1SG love-ACT-PST
'That's it! I love you two (women)!'

Occasional evidence for ambiguity avoidance constraining the use of DAM markers within a specific input subtype in addition to hard constraints along (4) with all the other input subtypes have been reported for Malayalam (Dravidian), Catalan, Spanish, Hup (Nadahup), Awtuw (Sepik), Tamasheq (Afro-Asiatic), Guaraní (Tupian), Telkepe (Semitic, Aramaic), Mam (Mayan), see the references in Seržant (2019), possibly also in Nepali (Schikowski & Iemmolo 2015), Eastern Armenian (Strauß 2021), Bininj Gun-Wok (if the pronominal prefix system leaves the ambiguity between A and P unresolved, the instrumental marking of the A is used here, Evans 2003: 139-140).

3.2 Diachronic evidence for ambiguity avoidance from instability of DOM systems

There is also diachronic evidence that the development of DOM is constrained by ambiguity avoidance. Already Bossong (1985a: 117) assumed that the emergence of DOM is primarily due to the need to discriminate between A and P. Differential marking of P is cross-linguistically strongly preferred as opposed to non-differential marking (Sinnemäki 2014). While marking the P differentially is a very strong cross-linguistic trend, differential P marking systems in themselves are extremely unstable diachronically when it comes to the exact conditions determining the presence vs. absence of DOM markers.

First, this high instability is seen in the fact that closely related languages often exhibit etymologically unrelated DOM markers. Thus, LaPolla (1992: 3-4) lists 50 languages that "cover almost the entire Tibeto-Burman geographic area", all of which have different and evidently etymologically unrelated DOM postpositions. In other words, differential marking is preferred but a specific differential marking system is not stable. Likewise, many Iranian languages have lost the oblique/direct case to mark P in the present tense but then subsequently developed DOM systems (based on different markers) to various degrees and, crucially, from different sources, cf. postfix / postposition *-(r)ā* in Persian, prefix/preposition *az-* in the Shughni-Roshani group of Pamir languages or the predix *ž-* in Yazghulami (Wendtland 2008: 421). Other markers are also found, such as *ba*, *ei*, *-ā*, e.g., incipiently in Bashkardi (Korn 2017). I conclude that, if DOM systems were diachronically stable, closely related languages would tend to exhibit etymologically related DOM markers across families and identical conditioning.

I turn now to presenting evidence for the second point of instability, namely, to the instability and variation of conditions. Even if the DOM markers are related within a group of

closely related languages, these markers themselves nearly never follow the same constraints across these languages. This likewise suggests that there is no universally stable attractor state for DOM. For example, in the Khoe subbranch of Khoisan languages, the differential object marker *-(ʔ)à* emerged from focus marker. Crucially, different Khoe languages attest different degrees of progress. While most Khoe languages tend to mark definite and specific objects, in the Khoekhoe subbranch this marker has been extended to all objects, e.g. in Nama-Damara (McGregor 2018: 271). Likewise, the DOM marker *-re* in different Tucanoan languages follows quite distinct constraints along (4): while personal pronouns are obligatorily marked in Barasano, Desano, Tucano, Cubeo, Coreguaje and in Siona, when it comes to nominal Ps the constraints are very different: Barasano and Desano are reported to depend on definiteness, Tucano and Coreguaje on specificity, but Cubeo and Siona on animacy/humanness (see Wheeler 1967, Zúñiga 2007: 219, Cook & Levinsohn 1985, Stenzel 2008: 164, Menschel, forthc.). Finally, a comparison of two diachronic stages of the same language also suggests instability. For instance, the Persian DOM marker *-rā(y)* developed from marking mostly animate Ps in Middle Persian (Key 2008: 244, cf. also Paul 2008: 152–153) to the inclusion of inanimate definite Ps in Modern Persian, in which language animacy is, therefore, no longer a factor. Likewise, different Slavic languages exhibit very different constraints on the use of DOM: while Old Church Slavic only allows the DOM flag (formally the genitive case) with proper names and rarely with some human-referring nominals, modern Russian requires the DOM marker with all animates including animals (in the relevant declensional classes). In turn, the closely related Ukrainian has extended the marker to some inanimate nominals as well. Dedicated diachronic studies of specific DOM systems, such as the diachrony of DOM across 118 Roma dialects (Kozhanov & Seržant, forthc.) and the diachrony of DOM in South Macedonian dialects (Kozhanov et al., forthc.), unequivocally show that different diachronic stages as well as different dialects diverge as to the conditions on the use of DOM markers.

Seržant (2019) claims that the continuous extension by which DOM markers keep crossing their original conditions might also be rooted in ambiguity avoidance. For example, the expansion of the DOM marker from nearly exclusively animate Ps in Middle Persian or Old East Slavic to the inclusion of inanimate Ps in Modern Persian and in Ukrainian, respectively, is not semantically straightforward. Animacy and inanimacy are rather antonymic in meaning. Thus, the general mechanisms of semantic extension cannot really explain the diachronic transition from marking only animate Ps to also marking inanimate Ps. However, given the synchronic evidence on ambiguity avoidance in specific input subtypes provided immediately above, I suggest that it is ambiguity avoidance that is diachronically responsible for incipient uses of the DOM marker with ever new input types on prominence scales: in addition to some input types in which the DOM marker is used conventionally (e.g. animate Ps), the same marker may be occasionally used with inanimate Ps with the disambiguating function; once it is found with inanimate Ps (under ambiguity) it may then start becoming the conventional marking of some subsets of inanimate Ps (e.g. definite inanimate). Accordingly, I suggest that it is the effects of ambiguity avoidance that contribute to instability of the constraints on DOM systems, because – in contrast to semantic extension – ambiguity avoidance is not associated with any specific input meaning (such as animacy) and can extend across semantically unrelated input subtypes, for example, from animate Ps onto inanimate Ps, or from definite to indefinite Ps. We observe this extension in examples discussed in §3.1 above.

3.3 Diachronic evidence on ambiguity avoidance from historical sources

Documented diachronies of some DOM systems also provide evidence for the role of ambiguity avoidance in their development. Very often DOM markers develop from topic-related markers and constructions (Iemmolo 2010; Dalrymple & Nikolaeva 2011) and then carry over conditions on topicality into the emerging DOM system. This is why ambiguity avoidance often cannot be clearly disentangled in such diachronies. In Slavic, however, a different developmental path for DOM is found. The new Slavic DOM system was based on gradual expansion of the genitive case onto Ps in affirmative clauses (Klenin 1983). The genitive case was already used to mark Ps in the context of verbal negation in Proto-Slavic. In contrast to the DOM markers stemming from topic markers, the genitive case did not bring with it any topic-related conditions on its use from the source construction and its expansion was not skewed by topic functions.

What the diachronic evidence from Russian shows is that the new DOM marker (the genitive case) expanded only onto those input types that have lost the original morphological distinction between A and P, i.e. between nominative and accusative cases (philologically profound evidence in Krys'ko 1994; 1997). Importantly, only those animate nouns and pronouns were affected by the new DOM marker which did not differentiate between the nominative and the accusative anymore.[7] Evidence for this is abundant: (i) The new DOM marker did not replace the old accusative in the *a*-declension because, in this declension, a non-zero accusative (*-ǫ* > *-u*) was retained. (ii) The first nominal input types affected by the new DOM marker were proper names while personal pronouns generally remained unaffected by it. This is quite atypical of DOM systems that tend to expand along the referential scales and start with personal pronouns. The reason for this is that personal pronouns had not undergone the phonetic conflation of the nominative (e.g. *azъ* '1SG.NOM') and the accusative (e.g. *mę* '1SG.ACC') and hence were not in need of disambiguation. (iii) The plural of the *o*-declension – in contrast to the singular – did retain the distinction between the old accusative (*-y*) and the old nominative (*-i*), which is why the old accusative was not replaced by the new DOM marker here. Only later, between the 14th and 16th c., were both the old accusative plural and the old nominative plural conflated into *-y* and thereby became indistinguishable in Old Russian. Precisely from this period onwards, the new DOM marker (the genitive plural) started to be used in the plural of animate nouns as well (Krys'ko 1994: 144). Interestingly, the potential ambiguity between A and P were not exclusively resolved via dedicated P marking (genitive-accusative) in Slavic but also, (iv) with the third person pronoun *j-*, by lack of a nominative form altogether (various demonstratives were used instead here). This pronoun acquired the new genitive-based DOM marker (*jego*) much later than the relative pronoun *ji-že* (both are etymologically related) because the relative pronoun *ji-že* had a nominative form as well, which was indistinguishable from the accusative form. Thus, the relative pronoun acquired DOM very early while the related third-person pronoun much later. (v) Finally, the Old Novgorodian dialect of Old Russian retained the marked nominative in *-e* with a subset of *o*-declension nouns in its earliest period while having zero in the accusative. As Krys'ko (1993) shows, it is this dialect in which the new DOM marker appeared on nouns later than elsewhere in Old Russian. This happened only once the dedicated nominative in *-e* disappeared here as well and turned into zero.

[7] The *i*-declension of Modern Russian also supports this. Even though in present-day Russian, there is no more distinction between A and P with animate nouns here (e.g. *mat'* 'mother'), Early Slavic did have a distinction between A and P forms here (Old Church Slavic *mat-i* NOM.SG vs. *mat-er-ь* ACC.SG).

To summarize, in the history of Slavic, the new DOM marker chronologically affected a specific input type only after the latter had lost the distinction between A (nominative) and P (accusative) (Dobrovsky 1834: 39, Krys'ko 1994: 156, Tomson 1908, 1909).

3.4 Diachronic-typological evidence

Finally, the cross-linguistic diachronic evidence at hand supports the observations from the relative chronology of Russian. I have collected a dataset with diachronic data on 53 DOM systems based on a convenience sample of languages across the world, comprising 14 families and three macroareas (Africa, Eurasia and South America), cf. Figure 1.

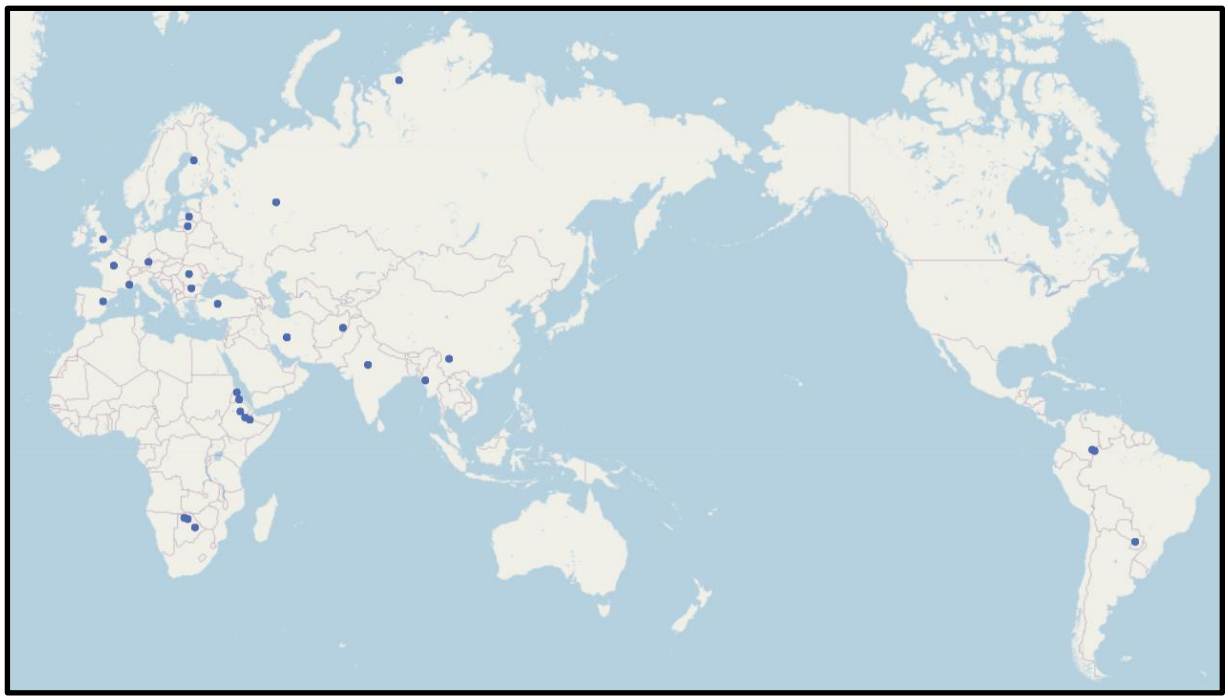

Figure 1. A sample of diachronies of DOM systems in different languages

This is work in progress. The sample contains descriptions of diachronies of etymologically unrelated DOM systems as found in the literature as well as comparative evidence from closely related languages. Thus, despite genealogical biases, most of the collected diachronies are independent from each other.[8]

Crucially, more than two thirds (72%, 38 out of 53) of the DOM systems in my sample developed DOM markers in situations in which – like in the history of Slavic – no other object marker was available to disambiguate P from A prior to the emergence of the new DOM marker.

4. Bringing the diachronic and typological evidence for ambiguity avoidance into the psycholinguistic context

There is obviously no ambiguity for the speaker in what they intend to communicate. Nevertheless, it is the speaker who can take measures for ambiguity avoidance. This additional

[8] I cannot, however, exclude some areal effects.

effort is made to maximize communicative success with the comprehender. Thus, ambiguity avoidance is genuinely a communicative effect of *audience design* in speech production. Audience design is "tailoring utterances to addressee needs" as Wardlow Lane & Ferreira (2008: 1) put it. It implies that the speaker not only economizes on their production costs due to limitations of the available resources (working memory, computational resources, articulatory efforts, etc.) but also bears in mind their communicative goals and, therefore, may implement additional strategies in order to maximize the probability of being correctly understood (Kurumada & Jaeger 2015). Thus, audience design may lead to some degree of redundancy and, thereby, to increased effort on the part of the speaker.

Roughly, there are two types of audience-design strategies: *utterance-specific audience-design strategies* and *generic audience-design strategies*.[9] Generic strategies aim at increasing the communicative effect without the speaker having to take into account the specifics of the particular utterance (cf. *generic-listener adaptations* in Dell & Brown 1991). Generic strategies are common in various domains of language use. They are found in avoiding structures that are difficult to parse, in adjusting the speech to the listener's age, in speaking louder in a noisy space, etc. (Dell & Brown 1991). By contrast, utterance-specific audience design implies *ad hoc* enhancement of comprehensibility of a particular utterance if this utterance is perceived by the speaker as ambiguous at the time of production.

When it comes to production costs, there is an important difference between the two types of audience design (Jaeger & Buz 2018: 58): Generic strategies do not require the speaker to monitor the comprehender's perspective with regard to the comprehensibility of the message being produced. The speaker just "automatically" applies a generic strategy about which they know from previous experience that it generally increases the probability of being understood regardless of the properties of the current utterance such as the lexical input. Utterance-specific audience-design strategies, in turn, are perhaps easier in articulation because they allow the speaker to save effort by avoiding redundancy, but at the same time they are cognitively resource-demanding. Utterance-specific audience-design strategies require the speaker's online monitoring of the production with respect to comprehensibility of the utterance being produced, which is costly.

Psycholinguistic evidence for generic audience design is abundant while evidence for utterance-specific audience design is scarce and not uncontroversial (Jaeger & Buz 2018: 58-59). For example, the alleged preference for the use of the English complementizer *that* in syntactically ambiguous relative and complement clauses did not find support. Jaeger & Buz (2018: 58-59) state that many of the specific audience-design effects claimed in earlier research to have an effect on syntax revealed themselves later as utterance-independent, generic audience-design strategies.

When it comes to argument marking, an utterance-specific audience-design strategy to avoid ambiguity in argument role identification would require the speaker to first assess the entire context given the lexical input (e.g. *I* vs. *snake* in (7) or *tree* vs. *house* in (8)) with respect to whether or not it provides sufficient disambiguation and, only if deemed necessary, produce the disambiguating marker. Such an utterance-specific strategy requires the speaker to rely on pre-planning of larger chunks which have to be stacked up in the working memory before the

[9] I use the notion *strategies* generally, meaning both specific and conscious actions of language users (e.g. speaking louder in a noisy environment) but also less conscious or unconscious actions, i.e., strategies that are conventionalized in the grammar but which are the result of adaptation of the grammar under the pressure for successful information transmission.

entire cue can go to articulation. These processes strain working memory which has very limited resources (ca. 3-sec. window of speech, Pöppel 2009, Wittmann 2011, Ferreira 2019, up to 8 sec., see Falandays et al. 2020). Overloading working memory on the lexical and category level hinders ahead-planning (Christiansen & Chater 2016: 5, neurolinguistic evidence in Swaab et al. 2003). Moreover, pre-assessment of potential ambiguities precludes incremental production and comprehension despite the general strive for incremental production (Bornkessel-Schlesewsky & Schlesewsky 2014, Jaeger & Tilly 2011) and incremental comprehension (Christiansen & Chater 2016). In sum, utterance-specific ambiguity avoidance in argument marking is very costly.

The claim that utterance-specific ambiguity avoidance strategies are too costly also finds support in typological evidence: these strategies indeed are cross-linguistically highly dispreferred. Argument marking systems that are primarily based on utterance-specific ambiguity avoidance such as the DOM system of Yongren Lolo (cf. (5) and (6) above) are extremely dispreferred cross-linguistically. Somewhat more stable cross-linguistic evidence for utterance-specific ambiguity avoidance is found only in specific scenarios and input-role combinations. Such is the combination of both A and P being indefinite and animate in Donno Sɔ, cf. ex. (20) above, while definite animate Ps are automatically, i.e. generically, marked in this language (see the examples above from Telkepe, Khwe and other languages, see §3.1).

Drawing on earlier research, I claim that languages adapt and respond to high production costs of utterance-specific ambiguity avoidance in argument marking by developing context-independent, generic coding strategies to resolve potential rather than actual ambiguity (see Durie 1995: 279, Malchukov 2008: 208, Seržant 2019), a subtype of feedforward, generic audience design, cf. Ferreira (2019). Above (§3), I have provided synchronic and diachronic evidence for the development away from utterance-specific ambiguity avoidance to generic ambiguity avoidance. I have claimed that this happens via automatization of originally occasional, online marking choices driven by the need to avoid ambiguity. Speakers automatize their choices on the basis of previous communicative successes and failures (Durie 1995: 281, Jaeger & Buz 2018: 59, Buz et al. 2016: 80). The reason for automatization lies in the minimization of production costs. Automatization allows the speaker to plan the encoding of the argument given only the smallest activation window, namely, just the very lexical input of the argument, without the need to pay attention to whether or not the entire clause let alone the entire discourse would sufficiently disambiguate the argument. This automatization of marking choices leads to generic audience-design strategies to avoid ambiguities in role-input combinations. The historical change from utterance-specific to generic is a process that has been discussed under the notion *over-generalization* or *overkill* already in Durie (1995: 278).

The suggested diachronic reduction of the activation window to just the nominal in question is efficient in incremental speech comprehension. It has been repeatedly argued that speakers tend to interpret chunks of speech right after these were uttered without having the entire context of the clause at disposal.

In Table 3 above (repeated here as Table 4 for convenience), I summarized generic audience-design strategies in argument marking under three types of strategies: (i) generally *no coding* regardless of the context, (ii) generally *consistent coding* regardless of the context and (iii) *differential automatized* coding generalized across all members of a particular input type (animate, definite, etc.):

No ambiguity potential	no coding	S
Low ambiguity potential, cues from the input available	differential coding (or even no coding)	A vs. P, Figure vs. Ground
High ambiguity potential, no cues from the input available	consistent coding	A vs. R, Comparee vs. Standard

Table 4: Ambiguity potential vs. the presence of coding

Crucially, all these strategies are utterance-independent and require a very small activation window not going beyond the very nominal being produced. Yet, these are the strategies that are preferred cross-linguistically. Thus, differential coding systems are mostly based on the properties of the input (e.g., definiteness and/or animacy with DOM) rather than on larger activation windows (Sinnemäki 2014, Witzlack-Makarevich & Seržant 2018). In effect, generic audience design unavoidably leads to redundancy and, specifically, DOM systems exhibit mismatches between the inferability of the roles given the specific context and their coding lengths in this context (Seyfarth 2014, Sóskuthy & Hay 2017).

These facts also now explain the mechanism underlying Haspelmath's (2021b: 125) universal in (4), repeated here as (23) for convenience:

(28) The role-reference association universal (Universal 1) (Haspelmath 2021b: 125)
„Deviations from usual associations of role rank and referential prominence tend to be coded by longer grammatical forms if the coding is asymmetric."

The speaker memorizes which input types (i.e. the "deviations from usual" in Haspelmath 2021b: 125) more often lead to ambiguity between arguments and turns this communicative experience into a generic audience-design strategy to avoid ambiguity.

On the population level, we observe this change when incipient DOM systems follow only soft constraints but then, with time, turn into DOM systems with hard constraints. For example, the diachronic corpus-based study of a very recent DOM system in South Macedonian dialects (Kozhanov & Seržant, forthc.) shows that texts from the 19th c. allow for more variation and optionality than the records from the 20th century which primarily attest the DOM marker (the preposition *na* lit. 'on, to') only with animate and definite input. Indirect evidence from dialectal variation of DOM systems in Romani dialects across Europe suggests the same diachronic "fixation" towards hard constraints (Kozhanov & Seržant, forthc.). Another example is Early Slavic (Old Church Slavic) which optionally attests the DOM marker with a few animate nouns but then fixes animacy as a hard constraint in the modern languages. More examples can be adduced.

5. Conclusions

Ambiguity avoidance has been discussed with respect to A vs. P in DOM systems or T vs. R in ditransitive constructions (*inter alia*, see Siewierska & Bakker 2009: 292), but less attention has been paid to ambiguity avoidance in other bivalent functional constructions. I have reviewed a set of functional constructions and have shown that ambiguity avoidance is the major pressure conditioning the presence vs. absence of argument marking in grammatical roles such as A, R, P, Figure, Ground, Standard or Comparee.

I have also reported diachronic, diachronic-typological and synchronic-typological evidence for a strong preference for ambiguity avoidance as a generic audience-design strategy in argument marking as opposed to utterance-specific audience design. This evidence converges with the evidence presented in psycholinguistic studies. The latter, however, are primarily based on a few Western European languages and might therefore be in need of cross-linguistic validation, which this paper does. As Seifart et al. (2018: 5723) put it with respect to psycholinguistic research in general, "[m]ost such work is still largely based on educated speakers of a small number of mostly Western European languages, and it remains unclear whether findings generalize beyond this." (similarly Jaeger & Norcliffe 2009, MacDonald 2013, Norcliffe et al. 2015). Specifically, I have claimed that ambiguity avoidance explains the following three strategies of argument marking in Table 4 repeated here for convenience as Table 5:

No ambiguity potential	no coding
Low ambiguity potential, cues from the input available	differential coding (or even no coding)
High ambiguity potential, no cues from the input available	consistent coding

Table 5: Ambiguity potential vs. the presence of argument coding in functional constructions

In this context, I proposed to slightly reconcile Haspelmath's (2021a, 2021b) account that is based on (un)expectedness of input-role associations with ambiguity avoidance. While it is undeniable that unexpected role-input associations tend to be marked cross-linguistically as opposed to the expected ones, ambiguity avoidance likewise shows effects across different constructions, as I have demonstrated above. I claim that ambiguity avoidance is a strong pressure that diachronically leads to the input-role associations like (4) or (17) via automatization of marking choices in those constructions in which strong cues from the argument input are available as per Table 4. I have also argued that automatization in, e.g., transitive constructions itself diachronically proceeds from those input configurations which have a higher ambiguity potential to the input configurations with a lower ambiguity potential. The reason for automatization of marking choices is that it makes speech production and comprehension computationally (but not necessarily articulatorily) more efficient. More generally, signaling unexpectedness of an input-role association is directly related to resolving ambiguity. Unexpected role-input associations tend to be marked because these are a potential source for ambiguities. Narrowing down the activation window (2) from clause-level to just the very nominal in question better aligns with the incremental nature of language production and comprehension.

These results resemble the psycholinguistic evidence on audience-design strategies (summarized in Jaeger & Buz 2018: 58-59) such that languages are shaped by generic audience-design strategies and the evidence for utterance-specific audience-design effects is very scarce. Diachronically utterance-specific audience-design strategies tend to be abandoned in favor of generic audience-design strategies. Generic audience-design strategies, in turn, unavoidably bring some redundancy with them. For example, since contextual ambiguity of argument roles in transitive constructions is in fact quite rare (Kozhanov et al. forthc.), most of the time (differential) argument marking can be said to involve a considerable amount of redundant coding. However, as has been suggested in Seržant & Moroz (2022), computational production

costs overweight articulatory costs and, therefore, speakers prefer to economize on computational costs at the expense of articulatory/coding costs.

To summarize, unexpectedness has undeniably an important effect on marking information in general (DeLong et al. 2005; Xu et al. 2023, among many others). However, in constructions which involve competition of a limited set of values (mostly two), ambiguity avoidance is an important factor constraining the presence vs. absence of marking.

Finally, I note that ambiguity avoidance and unexpectedness are not the only factors constraining the marking rules of arguments in functional bivalent constructions. Other factors such as the source functions of the marker (e.g. if a DOM marker develops from a topic marker), different discourse situations and interactional factors, language contact, etc., can also affect marking conditions.

Acknowledgements

This study has been supported by the German Science Foundation (Deutsche Forschungsgemeinschaft, Project-ID 317633480 – SFB 1287 "Limits of Variability in Language. Cognitive, Computational, and Grammatical Aspects", subproject "Limits of Variability of Valency Class Systems")

References

Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. Natural Language and Linguistic Theory 21(3). 435–483. DOI:10.1023/A:1024109008573

Arkadiev, Peter. 2008a. Chapter 7. Differential Argument Marking in Two-term Case Systems and its Implications for the General Theory of Case Marking. In: H. de Hoop and P. de Swart (eds.), *Differential Subject Marking*. Springer. 151–171.

Arkadiev, Peter. 2008b. Thematic roles, event structure, and argument encoding in semantically aligned languages. In: Mark Donohue & Søren Wichmann (eds.), *The Typology of Semantic Alignment*. Oxford: Oxford University Press, 101–117.

Arumaa, P. 1985: *Urslavische Grammatik*. Band III. Heidelberg: Carl Winter.

Asher, R., Kumari, T., 1997. *Malayalam.* Routledge, London.

Bornkessel-Schlesewsky, Ina & Matthias Schlesewsky. 2014. Competition in argument interpretation: Evidence from the neurobiology of language. In Brian MacWhnney, Andrej L. Malchukov & Edith A. Moravcsik (eds.), Competing motivations in grammar and usage, 107–126. Oxford: Oxford University Press.

Bornkessel-Schlesewsky, Ina & Matthias Schlesewsky. 2015. Scales in real-time language comprehension: A review. In: Ina Bornkessel-Schlesewsky, Andrej L. Malchukov and Marc Richards (eds.), *Scales and Hierarchies: A Cross-Disciplinary Perspective*. Berlin, München, Boston: De Gruyter Mouton. 321-352.

Bossong, Georg 1985a: *Empirische Universalienforschung. Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Narr.

Bossong, Georg 1985b: Markierung von Aktantenfunktionen im Guaraní. In F. Plank (ed.), *Relational typology*. Berlin: Mouton DeGruyter. 1–29.

Buz, Esteban, Michael K. Tanenhaus & T. Florian Jaeger. 2016. Dynamically adapted context-specific hyper-articulation: Feedback from interlocutors affects speakers' subsequent pronunciations. *Journal of Memory and Language* 89. 68–86.

Christiansen, M.H., & N. Chater 2016. The now-or-never bottleneck: a fundamental constraint on language. *Behavioral and Brain Sciences* 39, e62.

Coghill, Eleanor 2014: Differential object marking in Neo-Aramaic, *Linguistics* 52(2), 335-364.

Comrie, Bernard. 1989. Language universals and linguistic typology: Syntax and morphology. 2nd edn. Chicago: University of Chicago Press.

Cook, Dorothy & Stephen Levinsohn 1985: Coreguaje: The Domains of Focus Markers. In: Ruth M. Brend (ed.), *From Phonology to Discourse: Studies in Six Colombian Languages.* [Amerindian Series 9]. Dallas: Summer Institute of Linguistics.

Culy, Christopher 1995: Ambiguity and Case Marking in Donno Sɔ (Dogon). In: Akinlabi, Akinbiyi (ed.), *Theoretical Approaches to African Languages.* Trento: Africa World Press. 47-58.

Dahl, Östen. 2000. Egophoricity in discourse and syntax. Functions of Language 7(1). 37–77.

Dalrymple, M. & I. Nikolaeva 2011: *Objects and Information Structure.* Cambridge: Cambridge University Press.

Dell, G. S., & Brown, P. M. 1991. Mechanisms for listener-adaptation in language production: Limiting the role of the "model of the listener". In D. J. Napoli & J. A. Kegl (Eds.), Bridges between psychology and linguistics: A Swarthmore festschrift for lila gleitman (pp. 105–130). Hillsdale, NJ: Lawrence Erlbaum Associates.

DeLong, Katherine A., Thomas P. Urbach & Marta Kutas. 2005. Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience 8*(8). 1117–1121.

de Hoop, H. & M. Lamers. 2006. Incremental distinguishability of subject and object. Case, Valency, and Transitivity. Ed. by L.I. Kulikov, A.L. Malchukov and P. de Swart. Amsterdam/Philadelphia: John Benjamins, 269-287.

de Hoop, Helen & Andrej Malchukov. 2007. On fluid differential case marking: A bidirectional OT approach. *Lingua* 117. 1636–1656.

de Hoop, Helen & Peter de Swart, eds., 2008: *Differential subject marking*. Dordrecht: Springer.

Dixon, Robert M. W. 1979. Ergativity. *Language* 55. 59-138.

Dixon, Robert M. W. 1994. *Ergativity.* Cambridge: CUP.

Dobrovskij, I. 1834: *Grammatika jazyka slavjanskogo po drevnemu narečiju.* St.-Peterburg: Tipografija departamenta narodnogo prosveščenija.

Donohue, Mark. 2005. Configurationality in the Languages of New Guinea. *Australian Journal of Linguistics* 25(2). 181-218.

Durie, Mark. 1995. Towards an Understanding of Linguistic Evolution and the Notion 'X has a Function Y'. In: Werner Abraham, Talmy Givón & Sandra A. Thompson (eds.), *Discourse, Grammar and Typology: Papers in honor of John W.M. Verhaar*. Amsterdam: John Benjamins. 275-308.

Evans, Nicholas. 2003. *Bininj Gun-Wok: A Pan-dialectal Grammar of Mayali, Kunwinjku and Kune.* (Pacific Linguistics, 541.) Canberra: Research School of Pacific and Asian Studies, Australian National University.

Everett, Caleb. 2009. A reconsideration of the motivations for preferred argument structure. *Studies in language* 33(1). 1–24.
Falandays, J.B., S. Brown-Schmidt & J.C. Toscano. 2020. Long-lasting gradient activation of referents during spoken language processing. Journal of Memory and Language 112. 104088.
Fauconnier, Stefanie & Jean-Christophe Verstraete. 2014. A and O as each other's mirror image? Problems with markedness reversal. *Linguistic Typology* 18(1). 3–49.
Fedzechkina, Mariya, Florian T. Jaeger, & Elissa L. Newport 2012: Language learners restructure their input to facilitate efficient communication, *Proceedings of the National Academy of Sciences* 109, 17897–17902.
Ferreira, Victor. 1996. Is It Better to Give Than to Donate? Syntactic Flexibility in Language Production. Journal of Memory and Language 35, 724–755.
Ferreira, Victor. 2008. Ambiguity, accessibility, and a division of labor for communicative success. *Psychology of Learning and Motivation* 49. 209–246.
Ferreira, Victor. 2019. A Mechanistic Framework for Explaining Audience Design in Language Production. *Annual Review of Psychology* 70. 29-51.
Foley, William A. 1986. *The Papuan languages of New Guinea* (Cambridge language surveys). Cambridge: Cambridge University Press.
Fox, Barbara A. 1995. The Category S in English conversation. In Werner Abraham, Talmy Givón & Sandra A. Thompson (eds.), *Discourse Grammar and Typology*. 153-178. Amsterdam: Benjamins.
Gerner, Matthias. 2008. Ambiguity-Driven Differential Object Marking in Yongren Lolo. *Lingua* 118. 296–331.
Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In: Joseph H. Greenberg (ed.), *Universals of language.* 73–113. Cambridge, MA: MIT Press.
Handschuh, Corinna. 2014. *A typology of marked-S languages.* Berlin: Language Science Press.
Haspelmath, Martin. 2013. Ditransitive Constructions: The Verb 'Give'. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) *WALS Online* (v2020.3) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.7385533 http://wals.info/chapter/105
Haspelmath, Martin. 2019. Differential place marking and differential object marking. *Language Typology and Universals (STUF)* 72(3). 313-334.
Haspelmath, Martin. 2021a. Explaining grammatical coding asymmetries: Form–frequency correspondences and predictability. *Journal of Linguistics* 57. 605–633.
Haspelmath, Martin. 2021b. Role-reference associations and the explanation of argument coding splits. *Linguistics* 59(1). 123-174.
Hofling, Charles. 2003. Tracking the deer: Nominal reference, parallelism and Preferred Argument Structure in Itzaj Maya narrative genres. Du Bois, Kumpf & Ashby (eds.), 353–384.
Iemmolo, Giorgio 2010: Topicality and differential object marking : Evidence from Romance and beyond, *Studies in Language* 34(2), 239-272.
Iemmolo, Giorgio. 2013. Symmetric and asymmetric alternations in direct object encoding. *STUF - Language Typology and Universals* 66. 378-403.
Jäger, Gerhard 2004: Learning Constraint Subhierarchies: The Bidirectional Gradual Learning Algorithm, in Blutner, R. & H. Zeevat, *Optimality Theory and Pragmatics*, Palgrave MacMillan. 251-287.

Jaeger, T. Florian & Esteban Buz. 2018. Signal reduction and linguistic coding. In E. M. Fernández & H. S. Cairns (eds.), *The handbook of psycholinguistics*. First edition. Hoboken, New Jersey: Wiley Blackwell. 38–81.

Jaeger, T. Florian & Elisabeth J. Norcliffe 2009. The Cross-linguistic Study of Sentence Production. *Language and Linguistics Compass* 3/4, 866–887.

Jaeger, T. Florian, & Harry Tily. 2010. On language 'utility': Processing complexity and communicative efficiency. Wiley Interdisciplinary Revues: Cognitive Science 2(3). 323–335.

Key, Gregory 2008: Differential Object Marking in a Medieval Persian Text. In: *Aspects of Iranian Linguistics*, (ed.) Simin Karimi, Vida Samiian, & Don Stilo. Cambridge Scholars Publishing: Newcastle upon Tyne. 227-247.

Kilian-Hatz, Christa. 2006. Topic and focus in Khwe. In: Sonja Ermisch (ed.), *Focus and Topic in African Languages.* [Frankfurter Afrikanistische Blätter 18]. Köln: Rüdiger Köppe. 69-90.

Klenin, E. 1983. *Animacy in Russian: A new interpretation*. Columbus, Ohio: Slavica Publishers, Inc.

Kozhanov, Kirill & Ilja A. Seržant. Forthcoming. Evolution and areal expansion of differential object marking in Romani. *Journal of Language Dynamics and Change.*

Kozhanov, Kirill, Ilja A. Seržant, Eleni Bužarovska. Forthcoming. Evolution of differential object marking in Macedonian dialects. *Diachronica.*

Krys´ko, Vadim B. 1994. *Razvitie kategorii oduševlennosti v istorii russkogo jazyka.* Moscow.

Krys´ko, Vadim B. 1997. *Istoričeskij sintaksis russkogo jazyka. Ob"ekt i perexodnost'.* Moscow.

Kurumada, Chigusa & T. Florian Jaeger 2015: Communicative efficiency in language production: Optional case-marking in Japanese, *Journal of Memory and Language* 83, 152–178.

LaPolla, Randy. 1992. Anti-ergative Marking in Tibeto-Burman. *Linguistics of the Tibeto-Burman Area* 15.1. 1-9.

Levshina, Natalia. 2021. Communicative efficiency and differential case marking: A reverse-engineering approach. *Linguistics Vanguard*, 7(s3): 20190087.

MacDonald, M. 2013. How language production shapes language form and comprehension, *Frontiers in Psychology / Language Sciences* 4, Article 226.

Malchukov, Andrej L. 2008. Animacy and asymmetries in differential case marking. Lingua 118(2). 203–221. DOI:10.1016/j.lingua.2007.02.005

Malchukov, Andrey, Martin Haspelmath & Bernard Comrie. 2011. Ditransitive constructions: a typological overview. In: Malchukov, Andrey, Martin Haspelmath & Bernard Comrie (eds.), *Studies in ditransitive constructions: A comparative handbook.* Berlin: De Gruyter Mouton. 1-64.

Matisoff, James. 1973. *The grammar of Lahu.* Publications in Linguistics, vol. 75. Berkeley: University of California.

McGregor, William B. 2010. Optional ergative case marking systems in a typological-semiotic perspective. *Lingua* 120. 1610–1636.

McGregor, William. 2018. Emergence of optional accusative case marking in Khoe languages. In: Ilja A. Seržant & Alena Witzlack-Makarevich (eds.), *The diachronic typology of differential argument marking.* Berlin: Language Science Press. 217-251.

Menschel, Jakob. Forthcoming. Differential Object Marking in Tucanoan languages. A diachronic-typological account. A manuscript.

Næss, Åshild. 2006. Case semantics and the agent-patient opposition. In Kulikov, Leonid; Malchukov, Andrej & De Swart, Peter (Ed.), Case, Valency and Transitivity. John Benjamins Publishing Company. 309–327.

Norcliffe, E., A. C. Harris & T. F. Jaeger. 2015. Crosslinguistic psycholinguistics and its critical role in theory development: early beginnings and recent advances. *Language, Cognition and Neuroscience* 30, 1009-1032.

Ospina Bozzi, Ana María. 2002. *Les structures élémentaires du yuhup makú, langue de l'amazonie colombienne: morphologie et syntaxe*. Phd thesis. Université Paris 7.

Paul, Ludwig 2008: Some remarks on the Persian suffix *-rā* as a general and historical liningguistic issue. In: Simin Karimi, Vida Samiian & Donald Stilo (eds.), *Aspects of Iranian Linguistics*. Newcastle. 329-337.

Piantadosi, Steven T., Harry Tily & Edward Gibson. 2012. The Communicative Function of Ambiguity in Language. *Cognition* 122. 280–291.

Pineda, Anna & Carles Royo. 2017. Differential Indirect Object Marking in Romance (and How to Get Rid of It). *Revue Roumaine de Linguistique* 62(4). 445-462.

Pöppel, Ernst. 2009. Pre-semantically defined temporal windows for cognitive processing. *Philosophical Transactions of the Royal Society B* 364, 1887–1896.

Rose, Françoise. 2011. Who Is the Third Person? Fluid Transitivity in Mojeño Trinitario. *International Journal of American Linguistics* 77(4). 469-494.

Seifart, F., J. Strunk, S. Danielsen, I. Hartmann, B. Pakendorf, S. Wichmann, A. Witzlack, N. de Jong & B. Bickel. 2018. Nouns slow down speech across structurally and culturally diverse languages. *PNAS* 115. 5720–5725.

Seržant, Ilja A. 2019: Weak universal forces: The discriminatory function of case in differential object marking systems. In: Schmidtke-Bode, K., N. Levshina, S. M. Michaelis & I. Seržant (eds.), Explanation in typology. Berlin: LSP.149-78.

Schikowski, Robert & Giorgio Iemmolo. 2015. Commonalities and differences between differential object marking and indexing (Zurich: University of Zurich). Zenodo. https://doi.org/10.5281/zenodo.7738049

Seržant, Ilja A. 2019. Weak universal forces: The discriminatory function of case in differential object marking systems. In: Schmidtke-Bode, Karsten, Natalia Levshina, Susanne M. Michaelis & Ilja A. Seržant (eds.), *Explanation in typology: Diachronic sources, functional motivations and the nature of the evidence*. [Conceptual Foundations of Language]. Berlin: Language Science Press. 149-178.

Seržant, Ilja A. & George Moroz. 2022. Universal attractors in language evolution provide evidence for the kinds of efficiency pressures involved. *Humanities & Social Sciences Communications* 9. Article 58.

Seyfarth, Scott. 2014. Word informativity influences acoustic duration: Effects of contextual predictability on lexical representation. Cognition 133(1). 140–155.

Siewierska, Anna & Dik Bakker. 2009. Case and Alternative Strategies: Word Order and Agreement Marking. In: Malchukov, Andrej L. & Andrew Spencer (eds.), *The Oxford Handbook of Case*. Oxford: OUP. 290-303.

Silverstein, Michael. 1976. Hierarchy of features and ergativity. In R. M. W. Dixon (ed.), Grammatical categories in Australian languages, 112–171. Canberra: Australian Institute of Aboriginal Studies.

Sinnemäki, Kaius. 2014. A typological perspective on differential object marking. *Linguistics* 52(2). 281–313.
Stassen, Leon. 2013. Comparative Constructions. In: Dryer, Matthew S. & Haspelmath, Martin (eds.), *WALS Online* (v2020.3) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.7385533 (Available online at http://wals.info/chapter/121, Accessed on 2024-04-30.)
Stenzel, Kristine. 2008. Kotiria 'differential object marking' in cross-linguistic perspective. *Amerdindia* 32. 153-181.
Strauß, Silvia. 2021. Differential argument marking with a special focus on differential object marking in Eastern Armenian. Abschlussarbeit zur Erlangung des akademischen Grades Master of Arts (M.A.) im Fachbereich Sprach- und Kulturwissenschaften der Johann-Wolfgang-Goethe-Universität, Frankfurt am Main, am Institut für Empirische Sprachwissenschaft.
Sóskuthy, Marton & Jenifer Hay. 2017. Changing word usage predicts changing word durations in New Zealand English. Cognition 166. 298–313.
Tomson, A. I. 1908. Roditel´nyj-vinitel´nyj padež pri nazvanijach živych suščestv v slavjanskich jazykach. *Izvestija Otdelenija russkago jazyka i slovesnosti Imperatorskoj akademii nauk* 13(1–2), 232–264.
Tomson, A. I. 1909: K voprosu o vozniknovenii rod.-vin. p. v slavjanskix jazykax: Priglagol'nyj rod. p. v praslav. jazyke, *Izvestija ORJaS* t. 13, kn. 3, 281-302.
Torrego Salcedo, Esther 1999. El complemento directo preposicional. In: I. Bosque & V. Demonte (eds.). *Gramática descriptiva de la lengua española. Vol. 2: Las construcciones sintácticas fundamentales. Relaciones temporales, aspectuales y modales*. Madrid: Espasa-Calpe, 1780-1805.
Walker, Katherine, Pegah Faghiri & Eva van Lier. 2024. Argument indexing in Kamang. *Studies in Language* 48(2), 287–350.
Wardlow Lane, L. & V. S. Ferreira. 2008. Speaker-external versus speaker-internal forces on utterance form: Do cognitive demands override threats to referential success? Journal of Experimental Psychology: Learning, Memory, and Cognition 34(6), 1466–1481.
Wasow, Thomas. 2002. *Postverbal Behavior*. Stanford: CSLI Publications.
Wasow, Thomas, Amy Perfors & David Beaver. 2005. The Puzzle of Ambiguity. In: C. Orhan Orgun & Peter Sells (eds), *Morphology and the Web of Grammar: Essays in the Memory of Steven G. Lapointe*. Stanford, CA: CSLI Publications, 265–282.
Wheeler, Max W, Alan Yates & Nicolau Dols 1999: *Catalan: a comprehensive grammar*. London: Routledge.
Williams, Corinne J. 1980. *A grammar of Yuwaalaraay.* (Pacific Linguistics: Series B, 74.) Canberra: Research School of Pacific and Asian Studies, Australian National University.
Wittmann, Marc. 2011. Moments in time. Frontiers in Integrative Neuroscience 5, 66.
Witzlack-Makarevich, Alena & Ilja A. Seržant. 2018. Differential argument marking: Patterns of variation. In: Seržant, I. A. & A. Witzlack-Makarevich (eds.). Diachrony of differential argument marking. Berlin: LSP.
Xu, Weijie, Jason Chon, Tianran Liu & Richard Futrell. 2023. The linearity of the effect of surprisal on reading times across languages. *Findings of the Association for Computational Linguistics: EMNLP 2023*. 15711–15721.
Zúñiga, Fernando. 2006. *Deixis and Alignment: Inverse Systems in Indigenous Languages of the Americas*. Amsterdam, Philadelphia: John Benjamins.

Zúñiga, Fernando. 2007. The discourse-syntax interface in northwestern Amazonia. Differential object marking in Makú and some Tucanoan languages. In: edited by Leo Wetzels (ed.), *Language Endangerment and Endangered Languages: Linguistic and Anthropological Studies with Special Emphasis on the Languages and Cultures of the Andean-Amazonian Border Area*. Indigenous Languages of Latin America 5; Publications of the Research School of Asian, African, and Amerindian Studies (CNWS), University of Leiden. 209–227.