

# Non-structural Case Marking in Tibeto-Burman and Artificial Languages\*

Alexander R. Coupe  
Nanyang Technological University

Sander Lestrade  
Radboud University

*This paper discusses the diachronic development of non-structural case marking in Tibeto-Burman and in computer simulations of language evolution. It is shown how case marking is initially motivated by the pragmatic need to disambiguate between two core arguments, but eventually may develop flagging functions that even extend into the intransitive domain. Also, it is shown how different types of case-marking may emerge from various underlying disambiguation strategies.*

## 1. Introduction

This paper was motivated by the authors' serendipitous discovery that a computer simulation of the emergence of case marking produced results that are uncannily similar to attested grammaticalization pathways and patterns of core case marking observed in the grammars of many Tibeto-Burman (henceforth TB) languages with pragmatically-motivated case marking. When it manifests, core case marking in both the natural and artificial languages appears to be initially motivated by the same fundamental need to disambiguate semantic roles. In other words, whenever some meaning does not follow inherently from the semantics of the dependency relationship holding between a head and its arguments, it must be encoded explicitly in the grammars of both the simulated language and the natural languages.

Disambiguating case marking is most likely to appear in clauses with two core arguments when either one may potentially fulfill the role of the agent. Under these circumstances, some grammatical means of distinguishing an A argument from an O argument may be required to clarify meaning.<sup>1</sup> Conversely, if semantic roles can be unambiguously assigned to core arguments because of the semantic nature of their referents, then there may be no overt marking, even in bivalent clauses, so the use of core case marking is observed to have a pragmatic basis. In some TB languages one also finds metaphorical extensions of this basic disambiguating function. Agentive marking is additionally used to contrast one referent from another, to encode increased agency or volitionality, to signal a shift in perspective, or to encode the atypical or unexpected behavior of a referent. In short, there are a number of motivations for the presence of core case marking in these languages, the vast majority of which appear to be linked to pragmatic factors, not syntax.

The paper combines two complementary approaches, one based on an empirical typological and diachronic study of case marking patterns in a selection of TB languages, and the other based on a computer simulation. The aim of this dovetailed approach is to demonstrate that the grammaticalization pathways and functions of agentive marking observed in natural languages are replicated by the grammaticalization of case marking in the artificial language of the computer simulation, and to show that the findings have theoretical implications for our understanding of how core case-marking patterns evolve.

---

\* This division of labor for writing this paper is as follows. Coupe was responsible for writing the sections on the natural languages, Lestrade was responsible for the sections on the computer simulation, and the introduction and concluding sections were jointly written.

<sup>1</sup> A represents the agent-like argument of a transitive clause, O represents the patient-like argument of a transitive clause, and S represents the single argument of a transitive clause (after Dixon 1979, 1994).

The first part of the paper deals with the natural languages. Section 2 first discusses the functions of case marking and presents a succinct overview of case marking patterns from a typological perspective. Section 3 first provides a thumbnail sketch of the typological characteristics of TB languages, the areal influences that have resulted in the extensive typological diversity found in these languages, and why variable agentive marking appears to have been frequently overlooked in earlier studies. It also introduces the TB languages of this study. Next, it presents six pragmatic functions of agentive case marking on S and A arguments in TB languages exhibiting variable core marking patterns, and then illustrates each of these with examples. Section 4 discusses the diachronic development of agentive case markers in TB, their likely lexical sources, and the means of their metaphorical extension from an oblique function to a core A marking function. This section also contrasts differential marking of O arguments via an extended function of the dative case in Apatani and considers the pragmatic contexts in which this occurs.

The second part of the paper deals with the methodology and findings of the computer simulation. Section 5 briefly discusses the relevance of combining typological and computational methods, and Section 6 introduces *MoLE*, an R package designed by the second author to simulate language evolution. In Section 7, the actual experiment is discussed. It will be shown how different types of case-marking may emerge from various disambiguation strategies. The results of only one of these strategies will be shown to mimic the development of case marking in TB, including the extension into the S domain. We will end with a general discussion in Section 8.

## 2. Functions of core case marking

Core case marking serves a fundamental relational function of making apparent the relationship holding between the predicate and its dependent arguments; case marking thus makes explicit ‘who does what to whom’. In transitive clauses, case marking typically appears on just one argument of the predicate (Comrie 1989, Dixon 1994, Siewierska & Bakker 2009), as that is sufficient to discriminate between A and O core grammatical functions.<sup>2</sup>

Two cross-linguistically common case marking alignments are the nominative-accusative and ergative-absolutive patterns. In the nominative-accusative pattern, the S argument of an intransitive clause and the A argument of a transitive clause are case-marked identically by nominative case, prototypically by the absence of morphological marking. This is contrasted with overt accusative case marking on the O argument of a transitive clause. In the ergative-absolutive pattern, it is the S argument of an intransitive clause and the O argument of a transitive clause that are case marked identically by absolutive case (which is prototypically represented by the absence of morphological marking). This is contrasted by overt ergative case marking on the A argument of a transitive clause.

In his discussion of ‘main alignment types’, Comrie (2013) suggests that a language such as Mandarin is representative of a third type of alignment that he refers to as ‘neutral’. An alternative approach, and one supported by the data of the natural and artificial languages to be presented below, is that Mandarin is a language that actually has no alignment, simply because there has been no grammaticalization of the morphological means of encoding ‘who does what to whom’. There are also rarer patterns of case marking, such as a transitional tripartite pattern that uniquely case marks A, S and O arguments, split-S patterns, fluid-S patterns, and active-stative patterns in addition to splits between nominative-accusative and ergative-absolutive alignment

---

<sup>2</sup> Comrie (1978, 1989) uses P to denote the patient-like argument of a transitive clause, in contrast to Dixon’s (1979, 1994) use of O.

patterns based on tense, aspect, mood, or animacy (e.g. see Comrie 1978 and Dixon 1994), but these will not be described or discussed further, as they are orthogonal to the focus of our paper and not attested in the data to be discussed.

Yet another pattern addressed only briefly by Comrie (1978: 384; 1989: 130) and Dixon (1994: 32–33) concerns one in which overt case marking on an A argument appears only when there is a pragmatic need to disambiguate the semantic roles of core arguments.<sup>3</sup> If the nature of the referents of the core arguments and the type of the event in which they are involved renders explicit case marking redundant, then there may be no marking on the A argument, so the motivation for the marking or its absence is pragmatic, rather than syntactic. Furthermore, overt marking may additionally occur on S arguments under certain circumstances. This turns out to be an increasingly reported case-marking strategy in TB languages, as we demonstrate in Section 3. We will refer to this additional pattern as pragmatically-motivated agentive marking (henceforth PMA marking). As the presence or absence of PMA marking applies to both A and S arguments independently of syntactic valency considerations, it is formally distinct from the syntactically-motivated ergative-absolutive pattern described above.

The recent focus on the use of naturalistic textual data in detailed studies of TB languages has revealed the extent to which their case marking patterns diverge from canonical ergative-absolutive or nominative-accusative morphological alignments (e.g. see LaPolla 1992, 1995 for an overview of case marking in Tibeto-Burman, and LaPolla 2003 for Qiang; Chelliah 1997, 2009 for Meithei; Coupe 2007, 2011a, 2011b for Mongsen Ao; Hyslop 2010 for Kurtöp; Lidz 2011 for Yongning Na, Morey 2012 for Singpho; Peterson 2011 for Khumi, Teo 2012 for Sumi, Tournadre 1991 for Tibetan, and Willis 2011 for Darma, among others). Even a language like Tibetan, which was once widely characterized as a prototypical ergative language (e.g. DeLancey 1981; Dixon 1994; and Miller 1996: 757, cited in DeLancey 2011) has been shown to deviate significantly from a systematic case-marking alignment based on the transitivity status of the clause (cf. Tournadre 1991). While there are also convincing accounts of TB languages that do demonstrate a consistent, canonical pattern of ergative-absolutive case marking, such as Dolakha Newar (e.g. Genetti 2007), the increasing number of fine-grained grammatical descriptions emerging in recent years suggests that the pattern most common to TB is one in which non-syntactic factors motivate the presence or absence of core (and particularly agentive) case marking (LaPolla 1992, 1995, 2004; Coupe 2011a; DeLancey 2011).

Misconstrued interpretations of case marking patterns and the recognition of split ergative systems in earlier work can probably be attributed to the methodology employed: it is now known that analyses based on directly elicited data often fail to create the specific pragmatic contexts that motivate the use of core case marking in these languages. Elicited data may consequently produce regular paradigms that are not actually attested in narrated texts (Macgregor 2009: 493; Willis 2011: 103), or the structure of the contact language used may exert an adverse influence on the structure of the elicited data (Chelliah 1997: 129; Lidz 2011: 50; Willis 2011: 110).

Some authors refer to variable case marking on S arguments of monovalent clauses and A arguments of bivalent clauses as “optional ergativity” in the typological literature. Although he promotes this term, McGregor (2009: 493–494, 2010: 1615) concedes that it is misleading, as the presence or absence of the marking is not random and is convincingly shown to encode different nuances in meaning. He estimates that 10% of morphologically ergative languages show variable marking patterns. The divergent agentive marking pattern is almost certainly higher in TB languages, based upon the descriptions of case marking in the cited papers of this article and other emerging work. In the interests of clarity, the term ‘ergative’ is only used here to refer to

---

<sup>3</sup> Comrie (1978) identifies Hua as an example of a language that exhibits this pattern.

the paradigmatic, syntactically-motivated marking on an A argument of a bivalent clause, such as is found in Dolakha Newar (Genetti 2007), which consistently distinguishes it from the O argument of a transitive clause and the S argument of an intransitive clause. In contrast, ‘agentive’ is used for describing a type of core case marking that may be used on both an A argument of a bivalent clause and the S argument of a monovalent clause, and whose distribution is not dictated by clausal transitivity considerations. This usage is also reflected in the adjusted glosses of cited data.

### 3. PMA marking in Tibeto-Burman

The TB branch of the Sino-Tibetan family possibly includes up to three hundred languages scattered over a vast region of Asia. This extends westward from southern China and across the Tibetan Plateau to Baltistan (Pakistan), southward into peninsular Southeast Asia, Myanmar, and the northeastern periphery of South Asia, and into the valleys of the Himalayas. Because of the considerable time depth of the Sino-Tibetan family (comparable to that of Indo-European, according to Matisoff [1991: 470]), and the likelihood of repeated language contact scenarios as they have spread westward and southward in at least two separate migrations, TB languages show a great deal of typological diversity.

Those languages located in mainland Southeast Asia are most likely to share features in common with Sinitic languages, such as monosyllabicity and isolating word formation characteristics, potentially complex lexical tone systems, serial verb constructions, aspectual and mood contrasts rather than tense oppositions, a general absence of inflectional morphology, and rudimentary relational marking. Conversely, languages located closer to South Asia are more likely to demonstrate typological features common to the languages of this linguistic area, due to the effects of language contact. Thus, they tend to have moderately synthetic and agglutinative word formation, well developed case-marking patterns (including nominative-accusative, ergative-absolutive and split systems, as well as non-paradigmatic, pragmatically-based patterns), they are more likely to have inflectional morphology, and they may additionally exhibit a variety of head marking configurations (such as indexical nominative-accusative, inverse and hierarchical marking patterns). Like most languages of South Asia, they typically make great use of non-finite clause chaining, and they are more likely to be either atonal or have simple tone systems, depending on their relative proximity to Southeast Asia.<sup>4</sup> Matisoff (1991) attributes these divergent typological characteristics to two great domains of cultural and linguistic influence that he terms the ‘Sinosphere’ and the ‘Indosphere’. The languages to be discussed in this paper mostly share the Indospheric typological profile on account of their contiguity with the Indo-Aryan languages of South Asia.

The data of ten TB languages are presented in this paper. Their names, branch affiliations and the locations where they are spoken are as follows: Darma, Tibetan and Kurtöp are Bodish languages spoken in Uttarakand State (India), on the Tibetan Plateau, and in northern Bhutan respectively; in the mountains of the Indo-Burmese Arc can be found Mongsen Ao and Sumi in central and southern Nagaland, both of which are unclassified for branch affiliation, but may form an Indo-Burmic subgroup; a little further to the south in the same mountain range is the Meithei (Meeteilon) language of Manipur, also unclassified, but sharing a number of features with the Indo-Burmic languages of Nagaland; yet further south near the end of the Indo-Burmese Arc can be found Khumi, a Kuki-Chin language spoken in the Chittagong Hill Tracts of Bangladesh; Qiang is a Qiangic language spoken in Sichuan Province on the eastern edge of the

---

<sup>4</sup> Those TB languages spoken in the Indo-Burmese Arc are likely to have two to five lexical tones, whereas those spoken further to the west and in the Himalayas are more likely to be atonal.

Tibetan Plateau; Yongning Na is an unclassified language spoken in Sichuan and northern Yunnan with similarities to languages of the Lolo-Burmese branch; and lastly, Apatani is a Tani language spoken in Arunachal Pradesh, Northeast India.

Although genetic classification in Sino-Tibetan very much remains a nascent work in progress, and many languages are still unclassified with respect to their branch affiliation, it is likely that at least six branches of Sino-Tibetan are represented by the abovementioned languages. The occurrence of PMA marking in all but the Tani group suggests that it represents the archetypal marking pattern in TB, yet the ancestral languages of these geographically diverse branches are likely to have split off from the proto-language at a very early stage. Furthermore, despite the ubiquity of the PMA marking pattern, no common forms of core relational marking can be convincingly reconstructed beyond the individual branches to the level of Proto-Tibeto-Burman (LaPolla 1995). Spread through language contact also seems unlikely, given the different directions of westerly and southerly migration, so the parallel development of PMA marking in independent branches of TB is most plausibly attributed to drift (Sapir 1921, LaPolla 1994).

To demonstrate that the PMA pattern of case marking described in this paper is considerably more common in these languages than previously appreciated, the following sections will present examples of all the contexts in which it is reported to occur. We hold that the primary motivation for a core case marking function initially developing is for the disambiguation of semantic roles, in common with Comrie (1989), Dixon (1994), and Siewierska & Bakker (2009). However once this extended function of an erstwhile oblique case marker has grammaticalized, the case marker can be potentially used for other metaphorically extended functions that appear to be unconstrained by factors relating to syntactic valency and argument structure.

The following lists the reported contexts in which PMA marking pattern manifests in the TB languages of this study. PMA marking is used:

1. to disambiguate semantic roles in bivalent clauses, including when there are atypical constituent orders in which the patient NP precedes the agent NP;
2. for contrastive reference, i.e. to pragmatically distinguish a referent from one or more other potential referents, or to foreground a referent;
3. to encode extraordinary, willful, or atypical behavior, especially in the context of an unexpected turn of events;
4. to clarify the semantic role of a remaining core argument when another core argument undergoes ellipsis;
5. in some languages, to distinguish the causer argument in causativized clauses;
6. in some languages, to characterize a habitually acting NP referent.

Each of these uses of PMA marking are discussed in turn and illustrated with examples in subsections 3.1-3.6 below. Most of the presented data in this section necessarily comes from narrative texts, as the use of PMA marking is often dependent upon pragmatically-marked contexts; elicited examples are noted as such when used for drawing comparisons.

### **3.1 Disambiguation of semantic roles**

Many TB languages are reported to use disambiguating PMA marking when there is the possibility of confusion over semantic role assignment. The marking is particularly motivated when core arguments have human referents, as demonstrated by the Kurtöp example in (1) below. Conversely, when semantic role assignment is not in doubt because of the semantic

nature of the NP referents, then there may be no requirement for core case marking whatsoever, as in (2).

- (1) Kurtöp (Hyslop 2010: 12)

*yum-gi se ni emo yap zhuk-sa-na*  
 mother.HON-AGT prince and princess father.HON stay.HON-NMLZ-LOC

*zon-pala wenta la*  
 send-PFV COP.EQ.MIR POL

‘The mother sent the prince and princess to where the father was.’

- (2) Kurtöp (Hyslop 2010: 13)

*Rinzin gari ’lup*  
 Rinzin car learn

‘Rinzin was learning to drive.’ (lit. learn car)

PMA marking is likely to appear in bivalent clauses when expected semantic roles do not accord with expectations based on the animacy hierarchy (Silverstein 1976). Under these circumstances, disambiguating PMA marking serves to correct assumptions that are based on real world knowledge, but do not apply for a particular atypical state of affairs. When semantic roles do accord with assumptions concerning hierarchical relationships in the real world, then no disambiguating marking may be needed. For example, humans typically slaughter pigs, rather than vice-versa. This accounts for why both core arguments of ‘kill’ in the Mongsen Ao example of (3) are morphologically unmarked. However, if it were the case that the pig killed the 3<sup>rd</sup> person referent of that example, then this would require agentive marking on the A argument to clarify the unexpected semantic role of the pig in this extraordinary situation. This pattern is representative of many TB languages in which the presence or absence of core case marking is not determined by constraints based on syntactic valency.

- (3) Mongsen Ao (Coupe 2007: 158)

*təɪ, pa ʰpʰ ənsət.*

tə-əɪ pa ʰpʰ ənsət-Ø  
 thus-SEQ 3SG PROX kill-PST

‘And then, she killed this.’ [i.e. in context, the pig]

In Qiang, non-canonical constituent orders in which the actor argument is not in the topical clause-initial position obligatorily require agentive marking to distinguish an A argument from an O argument. This becomes necessary when both NPs have 3<sup>rd</sup> person referents, because in this situation person marking on the verb cannot reveal which core NP is the cross-referenced A argument. According to LaPolla (2003:78–79), inclusion of the agentive marker *-wu* in (4a) would clarify that it was *χumtʂi* and not someone else who did the hitting, whereas the non-canonical constituent order of (4b) renders agentive marking obligatory on the immediately preverbal NP to avoid ambiguity.

- (4) Qiang (LaPolla 2003: 78)

a. *χumtʂi-(wu) khumtʂi zə-dzə-u*  
*χumtʂi-(AGT) khumtʂi DIR-hit-VIS*  
 ‘χumtʂi hit Khumtʂi.’

- b. *khumtsi χumtʃi-wu zə-dzə-u*  
 khumtsi χumtʃi-AGT DIR-hit-VIS  
 ‘Khumtsi was hit by χumtʃi.’

The constituent order of core arguments in TB languages is generally determined by pragmatic considerations of topicality and focus. The default position for the verb is uniformly clause-final in a single intonation unit, but the order of NP arguments before the verb does not serve to distinguish grammatical functions, and OAV orders are likely to trigger PMA marking in these languages, in common with Qiang (e.g. see Coupe [2007: 151–153], Lidz [2011: 56–57], and Willis [2001: 104–105] for examples and discussion.

### 3.2 Contrastive reference

The contrastive function of PMA marking demonstrated above in (4a) is common in other TB languages and is widely reported across the TB-speaking area (e.g. Tibetan [Tournadre 1991], Mongen Ao [Coupe 2007], Kurtöp [Hyslop 2010], Yongning Na [Lidz 2011], Darma [Willis 2011] and Sumi [Teo 2012]). The following examples demonstrate how the referent of an agentive case-marked NP is pragmatically distinguished from one or more other referents. Again, this is unrelated to syntactic valency, as it applies to the core arguments of both monovalent and bivalent clauses, and even to arguments of verbless clauses (e.g. see [7]).

- (5) Tibetan (Tournadre 1991: 101–102)

- a. *khyed zhugs a nga:s phyin.dgos*  
 2SG stay(HON) PART 1SG-AGT go(past)-MODAL AUX  
 ‘(Please) you stay, I will go (for you).’
- b. *khong-gis phyin-pa.red*  
 he-AGT go(past)-AOR+GNOMIC  
 ‘He is the one who went.’

- (6) Mongsen Ao (Coupe 2007: 482)<sup>5</sup>

*təku, pùkphula sə nə “nì nə ajim.rù àw,” tətə sa.*  
 tə-ku pùkphu-la sə nə “nì nə ajim.rù àw” tətə sa-Ø  
 thus-LOC owl-F ANAPH AGT 1SG AGT cry.out-IMM PTCL thus say-PST  
 ‘Upon [the leader asking that], aforementioned Owl said “I will make a proclamation.”’

- (7) Sumi (Teo 2012: 61)

*Paza no Sümi.*  
 pa-ʒá = nò ʃimì  
 3SG-mother=AGT Sumi  
 ‘His *mother* is Sumi.’ (as opposed to his father)

<sup>5</sup> This Mongsen Ao example also demonstrates how a verb of vocalization commonly assigns agentive case marking to its actor argument. This is a widely observed phenomenon in TB languages reported to have PMA marking (e.g. Coupe 2007, Lidz 2011, Morey 2012, Peterson 2011, Teo 2012).

## (8) Kurtöp (Hyslop 2010: 13)

a. *tshe ge-shang khit*  
 DM go-PFV.EGO 3.ABS  
 ‘So he left.’

b. *khî ge-shang*  
 3.AGT go-PFV.EGO  
 ‘She went.’ (contrary to the interlocutor’s assumption)

While Peterson (2011) justifiably does not explicitly identify an agentive marker in Khumi, it is nevertheless significant that a case marking clitic =*mö*<sup>3</sup> (glossed FGR for ‘foregrounder’) mimics the contrastive function encoded by agentive case markers in other TB languages, in addition to frequently appearing in bivalent clauses with non-canonical constituent orders:

## (9) Khumi (Peterson 2011: 81)

*j’vo*<sup>4</sup>=*mö*<sup>3</sup> *kay*<sup>1</sup>=*mö*<sup>3</sup> *la*<sup>1</sup>...  
 husband=FGR 1SG=FGR take.IRR  
 ‘(Her) husband said, “I’ll [as opposed to someone else] get (=take) it...”’

In the following example from Darma, Willis (2011: 118) reports how a woman who was listening to the narration interjected with the clarifying qualification that it is *girls* who ritually toss the cake, not the men. The presence of PMA marking on the NP *tsemme* in (10b) specifically appears to encode this contrasted point of reference.

## (10) Darma (Willis 2011: 118)

a. *hã mala çya-lan ju, jo nini, ?a dulaŋ t<sup>h</sup>o-dan*  
 then garland put.on-CVB after HM HM cake toss.ritually-1PL.NPT  
 ‘Then after putting on the garland, um, (we) ritually toss the cake’ (Haircutting Ceremony)

b. *tsemme su!*  
 girl AGT  
 ‘The girls (do it)!’ (Hair-cutting Ceremony)

Related to this contrasting function is the use of PMA marking to encode a shift in perspective. Lidz (2011: 56) observes that the presence of agentive case marking in Yongning Na correlates with a change in viewpoint from one participant to another and suggests that the marking is contributing to a foregrounding function in discourse. She compares this foregrounding function to a similar use of agentive marking reported in Khumi (Peterson 2011), which also uses a case marker for foregrounding participants in narrative discourse, particularly when they are behaving contrary to expectations or notions of volitionality.

### 3.3 Encoding extraordinary, willful or unexpected behavior

Unusual turns of events are often signaled in TB languages by the use of PMA marking. Willis (2011: 114) describes how a core argument of ‘sit, live’ occurs with agentive marking in a narrative text because the NP referent, a local king, is hiding underground instead of living in his

palace. This represents an unexpected and marked situation, and its peculiarity is brought into focus by the speaker's use of the agentive marker *su*.

(11) Darma (Willis 2011: 114)

...*cet ru, ʔu su, ʃo nini, ki raʃa, ʃo nini, ʔu da, raʃa*  
 place.name LOC 3SG AGT HM that.LN king.LN HM 3SG CONT king.LN

*da andergrawnd ru ʃyʊŋ-ʃyɛ-n le-ʃu.*  
 CONT underground.LN LOC live-MID-NMLZ AUX.EX-PST

'...in Cet, he, that is, that the king, um he, the king was living underground.'

One of the functions of PMA marking in Mongsen Ao is to encode increased agency. As in Darma, the verb *li* 'stay, live' does not normally require an agentive-marked core argument, but in the following textual example a woman threatens to desert her husband if he does not acquiesce to her demand for him to kill her stepson. This represents a pragmatically-marked situation, and the narrator strongly encodes the stepmother's intent by assigning PMA marking to the 1<sup>st</sup> person pronoun in the apodosis.

(12) Mongsen Ao (Coupe 2017: 291)

*nàŋ nā ānū ī mətəpsət-tʃūk-pālā, nì nā nāthēn mətliʃùʔ, tət sā, ʃāpītsāi tʃū nā.*  
 nàŋ nā ānū ī mət-təpsət-tʃūk-pālā nì nā nā-thēn  
 2SG AGT NRL-child PROX NEG-kill-PFV-COND 1SG AGT 2SG.POSS/OBL-COM

*mət-lī-ì-ùʔ tət sā-Ø ʃāpītsāi tʃū nā*  
 NEG-stay-IRR-DECL thus say-PST stepmother AGT DIST

'“If you do not kill this child, I will not stay with you”, said the stepmother.'

Similarly, the Mongsen Ao verb *tʃāʔ* 'consume' does not normally require an actor argument to be assigned agentive case marking when reporting on a typical, everyday event of eating. However, the following pair of carefully elicited sentences demonstrates how the agentive marker *nā* can be used to encode a willful actor that is behaving with increased volitionality and intent. The context in which (13a) applies is a default situation in which chickens are fed a handful of paddy and the speaker reports on the event, whereas (13b) reports on a pragmatically-marked situation. Chickens freely wander around Ao villages and sometimes they opportunistically pilfer unguarded paddy, which is left exposed on mats to dry in the sun. One day a Mongsen Ao speaker was informed that some chickens were raiding his drying paddy and was asked what was happening, and (13b) was uttered in response as he chased the raiding chickens away. There is a clear difference in meaning between the two sentences, and this is created by the addition of the PMA marking in (13b). The chickens learn from experience that villagers will try to beat them for stealing the drying paddy, so it is plausible that their augmented agency is encoded in such pragmatic contexts by the presence of the agentive marker.

(13) Mongsen Ao (Coupe 2007: 157)

a. *ahən atʃak tʃāʔ-əi-ùʔ.*  
 a-hən a-tʃak tʃāʔ-əi-ùʔ  
 NRL-chicken NRL-paddy consume-PRES-DECL  
 'The chickens are eating paddy.' (elicited data)

- b. *ahən nə atʃak tʃàù?*  
 a-hən            nə        a-tʃak            tʃàù-àɪ-ù?  
 NRL-chicken    AGT    NRL-paddy    consume-PRES-DECL  
 ‘The chickens are eating paddy.’ (implying that they are stealing it)

According to Chelliah (2009: 387), extraordinary behavior is analogously signaled by agentive marking in Meithei, as captured by the contrast in the following pair of sentences.

- (14) Meithei (Chelliah 2009: 387)
- a. *tomba chá čá-í.*  
 Tomba    meat    eat-NHYP  
 ‘Tomba ate meat.’ (Tomba is a non-vegetarian.)
- b. *tomba-nə chá čá-í.*  
 Tomba-AGT    meat    eat-NHYP  
 ‘Tomba ate meat.’ (A noteworthy activity, not expected for this vegetarian.)

### 3.4 Clarification of semantic roles under ellipsis

Many TB languages freely permit the ellipsis of core arguments if they are recoverable from the context. When an argument of a bivalent verb is elided, the presence or absence of case marking on the remaining argument clarifies its semantic role, especially if this cannot be correctly deduced from the context. This function is demonstrated in the following pair of Kurtöp sentences.

- (15) Kurtöp (Hyslop 2010: 12)
- a. *nyarop zon phang-zi*  
 fisherman two feel.pity.for-NF  
 ‘The two fishermen were pitied...’ (elicited data)
- b. *nyarop zon-gi phang-zi*  
 fisherman two-AGT feel.pity.for-NF  
 ‘The two fishermen felt pity (for the prince and princess) ...’

Lidz (2011: 58) notes that agentive marking possibly plays a disambiguating role in some Yongning Na sentences, in addition to highlighting agency. Without the agentive marking on *mi<sup>31</sup> zɣ<sup>13</sup>* in (16), for example, the meaning could be understood as ‘Again (he) went to teach the woman’.

- (16) Yongning Na (Lidz 2011: 58)
- wɣ<sup>13</sup> mi<sup>31</sup> zɣ<sup>13</sup> nu<sup>33</sup> t<sup>h</sup>w<sup>33</sup>-sɔ<sup>13</sup> hu<sup>33</sup>*  
 again woman        AGT    DUR-teach    go  
 ‘Again, the woman went to teach (him).’

### 3.5 PMA marking in causativized clauses

Just as the occurrence of a single agentive-marked NP in a clause with a bivalent verb serves to indicate an actor semantic role, so can the absence of case marking on an NP indicate an undergoer semantic role. This function of core case marking is particularly important for clarifying meaning in Mongsen Ao causativized clauses with elided arguments. Causativization obligatorily triggers agentive marking on a causer argument and either zero or dative marking on a causee argument;<sup>6</sup> this represents one of only two situations in which the disambiguating function of agentive marking is consistent and mandatory. The referent of the zero-marked 3<sup>rd</sup> person pronoun *pa* is accordingly identified as a causee in (17a), whereas the presence of agentive marking on *wàzàʔ* in the consecutive sentence of (17b) establishes that NP as the causer argument.

(17) Mongsen Ao (Coupe 2011b: 514)

a. *təɪ pa tsəpàʔ ku atsə mətʃəməjùʔ.*

tə-əɪ pa tsəpàʔ ku a-tsə mə-tʃəm-iʔ-i-ùʔ  
 thus-SEQ 3SG pond LOC NRL-water NEG-drink-CAUS-IRR-DECL  
 ‘And, [they] would not let her drink at the pond.’

b. *təɪ wàzàʔ nə tsəpàʔ ku atsə mətʃəmilà.* [mə-tʃəm-iʔ-là]

tə-əɪ wàzàʔ nə tsəpàʔ kũ ā-tsə mə-tʃəm-iʔ-là  
 thus-SEQ bird AGT pond LOC NRL-water NEG-drink-CAUS-NEG.PST  
 ‘And, the birds didn’t let [her] drink at the pond.’

Obligatory disambiguating agentive marking in causativized clauses could plausibly provide a pathway to the development of consistent marking of all A arguments if it is generalized to non-causativized clauses, possibly resulting in the grammaticalization of a paradigmatic ergative-absolutive alignment pattern (cf. the related discussion of *generalization* in Section 6).

### 3.6 Characterizing habitual activity

Some TB languages use agentive marking in generic statements not to express meanings associated with any of the situations described above in Sections 3.1-3.4, but to encode instinctive or habitual behaviors. Such usages seem far removed from the functions of PMA marking seen in preceding examples, or from the expression of increased agency, personal choice or unexpected behavior often associated with this type of case marking. Chelliah (2009: 387) accounts for the use of agentive marking for this purpose in Meithei by proposing that agents case-marked by the agentive marker *-nə* can be understood as performing an activity that characterizes the class of referent or the force of tradition, e.g.

(18) Meithei (Chelliah 2009: 391)

a. *hindu-síŋ-nə lukun-si tháj-í*  
 Hindu-PL-AGT sacred.thread-PROX wear-NHYP  
 ‘Hindus wear the sacred thread.’

<sup>6</sup> Indirectly caused events involving a willing causee are differentially encoded with dative case marking, whereas a zero-marked causee may or may not be interpreted as a coerced participant. See Coupe (2007: 190ff.) for details.

- b. *učék-siŋ-nə sul atiya-tə pay-í.*  
 bird-PL-AGT air sky-LOC fly-NHYP  
 ‘Birds fly in the open air.’

Coupe (2011b: 499) compares a very similar meaning encoded by agentive marking in generic statements in Mongsen Ao. Possibly the agentive marker encodes a type of referentiality in this instance of use. If so, it somewhat mirrors the use of dative marking on O arguments in some languages to encode increased referentiality, as in Spanish and Hindi. The difference between the two sentences of (19a-b) below is that example (19a) without the agentive marking reports an activity in progress (which was uttered in response to the question “What are you doing?”), whereas the elicited example of (19b) with agentive marking on the 1<sup>st</sup> person argument expresses a characterizing meaning consistent with the preceding Meithei sentences (cf. also Lestrade and de Hoop 2016).

(19) Mongsen Ao (Coupe 2011b: 499)

- a. *nì asəŋ səið*  
 nì a-səŋ sə-əɪ  
 1SG NRL-wood chop-PRES  
 ‘I’m chopping wood.’
- b. *nì asəŋ səið*  
 nì nə a-səŋ sə-əɪ  
 1SG AGT NRL-wood chop-PRES  
 ‘I chop wood.’ (i.e. habitually, as an occupation)

Having illustrated the wide variety of non-syntactically determined situations in which PMA marking appears in TB languages, our attention will now turn to a consideration of the diachronic sources of core relational morphology – how spatial case marking first develops from relational nouns expressing locations, and how this is metaphorically extended to encoding even more abstract uses in grammar. This can be most conveniently demonstrated with reference to the development of agentive case markers in Mongsen Ao and Tibetan.

#### 4. The diachronic development of PMA marking

Grammaticalization studies consistently demonstrate pathways of development from more concrete concepts to more abstract concepts in grammar. This trajectory has been noted by numerous authors who have sought to explain how complexity develops in grammar (e.g. Givón 1971, 1979, 2009; Heine et al. 1991; Heine & Kuteva 2002, 2007; Dahl 2008, among others). In TB languages as well as in many other head-final languages of the world, nouns are typically found to grammaticalize as case-marking postpositions from appositional or genitival constructions that initially encode concrete spatial meanings, such as location, source and goal (DeLancey 1984: 62; Aristar 1991; LaPolla 1995: 214, 2004; and Coupe 2011b: 506).

Once a spatial noun is reinterpreted as a postposition encoding an oblique relational meaning, it may then be used in an innovative way for encoding even more abstract functions of grammar, such as for marking a core grammatical case relation. Grammaticalized relational morphology is additionally often the target of further cycles of reanalysis that develop yet more syntactic uses. For example, space > time grammaticalizations are a particularly common pathway for the development of reanalyzed functions of case marking morphology in many of the world’s

languages – see Genetti 1986, 1991 and Coupe 2007, 2017 for accounts of how case markers become reinterpreted as non-finite converbal suffixes encoding a range of temporal and/or adverbial presuppositions in TB and other head-final languages of South Asia.

Given the ubiquity of a syncretic agentive/instrumental form *na~nə* in all but one language of the Ao group of central Nagaland (viz, Mongsen Ao, Lotha, Sangtam and Yimchungrü), it can be reconstructed to the level of Proto-Ao as *\*na* (Coupe 2011a: 32). The Ao dialects go one step further and are universally unique in sharing a case syncretism involving not only the above-mentioned agentive and instrumental cases, but also an allative case. The following Mongsen Ao proverb demonstrates these three case relations in the one clause:

(20) Mongsen Ao (Coupe 2007: 172)

*aji nə tuɪ nə athùtʃən nə anətpən wa mətəm.*

*ā-jī      nə      tūɪ      nə      āthùtʃ-ʃən      nə      ā-nət-pən      wā      mətəm*

NRL-dog    AGT    GPN    INST    vomit-LNMLZ    ALL    NRL-two-ORD    go    like

‘Like a dog going back to its vomit by itself for a second time.’

(= to eat one’s words; to reject something and then want it later)

The likely lexical source of this relational morpheme is Proto-Tibeto-Burman *\*ʔ-nam* ‘side/rib’, as reconstructed by Matisoff 2003 for Proto-Lolo-Burmese. A cognate form *tə-na* occurs in the Chungli dialect of Ao, defined by Clark (1911: 626) as meaning ‘side at the waist where there are no ribs’. The noun ‘side’ is also a known lexical target for the grammaticalization of locatives in Chinese and some French-based pidgins and creoles (Heine & Kuteva 2002: 272). Similarly, Beames (2012[1875]: 257) proposed that the lexical origin of the Hindi dative marker *ko* is the locative declension of Sanskrit *kākṣa-* ‘armpit’. Reinöhl (2016: 58–59) suggests that this lexical meaning was metaphorically extended first to ‘side of body, flank’, then to the marking of goals and recipients, and finally to the marking of human and definite patients in a further metaphorical extension of function.

The Proto-Ao form *\*na* plausibly grammaticalized as a type of local case marker from the relational head of a compound noun. This is consistent with LaPolla’s (1995) observations concerning the diachronic origins of TB case-marking systems, except for the fact that an allative/agentive/instrumental isomorphism is an unattested syncretism in his survey; neither was it encountered in the study of Bodic case marking syncretisms by Noonan (2009). It is much more typical for the grammaticalized marker of a goal to be metaphorically extended to marking a core O argument, as demonstrated by the Hindi dative/accusative marker *ko*. However, at the initial stage of grammaticalization, Proto-Ao *\*na* presumably functioned simply as a semantically unspecified marker of a local case relation, as suggested by the multiple case roles encoded by a syncretic form in the synchronic grammar of the Mongsen dialect (e.g. [20]), so theoretically any oblique case role could have been metaphorically extended to a core marking function (cf. Lestrade 2016b for a discussion of unspecified local relations). Because agents and instruments are in some sense both effectors of events and activities, logically it must have been the instrumental meaning of *\*na* that was the target for the grammaticalization of the agentive marking function.

A similar development is suggested by the isomorphic form of the agentive/instrumental marker in Tibetan. LaPolla (1995: 191) reconstructs a syncretic agentive/instrumental marker *-s(V)*<sup>7</sup> in sub-groups of the Bodish branch of TB and observes that this morpheme is also partially isomorphic with the ablative marker (in common with Mongsen Ao, whose ablative form is *phinə*). The probable lexical source is an earlier Proto-Tibetan relational noun *\*-sa*

<sup>7</sup> (V) represents an optional unspecified vowel.

‘place’, which is also used to derive nominalizations such as *khuj* ‘hole, pit, cavity’, *khuj-sa* ‘hole place’ > *khuj-s* ‘mine’; *khonj* ‘the inside’, *\*khonj-sa* ‘inside place’ > *khongs* ‘middle, midst’; and *ŋo* ‘face’, *ŋo-sa* ‘face-place’ > *ŋos* ‘side, direction, surface’ (Beyer 1992: 118).

The agentive marker *-s* of modern Tibetan is very similar in form to the grammaticalized relict of ‘place’ in Classical Tibetan, so this constitutes another instance of the relational noun > oblique postposition > core case marker development that is attested in many head-final languages.

- (21) *kho-s bu.mo snying.rje.po de-la btlas-song*  
 he-AGT girl nice this-OBL look-AOR+EVID  
 ‘He looked at the nice girl.’ (Tournadre 1991: 94)

Once a lexical item has grammaticalized as a postposition, its new grammatical roles are open to a range of relational interpretations. This accounts for the frequently observed case syncretism in TB languages (e.g. Noonan 2009, Coupe 2011a), as well as in many other language families of the world – isomorphic ergative/instrumental and ergative/ablative markers are reported to be especially common (LaPolla 1995, 2004). Since only one argument may need to be grammatically marked for the disambiguation of meaning, a language has the option of overtly marking either the A argument *or* the O argument of a bivalent clause, and this choice potentially has ramifications for whether a language may eventually grammaticalize a paradigmatic nominative-accusative or ergative-absolutive alignment pattern. Both are possible eventual outcomes of the grammaticalization of core argument marking.

The Apatani examples of (22) are illustrative of the disambiguating O marking pattern. In this language, dative case is extended to marking an O argument to disambiguate semantic roles in bivalent clauses. The innovative dative marking is only obligatory when the referents of both core arguments are possible instigators of an event. For this reason, LaPolla (1992, 1994) refers to the pattern as ‘anti-ergative’ marking, as it distinguishes an NP argument as *not* being in A function. This term was later replaced by ‘anti-agentive’ to make it more widely applicable to TB languages. The current fashion is to refer to the pattern as ‘differential object marking’ (DOM) in the typological literature, principally after Bossong 1985.<sup>8</sup>

- (22) Apatani (Abraham 1985: 38-40)
- a. *ŋi-ka ani tade mi yo bi-bi-ne*  
 1POSS-GEN mother pers.name DAT meat give-EVID-NPROX  
 ‘My mother gave meat to Tade’
- b. *sihini mó mi alitu-bi-ne*  
 cow 3SG DAT kick-EVID-NPROX  
 ‘The cow kicked him’
- c. *mó sihini pa-bi-ne*  
 3SG cow kill-EVID-NPROX  
 ‘He killed the cow.’

<sup>8</sup> The term is purposely avoided here, because it makes unsubstantiated assumptions that all languages exhibiting this pattern must have a grammatical relation of ‘object’. We prefer to remain uncommitted regarding the existence of grammatical relations in the scantily described Apatani language.

If semantic roles can be assigned unambiguously because of the semantic nature of the referents, and the A argument also precedes the O argument, then no marking appears on a core argument.<sup>9</sup> But if the preverbal constituent order is reversed, so that the O argument precedes the A argument, then dative marking on the clause-initial undergoer argument becomes obligatory to override assumptions that the most topical NP represents the A argument. Example (22a) demonstrates the expected use of dative case to mark the recipient NP of a three-place predicate, and example (22b) shows its innovated use as a disambiguating marker on the O argument in a bivalent clause. This is motivated by the fact that the agent of ‘kick’ could be either of the referents, as both are animate. Conversely, (22c) demonstrates that the dative marking does not occur on the O argument when referents are acting in ways consistent with expected behavior.

The variable use of dative case marking on a core argument is widely reported across many languages, including Bantu and the Romance and Indo-Iranian branches of Indo-European (e.g. Hopper & Thompson [1980: 256]; Bosson 1991, *inter alia*). Since dative marking on an O argument can also be exploited for the discrimination of semantic roles – as demonstrated by the Apatani data – in many respects it mirrors the disambiguating marking functions of PMA marking outlined in Section 3.

## 5. Case marking in artificial languages

To summarize our findings thus far, it has been shown that non-paradigmatic PMA marking is rather more widespread in TB languages than was previously appreciated. Our fundamental conclusion is that the purpose of the agentive marking is disambiguation, which develops from the innovative use of a spatial case marker in a metaphorically extended core function.

Diachronic evidence from Tibetan and Ao suggests that a lexical noun in an appositional or genitival construction first becomes grammaticalized as a local case marker/postposition, and then is subsequently pressed into a more abstract role as a disambiguating core case marker, but only when non-canonical constituent orders or possible confusions over the assignment of semantic roles may result in misinterpretations of meaning. Virtually all of the other uses of agentive marking discussed in Section 3 can be attributed to further extensions of this basic disambiguating function. The presence of agentive marking in these languages is convincingly demonstrated to be unrelated to syntactic valency, which accounts for why it is found on both A and S arguments of bivalent and monovalent clauses under pragmatically predictable conditions.

The question now is whether these findings can be generalized, and whether other case-marking systems can be shown to develop in similar ways. The next sections will discuss a series of computer simulations of the emergence of case marking leading to similar results indeed. These computer simulations are run using the R package *MoLE*, for modelling language evolution, which is available from CRAN (<http://www.cran.r-project.org/>).

Given the methodological break, it is probably best to start this second part of the paper with a few disclaimers and general remarks on using computer simulations in linguistics. The main goal of computational models of language evolution is to show how aspects of language may develop as a result of more general principles (for recent overviews, see e.g. Smith 2014, Gong & Shuai 2013, Jäger et al. 2009, Cangelosi & Parisi 2001). Note that this can only be used as a proof of concept: a positive result does not mean that this is the way things must have gone, it only shows the feasibility of certain phenomena to emerge from the assumptions implemented in the model. But the more linguistically and typologically informed the model, the more interesting the results, of course. Unfortunately, evolutionary models mostly are fairly abstract. In *MoLE*, for example, agents lack spatial concepts and hence case markers cannot be derived

---

<sup>9</sup> The only exception to this is when the referent of the O argument is specific, in which case dative marking is used.

from spatial notions (cf. Section 4). Also, at least in the present study, there is no interaction between the use of case marking and word order. Thus, whereas in Qiang the less agent-like participant is marked if it qualifies for the more agent-like role in principle *and* precedes the actual agent in linear order (cf. example [4]), only the former motivation plays a role in the model (but see Lestrade *to appear*). Finally, the present simulation is restricted to the marking and disambiguation of generalized semantic roles. Note however that the encoding of other meaning aspects can be understood in the same vein: speakers check whether all meaning aspects of a proposition are likely to be understood by their hearers. If they think some relevant ingredient will not come across, whether this concerns a role or something like intentionality, this needs to be encoded explicitly.

On the basis of the previous section, one may be led to think that languages simply choose between marking the A or O argument if an utterance needs elaboration, the options being equivalent for disambiguation purposes. As will be shown in the simulations, however, things are a bit more complex: The available markers to choose from first have to be developed themselves, and their availability obviously constrains the options. In the beginning of the simulations, as for natural language, arguably, there are no case markers available. Instead, lexical ad-hoc solutions are used to resolve imminent communicative failure. For example in (23), the noun *pokynut* ‘teacher’ is used to mark *dutypog* ‘boy’ for its teaching role. Note that these glosses are for convenience only, as meaning in the model is completely abstract (cf. Section 7.1).

- (23) *dutypog pokynut nisagufy rapykume*  
 boy teacher man teaches.V  
 ‘The boy is teaching the man.’

In the model, the eventual development of case markers involves the initial recruitment and subsequent grammaticalization of lexical items such as *pokynut* in (23). These are words with a very specific referential meaning, which are used in combination with small sets of verbs only (viz. those with similar role specifications; e.g. *instruct* for the current example). Only later some of these words may become more general in meaning and shorter in form in the process of grammaticalization. Importantly, recruitment itself depends on more general solution strategies: If no grammatically determined solution method has been arrived at yet, which argument should be marked (using ad-hoc lexical expressions): the worst performer, the one for which the simplest solution is available, or just either of them? These different options will be entertained in Section 7 below. Before we can discuss the results of these strategies, however, a brief introduction to the model will be necessary.

## 6. Introduction to MoLE

The general idea and architecture of the model have previously been introduced (Lestrade 2015a/b, 2016a, to appear), and this section will largely be an (updated) recapitulation of that. Readers already familiar with the model may wish to skip this section, picking up at Section 7.2, which introduces the *solutionMethod* parameter that is newly developed for the purpose of this paper.

In the model, agents talk, procreate halfway through their lives, and die after 2000 utterances. They consist of a lexicon, distinguishing between object and action words, a usage history that keeps track of all contexts in which words have been used (cf. Bybee 2010), and a “language-ready” brain, which basically means that the agents have a desire for communicative success and are capable of joint attention (Tomasello 2003, Arbib 2015). In addition, during conversation, agents build up a common ground consisting of the set of recently used referential expressions.

The general conversation procedure is given in Figure 1. Two agents are selected for conversation and find themselves in a situation that consists of a set of randomly created events (to be qualified below). One of the agents wants to refer to one of these events, the *target event*, for which it formulates an utterance (roughly using the incremental and modular procedure described in Bock & Levelt 1994). If the speaker agent thinks some part of the intended meaning is insufficiently clear, it will elaborate to make it explicit (Steels 2003, Blutner 2007, Zeevat 2007, de Swart 2007, Lestrade 2010). Initially this involves the use of lexical ad-hoc marking through simple lexical apposition for the lack of conventional solutions, as shown in (23) above; once case marking develops, however, this will be used instead. Next, the hearer has to analyze the utterance by identifying the words and their functions, determine the intended meaning, and point out which event in the situation it thinks the speaker referred to. If correct, the common ground, usage history and lexicon will be updated, and the conversation will continue. If they fail, the conversation will continue without updating.

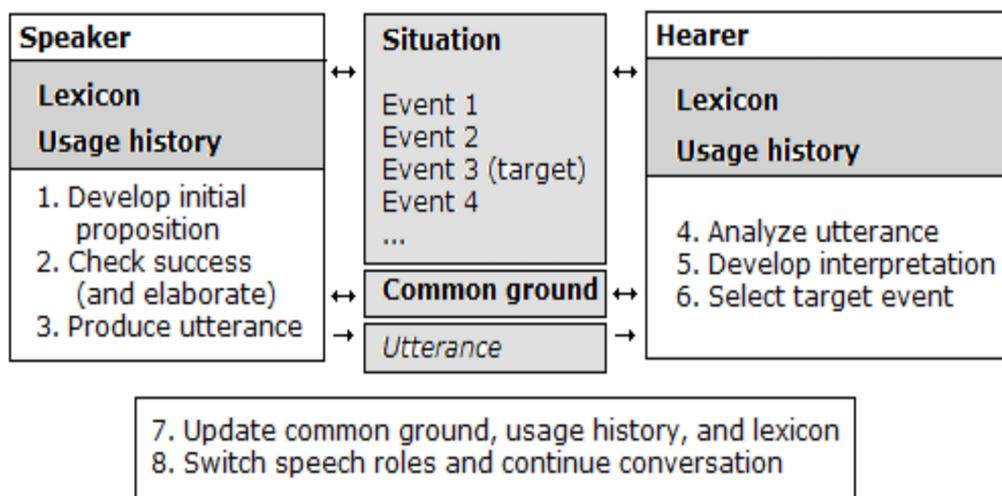


Figure 1: Conversation procedure

In what follows we will go through the different steps and ingredients in more detail. We will start with the lexicon, which, as in other recent proposals about the cognitive architecture of language, is put center stage (Vosse & Kempen 2000, Kempen 2014; Jackendoff 2002, 2011; Culicover & Jackendoff 2005). Since agents live in a virtual world, their meaning representations are rather different from ours. Fortunately, it seems possible to implement this in a relatively valid way nevertheless. According to Wierzbicka (1996), natural-language concepts can be decomposed into meaning primitives such as CONCRETE, HUMAN, MALE, etc. (cf. also Katz & Fodor 1963 and Guiraud 1968). Similarly, in a way, Gärdenfors (2000) argues that concepts are sets of values on different meaning dimensions. Thus, we can think of a cat as something that is time-stable, concrete, alive, four-legged, tailed, etc. Note that whereas the initial meaning dimensions in such characterizations are very general and bisect the world (e.g. time-stability), eventually meaning dimensions become more and more specific in order to single out a concept (e.g. having a tail). Abstracting away from the quality of the dimensions that organize our mental lexicon, the object lexicon of the agents is modeled as a list of randomly generated forms with values on nine numerical meaning dimensions (their *vector representation*). An example is given in (24). Following the observation just made, the dimensions make an increasing number of distinctions (with values between zero and one). These dimensions may be taken to represent whatever properties are (grammatically) relevant for the linguistic behavior of words in natural language, but the model does not commit to any such specific interpretation *per se*. Action words

can be similarly specified (their dimensions representing different semantic features), with the eventual addition of one (for one-place predicates) or two predicate roles (for two-place predicates), which are represented using vectors as well, cf. (25). In the remainder, object words are sometimes called *nouns* and action words *verbs* for the sake of convenience, which should be understood in a theory-neutral way.

(24) Representation of nouns  
*dotyjytit*      1 1 0 1 0 0.000 0.250 0.25 0.00

(25) Representation of verbs  
*tomorawog*    1 0 0 1 1 0.375 0.875 0.000 1.000  
 role 1        1 1 1 1 1 0.375 0.875 0.750 0.250  
 role 2        0 0 0 0 0 0.500 0.50 0.875 1.000

The meaning dimensions of nouns and verb roles correspond to one another in the model. That is, the first predicate-role dimension concerns the same property as the first noun dimension. These correspondences are used to determine the role qualifications of the nouns (or *typing scores*, after Aristar 1997). The more the vector of a noun resembles that of a verb role, the more it qualifies for it.

In the model, higher values, i.e. those closer to 1, are considered more prominent (for nouns) or agent-like (for roles) than lower values. Note that a correspondence is thus assumed between referential prominence and agentivity. That is, being prominent means being a typical A, and vice versa, and similarly for non-prominent Os.<sup>10</sup>

As exemplified by the verb representation in (25) already, if a verb has two role vectors, one of them contains higher values than the other on average. This is meant to reflect the semantic role distribution in natural languages, in which the predicate roles of two-place verbs generally can be distinguished in terms of agentivity, one role being more agent-like than the other (Dowty 1991, Van Valin 1999, Yip et al. 1987).

Just like meaning representations, events consist of sets of vectors, one containing the referential properties of the action, one for the more agent-like event participant, corresponding to the A argument, and one for the less active one, which will be expressed as the O. Events are generated *randomly*, as said above, but not entirely *at random*. In principle, events are made up of actions and objects that could be expected to occur together, subject to a bit of noise. Nouns in the common ground are more likely to figure in the events, and the distribution of prominence values over A and O roles follows that of natural-language (as established by Dahl 2000). The distribution for two-participant events is given in Table 1. Roughly, and as well-known, As are mostly prominent (speech-act participants or animate), Os mostly non-prominent (inanimate).<sup>11</sup>

<sup>10</sup> In Lestrade (2015, 2016a/b, to appear), these roles are referred to as the *Actor* and *Undergoer* participant, respectively (after Van Valin 1999). For consistency with the first part of the paper, *A* and *O* are used here.

<sup>11</sup> The numbers in Table 1 are based on the following distribution of person over roles as established by Dahl (2000), in which all non-prominent third person participants are classified as *inanimate* and all others as *animate*. Note that Dahl's study really concerns grammatical rather than semantic roles, which are equated here for convenience.

Role	Person			
	1	2	prominent 3	non-prominent 3
Subject	38%	22%	33%	7%
Direct object	3%	3%	10%	84%

role	prominent	non-prominent
A	93%	7%
O	16%	84%

Table 1: Animacy and (generalized) semantic role (after Dahl 2000)

When developing an utterance, the agents have to select referential items on the basis of the vector match between their meaning representations and the actual referents to be described. In principle, the form of the concept that best describes a referent will be selected to describe it. In addition to semantic match, however, frequency and economy considerations play a role: Frequently and recently used (“activated”) words are preferred, as are semantically and formally light expressions (see below for how words may become “light” as a result of grammaticalization). For each meaning, word candidates are ranked according to their activation and vector match, and the first word found to be *sufficiently distinctive* is selected for expression.

Whether a word is sufficiently distinctive depends on the number of distractor events and objects in the contextual situation. If there are many events ongoing, it is generally necessary to be more specific to prevent referential ambiguity (e.g., if there’s more than one object present, *thing* does not distinguish between them). Once the referential expressions have been selected, the initial proposition is ready (cf. step 1 in Figure 1).

Although not entirely uncontroversial, it seems that human communication is bidirectional, speakers taking into account their hearers in order to be understood (as was already assumed in the previous step, in fact). In many theoretical and computational models, this involves speakers pretending to be hearers to check if they themselves would get the right meaning (cf. Grice 1975, Levelt 1983, Hurford 1989, Steels 2003, Blutner et al. 2006, de Swart 2007, Zeevat 2007, Lestrade 2010). Economy and predictability are generally considered two important factors in this process. Speakers try to use as economic an utterance as possible (i.e. costing little pronunciation effort). Only if this leads to the wrong result, the utterance should be made more explicit (cf. Grice 1975). As discussed above already, the meaning dimensions of nouns and verb roles correspond to one another in the model. That is, the first predicate-role dimension concerns the same property as the first noun dimension. By comparing these vectors, the degree to which an argument qualifies for its role can be determined. Whenever a participant’s typing score for its intended predicate role is significantly higher than that of the other participant and that for the other role (i.e. under the reverse distribution of roles), an agent can simply combine the words that refer to the event and event participants. The argument structure then follows from their semantics (i.e., even in the absence of functional word order, we know who’s doing what in *book man read*). But if the intended role filler does not qualify better than the other argument and hence the argument structure does not follow automatically (cf. *man boy see*), something extra needs to be done. One solution is to make the role of one of the participants explicit by naming it using another word that describes this role best (cf. the notion of *differential case marking*, cf. e.g. Bossong 1985, Aissen 2003, de Hoop and de Swart 2008; another solution, which is not entertained in the present experiment, is using word order). Once the role distribution is checked and, if necessary, made explicit through additional marking, the (elaborated) proposition can be produced (steps 2 and 3; cf. Section 7.2).

The hearer now has to decompose the utterance, distinguishing the referential expressions from the (ad hoc) role markers. First, it has to be determined what the verb is in the utterance. This is done by first determining which verb lexeme in the lexicon each word in the utterance could correspond to, and next determining which of the actions in the ongoing events are best

described by the identified lexemes. The word with the best-combined form and action match is considered the verb. Depending on the type of verb (i.e. one- or two-place) and the number of remaining words, different analyses might be possible. If, for example, the first word has been identified as a two-place verb and three words remain, either the third word could be a role marker of the second, or the fourth could be the role marker of the third (assuming that role markers are adjacent and follow their hosts; cf. Givón's 1995:188 *proximity-relevance* constraint). Alternatively, one or more words could be left unanalyzed, which is penalized but may eventually lead to the preferred interpretation nevertheless. All grouping possibilities are entertained and interpreted by looking up and combining the semantics of the ingredients (Steps 4 and 5).

Next, for each interpretation it is determined how well it describes each of the ongoing events. The interpretation with the best match is considered the intended meaning and the corresponding event is considered the target event (following the setup of Steels 1997). If the hearer correctly identifies the target event, then the lexicon, common ground, and usage history are updated (Step 7). If not, nothing happens. Finally, agents may switch speech roles and continue their conversation (Step 8) or stop talking, after which two new agents are selected.

Initially, all words are fully specified on all meaning dimensions and have a word length of between 8 and 10 characters (as in examples [24] and [25] above). Over time, however, words can change. For this, a number of grammaticalization principles are implemented (Heine & Kuteva 2007), most importantly *erosion* and *desemanticization*. Words can be "pronounced" sloppily if they are frequent or predictable (cf. Jurafsky et al. 2001 and Balota & Chumbley 1985). In the model, sloppy pronunciation is instantiated simply by deleting the final character of a word. This does not lead to a change of lexical representation for the agent using the form, but if new generations of speakers attest only the shorter variant, they may set their lexical representation accordingly, because of which a form may erode (Nettle 1999). If words become too short to stand on their own, they are fused with their host, first as a clitic, and eventually as a suffix (Bybee 1985). In the model, this concerns those words for which the total production effort (i.e. the number of characters) has fallen below the suffix threshold of 6. Note that, in principle, the difference between bound and non-bound marking (or between suffixes and postpositions) is thus understood as one of "phonological" degree only. However, for erosion to get this far, something has to happen on the meaning side too.

Frequent words are likely to desemanticize. At first, they may extend their meaning range incidentally (if the context does not require a more specific description or in the absence of a better expression). Eventually, such an extension may become a standard part of a word's meaning, as a result of which it becomes more general. In the model, desemanticization involves the progressive removal of the meaning dimensions of a word. Shortly before they procreate, agents consider the variation in their usage history and may decide that certain meaning dimensions should either be changed or be removed from the meaning representation of a word (Bybee 2010).<sup>12</sup> A meaning dimension of a word is changed into another value if this latter value is found significantly more representative for the meanings for which this word has been used. Deletion takes place if some meaning dimension is found to be unrepresentative or inconsistent with the uses of the word, but there is no dominant value on that dimension among the uses. In addition, certain frequency thresholds need to have been reached: for a first dimension to be removed, a word has to be combined with 2% of the relevant predicates. This proportion grows exponentially to 40% for the last dimension to be removed. The less dimensions specified, the more general or abstract the meaning, and the more usage possibilities. Together with the

---

<sup>12</sup> In reality, updating meaning representations on the basis of actual usage is probably continuously ongoing, but using a single moment for this has the same effect and is easier for modeling purposes.

attenuated phonological form, a generalized meaning constitutes what could be understood as the model equivalent of a case marker: a maximally short form with a maximally general meaning (Lestrade 2010).

In addition to erosion and desemanticization, a third grammaticalization principle is implemented, viz. *generalization*. Agents not only keep track of the referential meanings of words, they also do this for the type of constructions they have used. Thus, a score is kept for how often As and Os were marked for their role. More specifically, agents keep track of the meaning values that are involved in this. For example, they know how often As with a 0 value on the first meaning dimension were assigned a marker (and similarly for all other dimensions, values, and Os). On the basis of these scores, rules may develop. If, for example, an agent finds out that certain As virtually always turned out to need marking, it may decide to do so straight away, irrespective of the necessity in context (cf. Durie's 1995 *functional overkill*; cf. also Dolan and Dayan 2013, and Lestrade et al. 2016). Obviously, once a parent generation decides to do this, the input language for the children becomes even more skewed, increasing the probability for the next generation to make the same generalization.

Note that generalization can result in case-marking systems in which certain intransitive subjects are marked too, for example marking them for their willful involvement, as in Mongsen Ao discussed above: if marked As happen to be almost always volitionally involved, their marker may become associated with volitionality, because of which it is no longer primarily used for disambiguation. If an S now is found to be exceptionally volitional, this can be expressed using the original A marker. In terms of the model, if virtually all marked As have a 1 on the third meaning dimension, this meaning may trigger the use of A marking irrespective of ambiguity and transitivity.

This much technical background should suffice to appreciate the results of the simulations discussed below. More information can be found in the help files of the R package.

## 7. Present experiment

### 7.1 Underlying solution methods for resolving ambiguity

In the present experiment, the effect of different type of disambiguation strategies will be investigated by manipulating the *solutionMethod* parameter that was developed for the purpose of the present paper. As explained in the previous section, a speaker checks whether the intended distribution of roles follows automatically from lexical semantics or needs to be made explicit. For this, different strategies are conceivable. In previous MoLE experiments, it was always the second argument that was marked for its role in case of ambiguity, the idea being that the ambiguity only becomes evident once this second participant is introduced, and only has to be resolved there and then. But the role distribution could also, and maybe more likely so, be made explicit by using the best marker available in the lexicon, marking the worst performer, or randomly marking either of the arguments. Other solution methods are conceivable and indeed possible to test, such as marking only the internal argument, or both arguments. These strategies seem to be less plausible, however, and therefore will not be considered here (for example because internal roles are generalizations that first have to be developed itself, and marking either of the arguments suffices to resolve ambiguity, as observed in Section 2 already).

The different types of language resulting from the different disambiguation strategies that are investigated in this study are listed in (26). All other model parameters are held constant between lineages.

## (26) Disambiguation strategies

- a. *Best marker (B)*: use the best marker available
- b. *Random (R)*: mark either the A or O
- c. *Second argument (S)*: mark the second argument
- d. *Worst performer (W)*: mark the worst performing participant

## 7.2 General results

Each solution type is tested with 6 lineages, each of which is run for 25 generations. As shown by the development of communicative success in Figure 2, three out of four lineages maintain a communicative-success rate of roughly 90% over time. This is as expected: The strategies themselves do not change over time and the different strategies are all viable solutions in principle. In the other 10% of the cases, the speaker wrongly believes the hearer should be able to find out who does what to whom. This is most probably due to changing lexical representations (because of desemanticization), because of which the estimations of role qualifications may differ between agents. Interestingly, however, this does not apply to the B lineages, in which communicative success does increase over time.

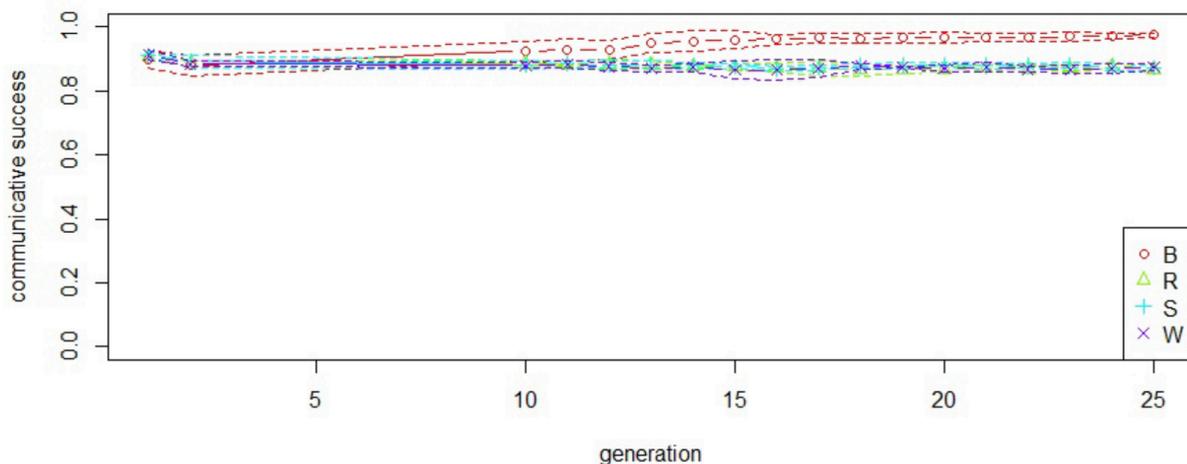


Figure 2. Development of communicative success per strategy type over time.  
(Each type is instantiated by six lineages. Dashed lines show standard deviations).

Given the default noise settings and prominence over role distribution, some form of role marking is used in approximately 42% initially in all lineages (with a standard deviation of less than 4%; see Figure 3). As noted above, this involves events in which one of the arguments qualifies at least equally well for the other verb role, which is the case whenever the O argument is equally prominent as or more prominent than the A argument (which, again, is generally the case for prominent Os and non-prominent As). As can be seen in Figure 3, this proportion remains stable over time in three out of four lineages. In the B lineages, however, the use of case marking is generalized beyond communicative necessity, because of which it becomes obligatory eventually. As a result, the role distribution is always made explicit, even if deemed unnecessary for disambiguation. Since this also applies to the 10% of cases that are wrongly judged unambiguous, this leads to maximal communicative success (cf. Figure 2).

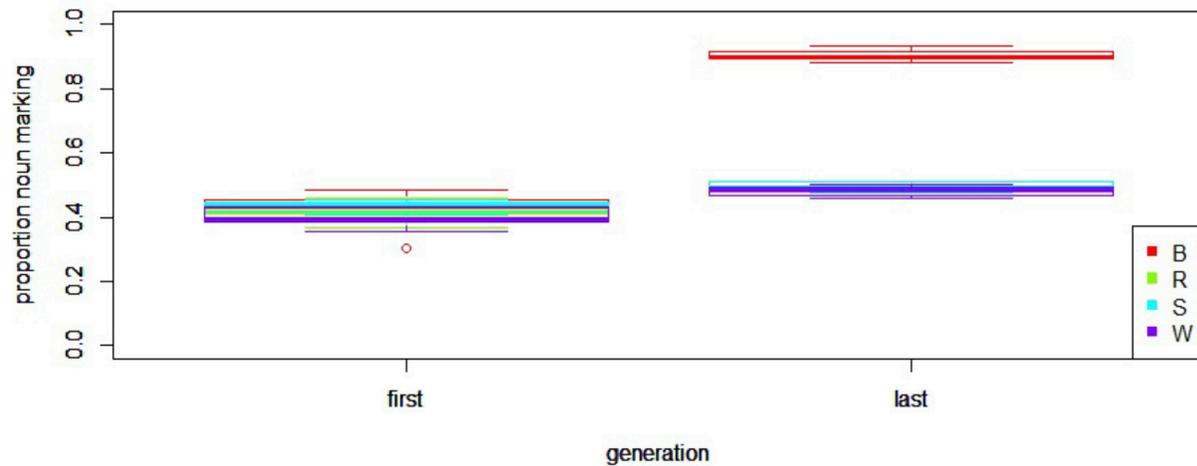


Figure 3: Proportion of noun marking per strategy type for the first and last generation.  
(Only in the B lineages, the use of case marking becomes generalized)

### 7.3 Emerging case systems

In the tables below, the case systems that have emerged under the various solution strategies are shown. Case markers are identified as such if they are used as role markers more than as referential expressions, if they have lost all four of their non-binary meaning dimensions, and if their production effort has fallen below the suffix threshold. The *lin(eage)* column shows the lineage, *form* gives the actual form and *etymology* shows the original form of the marker. The *frequency* column shows the total frequency of use of the word in any function (out of the 2000 utterances used by the agent), and the NM column shows the usage as a noun marker. The *weight* column shows the proportion of dimensions that are still specified (recall that each word starts out fully specified; the lower the weight, the more generalized the semantics). The *meaning* column, finally, gives the mean value on these remaining meaning dimensions. Again, higher values correspond to actor properties and therefore markers with scores close to zero can be understood as O markers (or “accusative cases”, once grammatical roles are distinguished), whereas markers with scores close to one can be understood as A markers (“ergative cases”).

Table 2 shows the result for the lineages of the random type (i.e. those that randomly mark either of the arguments in case of imminent communicative failure). As can be seen, all six lineages developed two role markers, one for each generalized semantic role.

lin.	form	etymology	freq.	NM	weight	score
1	-po	pokynutuj	458	372	0.56	1
	-yj	yfarunapow	490	394	0.56	0
2	-ub	umymirem	540	414	0.44	0
	-ry	rylifupywy	367	301	0.56	1
3	-mi	mipumigo	438	369	0.56	1
	-my	myfapyjes	415	335	0.56	0
4	-iw	ifakulihu	449	377	0.56	1
	-ik	iwobyfufe	496	381	0.44	0
5	-hu	hywanarin	453	328	0.56	0.2
	-ny	nyjanago	443	387	0.56	1.0
6	-ta	talodadahe	380	315	0.56	1
	-sy	syjikogesu	508	390	0.44	0

Table 2: Most frequently used noun markers of lineages with random solution method.

[*Lin(eage)* gives the lineage; *form* shows the actual form and *etymology* the original form of the marker; *freq(ueency)* shows the total frequency of use of a word (out of 2000 utterances), *NM* the frequency of use as a noun marker. The *weight* column shows the proportion of dimensions that are still specified (the lower the weight, the more generalized the semantics); the *meaning* column gives the mean value on these remaining meaning dimensions.]

Table 3 shows the case systems that develop under the *second-argument* strategy. The results are very similar, except for the absence of a dedicated case marker for the A role in the sixth lineage. This absence follows from the technical requirements that markers should no longer be specified for a non-binary meaning dimension. Because of this, the suffix *-ro*, which otherwise is a perfect A marker but still has one such specification, is not listed.

lin.	form	etymology	freq.	NM	weight	score
1	-py	pijyfiwud	220	178	0.56	1
	-af	afutomabuj	647	489	0.44	0
2	-oh	ohusamido	628	533	0.44	0
	-om	omypyhub	240	184	0.56	1
3	-ub	uwegaser	257	197	0.56	1
	-ku	kuheleki	626	534	0.44	0
4	-if	ifinygobis	625	538	0.44	0
	-as	asunilare	222	159	0.56	1
5	-my	mygutawil	646	525	0.44	0
	-im	imisilet	291	228	0.56	1
6	-ym	ymisadigi	606	493	0.44	0

Table 3: Results for lineages with second-argument strategy.

Also the results for the lineages that resolve ambiguity by marker the worst performer are largely equivalent to the previous ones. Each lineage developed markers for both roles, using them roughly equally often.

lin.	Form	Etymology	freq.	NM	weight	score
1	-yk	yharenyp	489	364	0.44	0
	-no	nubiwewym	441	368	0.56	1
2	-by	balapygu	413	333	0.56	1
	-be	betifofup	455	391	0.56	0
3	-uh	uniharij	494	388	0.56	1
	-yb	ybihowab	460	363	0.56	0
4	-ak	awigarynu	503	370	0.56	0
	-oj	ohetuwur	334	282	0.56	1
5	-ud	udolujef	392	314	0.56	1
	-ag	arasipew	507	371	0.44	0
6	-gy	gynadobo	470	394	0.56	0
	-ig	ifibiwetaf	454	324	0.56	1

Table 4: Results for lineages with worst-performer strategy.

The results of the *best-marker* lineages are qualitatively very different from the other three lineages. Most strikingly, only one type of case marker developed per lineage. Since it is always the best marker that is selected, a small initial advantage may lead to exclusive usage. As was observed in Figure 3 already, the use of these markers becomes generalized beyond communicative necessity only for this type of lineage. Because of this, these markers are much more frequent than the markers in the other lineages. And as a result, they are all maximally desemanticized. Whereas some of the previous markers have a semantic weight of .56; here, all markers are at .44.

lin.	form	etymology	freq.	NM	weight	score
1	-uh	uhemabir	1782	1609	0.44	0
2	-ab	anoruduwim	1677	1511	0.44	0
3	-ji	jikemywo	1639	1518	0.44	0
4	-ih	ihifysowa	1433	1341	0.44	1
5	-ow	owupahywyk	1756	1603	0.44	0
6	-op	opelotato	1759	1602	0.44	0

Table 5: Results for lineages with best-marker strategy.

#### 7.4 Disambiguation of semantic roles in artificial languages

In the previous sections, it was explained how the distribution of semantic roles is disambiguated in the various lineages, and what types of case systems emerge from this over time. In this section, some concrete examples of the end results in the different lineages are provided. To construct these examples, the argument structure of a proposition that led to a marked expression is simply reversed, after which the speaker agent develops a new utterance for that meaning. (Using example [3]: if a speaker uses role marking to make explicit that it really is a pig killing a woman, how do they express a woman killing a pig?) Recall that no case marking is available whatsoever initially (cf. [23]), word order is not informative at any time, and translations can

only give an idea of the argument structure, as the meanings cannot be translated beyond their numerical representations (which would not be of much help for intelligibility).

In (27), an example from one of the R lineages is given, in which it is randomly determined which of the arguments is marked if the role distribution does not follow automatically. *Opifep* (say ‘pig’) is a much better performer of the O role of the verb *oli* (‘kill/bite’) than of its A role (as can be seen when checking the respective typing scores of .78 and .42), and, equally importantly, a much better performer of the O role than the other argument *huwenahe* (‘woman’, with a score of .54). Without explicit marking, then, the interpretation is that *opifep* is *oliing* rather than being *olied*, as shown in (27a). To overrule this default interpretation, explicit marking of the role distribution has to be used. Given the random solution strategy of this lineage, this could either involve marking *opifep* for its A role, using the A marker *-po* (27b) or marking *huwenahe* for its O role, by means of the O marker *-yi* (27c). Again, these markers were not available in the beginning, but developed as a result of erosion and desemanticization in the course of the simulation only.

(27) Random marking

- a. *opifep oli huwenahe*  
*opifep oli.V huwenahe*  
 ‘Huwenahe olies opifep.’ (cf. *the woman killed pig*)
- b. *opifep-po oli huwenahe*  
*opifep-A oli.V huwenahe*  
 ‘Opifep olies huwenahe.’ (cf. *the pig bit the woman*)
- c. *opifep oli huwenahe-yj*  
*opifep oli.V huwenahe-O*  
 ‘Opifep olies huwenahe.’ (cf. *the pig bit the woman*)

Lineages of the S type always mark the second argument if a meaning does not follow automatically. An example is given in (28). Without marking, the interpretation is that *wywakak* is the A and *kalakip* is the O (28a). To say that *kalakip sobojifeys wywakak*, the O marker *-af* is used, since it’s the O argument that comes second (28b). If the A argument comes second, the A marker *-am* can be used, as shown in (28c). Thus, in this lineage too, distinctive markers for the A and O role developed, which are only used when necessary for communication.

(28) Second argument

- a. *kalakip sobojifehy wywakak*  
*kalakip sobojifey.V wywakak*  
 ‘Wywakak sobojifeyes kalakip.’
- b. *kalakip sobojifehy wywakak-af*  
*kalakip sobojifey.V wywakak-O*  
 ‘Kalakip sobojifeyes wywakak.’
- c. *wywakak kalakip-am sobojifehy*  
*wywakak kalakip-A sobojifey.V*  
 ‘Kalakip sobojifeyes wywakak.’

The lineages of the W type always mark the worst-performer if disambiguation is necessary.

The minimal pair in (29a,b) shows the use of the O marker *-yk*. If it is the performer of the A role that qualifies worst instead, the A marker *-no* can be used, as shown in (29c).

- (29) Worst performer
- a. *ro dadises-yk habages*  
 ro.V dadises-O habages  
 ‘Habages roes dadises.’
- b. *ro dadises habages*  
 ro.V dadises habages  
 ‘Dadises roes habages.’
- c. *onoteuk dywes-no lahogohe*  
 onoteuk.V dywes-A lahogohe  
 ‘Dywe onoteuks lahogoe.’

Differently from all lineages of the previous types, the lineages using a best-marker strategy all generalized their use of case marking. As a result, both O arguments of the minimal pair in (30) receive case marking, irrespective of their role qualifications and irrespective of any possible ambiguity.

- (30) Best marker
- a. *jajuhup yhimomyd-uh ygoj*  
 jajuhupuh yhimomyd-O ygoj.V  
 ‘Jajuhupuh ygojes yhimomyd.’
- b. *Jajuhup-uh yhimomyd ygoj*  
 Jajuhup-O yhimomyd ygoj.V  
 ‘Yhimomyd ygojes jajuhupuh.’

Interestingly, generalization in the best-marker lineages leads to a use of agentive case marking comparable to the PMA marking discussed in Section 3. For example in (31), the intransitive subject is marked for its actorhood, where there obviously is no role ambiguity possible:

- (31) *yluweju-ih uryboroju*  
 yluweju-A uryboroju  
 ‘Yluweju uryborojues.’

## 7.5 Discussion of modeling results

In the previous sections it was shown how case marking may emerge in artificial computer languages in which initially only lexical means are available. Different solution methods for disambiguation that could be hypothesized to underlie the eventual use of case marking were tested. Intuitively, the strategy to randomly mark either of the arguments seems least plausible. That is, it can be assumed that in the absence of grammatical rules, speakers still use a culturally determined convention as to how to say things properly. Cross-linguistically, next, it seems the strategy to always mark the second argument is not very likely either: As far as we know, no such correlation has ever been reported in the typological literature. Instead, there is something to say for the other two strategies: as for the best-marking strategy, it seems a valid linguistic

procedure to reuse solutions that have proven to work before previously, whereas the worst-performer strategy squares well with independently motivated *relevance* principles, which say to mark features on the relevant constituents (cf. Malchukov 2006). Empirically, it seems, the best-marker strategy gave the best results. As far as we know, most natural languages develop only one type of case marking for either generalized role, not for both. Moreover, it resulted in case marking very comparable to PMA marking in TB. Instead, all but the best-marker lineages developed both types of marking, making them less likely. Note that at the same time, however, most languages do *not* seem to generalize their use of case marking across the board (Sinnemäki 2014), suggesting a counterforce that is not properly presented in the model (yet).

Obviously, such findings should be treated with care indeed and the results model cannot be translated to natural-language change directly. For example, the solution strategies should probably be considered stochastic and interacting rather than absolute and mutually exclusive parameters. Also, alternative marking strategies (such as word order or verb marking), and person distinctions (on the basis of which different marking systems may develop) should be taken into consideration (as is done in current experiments). Nevertheless, it is important to notice that the emergence of case marking can be modeled as resulting from a clear pragmatic motivation, which nicely corresponds to the findings in natural language discussed in the first part of the paper.

As a final point of discussion, the predominance of O marking in the best-marker lineages should be mentioned. This could be merely coincidence, of course, and six lineages are not enough to tell whether one is more likely than the other. Cross-linguistically, at least, there seems to be no clear preference for either type of marking. If the preference for O marking in the model is found genuine, this means disambiguation is not enough to explain the emergence of case marking.<sup>13</sup>

## 8. General conclusions

In this paper we discussed the development of case marking in Tibeto-Burman and artificial computer languages. Our fundamental conclusion is that the purpose of the agentive marking is disambiguation, which develops from the innovative use of a more concrete (spatial) marker in a metaphorically extended core function. Our survey also shows that pragmatically motivated case marking is more widespread in TB languages than was previously appreciated. Examples were given of uses to disambiguate semantic roles in bivalent clauses; for contrastive reference, i.e. to pragmatically distinguish a referent from one or more other potential referents, or to foreground a referent; to encode extraordinary, willful, or atypical behavior, especially in the context of an unexpected turn of events; to clarify the semantic role of a remaining core argument when another core argument undergoes ellipsis; to distinguish the causer argument in causativized clauses, and to characterize a habitually acting NP referent. In the simulation experiment, several underlying strategies that may lead to the development of pragmatically motivated case marking

---

<sup>13</sup> Note that the absence of a cross-linguistic preference for O marking could be used to further argue against the W and S strategies. Given the Dahl numbers of Table 1, most confusion is due to prominent Os (viz. .93x.16 vs .07x.84). Under a worst-performer strategy, O marking should then be more likely to develop than A marking. Indeed, in most W lineages, O seems more frequent than A marking (cf. Table 4), although it would require further investigation to show this effect is statistically significant. (This reasoning does not apply to the B type, which does not care about the source of the confusion but only considers the best solution.) As for the S strategy: if word order played a role in the simulations (which it does in current MoLE experiments, as it does in natural languages; cf. the first part of this paper), O marking should be predominant too, since As tend to appear before Os cross-linguistically.

were tested. It was shown that a strategy that employs the best marker available leads to the most promising results.

## Acknowledgements

We are grateful to two anonymous reviewers for their suggestions and comments on an earlier draft, but fully absolve them of responsibility for the analysis and conclusions. Research by Coupe for this paper was generously supported by an Alexander von Humboldt Fellowship for Experienced Researchers (2016–2018), and a Singapore Government Ministry of Education AcRF Research grant (MOE 2016-T1-001-220). Research by Lestrade was supported by the Netherlands Organisation for Scientific Research (*NWO*; grant 275-78-001).

## Abbreviations

ABL	ablative	EVID	evidential	NRL	non-relational noun prefix
ABS	absolutive	EX	existential	NZP	nominalizing prefix
ANAPH	anaphoric demonstrative	EXCLM	exclamation	OBL	oblique
AOR	aorist	F	feminine	ONOM	onomatopoeia
ASSURE	assurance particle	FGR	foregrounder	ORD	ordinal
AUX	auxiliary	GPN	generic pronoun	PFV	perfective
AGT	agentive	HM	hesitation/pause marker	PL	plural
CAUS	causative	HON	honorific	POL	polite particle
COM	comitative	IMM	immediate future tense	POSS	possessive
COMPAR	comparative derivation	INST	instrumental	PROX	proximate demonstrative
CONT	contrastive	IRR	irrealis mood	PART	participle
COND	conditional	LN	loan word	PRES	present tense
COP	copula	LNMLZ	locative nominalizer	PST	past tense
CVB	converb	LOC	locative	PTCL	speech act particle
DECL	declarative	LOC.CVB	locative converb	RED	reduplication
DAT	dative	MID	middle	RPET	repetitive aspect
DIR	direction	MIR	mirative	SEQ	sequential converb
DIST	distal demonstrative	NEG	negative	SG	singular
DM	discourse marker	NF	non-finite	SIM	simultaneous converb
DUR	durative	NHYP	non-hypothetical	VIS	visual (direct) evidential
EGO	egophoric	NMLZ	nominalizer	VOC	vocative
EMPHAT	emphatic	NPROX	non-proximate		
EQ	equational	NPT	non-past		

## References

- Abraham, P.T. 1985. *Apatani grammar*. Mysore: CIIL.
- Arbib, M. A. 2015. Language evolution. An emergentist perspective. In: MacWhinney, B. & W. O'Grady (eds.), *The handbook of language emergence*, 600–623. West Sussex, UK: Wiley/Blackwell.
- Aristar, Anthony Rodrigues. 1991. On Diachronic Sources and Synchronic Pattern: An Investigation into the Origin of Linguistic Universals. *Language*, Vol. 67(1): 1–33.
- Aristar, Anthony Rodrigues. 1997. Marking and hierarchy. Types and the grammaticalization of case markers. *Studies in Language* 21(2): 313–368.

- Balota, D.A. & J.I. Chumbley. 1985. The locus of word-frequency in the pronunciation task: Lexical access and/or production? *Journal of memory and languages* 24: 89–106.
- Barðdal, Jóhanna & Shobhana L. Chelliah (eds.) *The role of semantic, pragmatic, and discourse factors in the development of case*. Amsterdam/Philadelphia: John Benjamins.
- Beames, John. 2012. Comparative grammar of the modern Aryan languages of India, to wit, Hindi, Panjabi, Sindhi, Gujarati, Marathi, Oriya, and Bangali. Volume 2: the noun and the pronoun. Cambridge: Cambridge University Press. First published 1875.
- Beyer, Stephan B. 1992. *The classical Tibetan language*. Albany: State University of New York Press.
- Blutner, R. 2007. Optimality theoretic pragmatics and the explicature/implicature distinction. In N. Burton-Roberts (ed.), *Advances in pragmatics*, 45–66. Basingstoke, etc.: Palgrave/MacMillan.
- Blutner, R., H. de Hoop, & P. Hendriks. 2006. *Optimal communication*. Stanford: CSLI.
- Bock, K. & W.J.M. Levelt. 1994. Language production: Grammatical encoding. In: M.A. Gernsbacher (ed.) *Handbook of Psycholinguistics*, 945–984. London: Academic Press.
- Bossong, Georg. 1985. *Empirische Universalienforschung. Differentielle Objektmarkierung in der neuiranischen Sprachen*. Tübingen: Narr.
- Bossong, Georg. 1991. Differential object marking in Romance and beyond. In Wanner, D. & D. Kibbee (eds.) *New analyses in romance linguistics: Selected papers from the XVIII Linguistic Symposium on Romance Languages 1988*, 143–170. Amsterdam: John Benjamins.
- Bybee, J. 1985. *Morphology. a study of the relation between meaning and form*. Amsterdam/Philadelphia: John Benjamins.
- Bybee, Joan. 2010. *Language, Usage and Cognition*. New York: Cambridge University Press.
- Cangelosi, A. and D. Parisi (2001), "Computer simulation: A new scientific approach to the study of language evolution". In: A. Cangelosi and D. Parisi (eds), *Simulating the Evolution of Language*. London: Springer.
- Chelliah, Shobana L. 2009. Semantic role to new information in Meithei. In Barðdal, Jóhanna & Shobhana L. Chelliah (eds.), 377–400.
- Chelliah, Shobana L. 1997. *A grammar of Meithei*. Berlin/New York: Mouton de Gruyter.
- Clark, Edward Winter. 1911. *Ao-Naga dictionary*. Calcutta: Baptist Mission Press.
- Comrie, Bernard. 1978. Ergativity. In Lehmann, Winfred P. (ed.) *Syntactic typology: Studies in the phenomenology of language*, 329–374. Austin: University of Texas.
- Comrie, Bernard. 1989. *Language universals and linguistic typology*, 2<sup>nd</sup> edn. Oxford: Basil Blackwell.
- Comrie, Bernard. 2013. Alignment of Case Marking of Full Noun Phrases. In Dryer, Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/98>, Accessed on 2017-08-28.)
- Coupe, Alexander R. 2007. *A grammar of Mongsen Ao*. Berlin/New York: Mouton de Gruyter.
- Coupe, Alexander R. 2011a. On core case marking patterns in two Tibeto-Burman languages of Nagaland. *Linguistics of the Tibeto-Burman Area* 34(2): 21–47.
- Coupe, Alexander R. 2011b. Pragmatic foundations of transitivity in Ao. *Studies in Language*. 35(3): 492–522.
- Coupe Alexander R. 2017. Mongsen Ao. In Thurgood, Graham & Randy LaPolla (eds.) *The Sino-Tibetan languages* 2<sup>nd</sup> Edn, 277–301. Oxford & New York: Routledge.
- Coupe Alexander R. 2017. On the diachronic origins of converbs in Tibeto-Burman and beyond. In Ding, Picus & Jamin Pelkey (eds.) *Sociohistorical linguistics in Southeast Asia: New horizons for Tibeto-Burman Studies in honor of David Bradley*, 211–237. Leiden: Brill.

- Culicover, P.W. & R. Jackendoff, 2005. *Simpler syntax*. Oxford, etc.: Oxford University Press, 2005.
- Dahl, Östen. 2000. "Egophoricity in discourse and syntax". *Functions of Language* 7(1): 37–77.
- Dahl, Östen. 2008. Two pathways of grammatical evolution. *Symposium on the genesis of Syntactic Evolution*. Rice University, Houston.
- DeLancey, Scott. 1981. An interpretation of split intransitivity and related patterns. *Language* 57.3: 626–657.
- DeLancey, Scott. 1984. Etymological notes on Tibeto-Burman case particles. *Linguistics of the Tibeto-Burman Area* 8(1): 59–77.
- DeLancey. 2011. "Optional" "ergativity" in Tibeto-Burman languages. *Linguistics of the Tibeto-Burman Area* 34(2): 9–20.
- Dixon, R.M.W. 1979. Ergativity. *Language* 55: 59–138.
- Dixon, R.M.W. 1994. *Ergativity*. Cambridge: Cambridge University Press.
- Dolan R.J. & Dayan P. 2013. Goals and habits in the brain. *Neuron* 80(2):312–325.
- Dowty, D. 1991. Thematic proto-roles and argument selection. *Language* 67(3), 547–619.
- Durie, M. 1995. Towards an understanding of linguistic evolution and the notion 'X has a function Y'. In: W. Abraham et al. (eds), *Discourse Grammar and Typology*, pp. 275–308.
- Feldman, H. 1986. *A grammar of Awtuw*. Canberra: Research School of Pacific Studies.
- Gärdenfors, P. 2000. *Conceptual Spaces: The Geometry of Thought*. Cambridge, MA: MIT.
- Genetti 1986. The development of subordinators from postpositions in Bodic languages. In *Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society*, 387–400. Berkeley: Berkeley Linguistics Society.
- Genetti 1991. From postposition to subordinator in Newari. In Traugott, Elizabeth Closs & Bernd Heine (eds.) *Approaches to grammaticalization, Vol. II: focus on types of grammatical markers*, 227–255.
- Genetti, Carol. 2007. *A grammar of Dolakha Newar*. Berlin/New York: Mouton de Gruyter.
- Givón, Talmy. 1971. Historical syntax and synchronic morphology: An archeologist's field trip. CLS 7, Chicago: University of Chicago, Chicago Linguistics Society.
- Givón, Talmy. 1979. *On understanding grammar*. New York NY: Academic Press.
- Givón, Talmy. 1995. *Functionalism and grammar*. Amsterdam and Philadelphia: PUB.
- Givón, Talmy. 2009. *The genesis of syntactic complexity*. Amsterdam/Philadelphia: John Benjamins.
- Gong, T. and L. Shuai (2013), "Computer simulation as a scientific approach in evolutionary linguistics". *Language Sciences* 40: 12-23.
- Grice, H.P. 1975. Logic and conversation. In: P. Cole & J. L. Morgan (eds), *Syntax and Semantics*, Vol. 3, *Speech Acts*, 41–58. New York: Academic Press,
- Guiraud, P. 1968. The semic matrices of meaning. *Social Science Information*, 7(2), 131–139.
- Heine, Bernd, Ulrike Claudi & Friederike Hünemeyer. 1991. *Grammaticalization: a conceptual framework*. Chicago: University of Chicago Press.
- Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization*. New York: Cambridge University Press.
- Heine, Bernd & Tania Kuteva. 2007. *The genesis of grammar*. Oxford: Oxford University Press.
- Hopper, Paul J. & Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56.2: 252–299.
- Hyslop, Gwendolyn. 2010. Kurtöp case: The pragmatic ergative and beyond. *Linguistics of the Tibeto-Burman Area* 33(1): 1–40.
- Hurford, J. R. 1989. Biological evolution of the saussurean sign as a component of the language acquisition device. *Lingua* 77(2):187-222

- Jackendoff, R. 2002. *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.
- Jäger, H. et al. 2009, "What can mathematical, computational and robotic models tell us about the origins of syntax?" In: D. Bickerton and E. Szathmáry (eds), *Biological Foundations and Origin of Syntax* [Strüngmann Forum Reports, vol. 3.], 385-410. Cambridge, MA: MIT Press.
- Jurafsky, D. et al. 2001. Probabilistic relations between words: Evidence from reduction in lexical production. In: J. Bybee and P. Hopper (eds.), *Frequency and the emergence of linguistic structure*, 229–255. Amsterdam/Philadelphia. John Benjamins.
- Katz, J.J. & J.A. Fodor. 1963. The Structure of a Semantic Theory. *Language* 39(2): 170-210.
- Kempen, G. 2014. Prolegomena to a neurocomputational architecture for human grammatical encoding and decoding. *Neuroinform* 12: 111–142.
- Lestrade, S. 2010. *The space of case*. PhD thesis, Radboud University Nijmegen.
- Lestrade, S. 2015a. Simulating the development of bound person marking. In: H. Baayen, G. Jäger, M. Köllner, J. Wahle, T. Baayen-Oudshoorn (eds), *Proceedings of the 6th Conference on Quantitative Investigations in Theoretical Linguistics*. Tuebingen: University of Tuebingen.
- Lestrade, S. 2015b. A case of cultural evolution: The emergence of morphological case. *Linguistics in the Netherlands [AVT 32]*: 105-115.
- Lestrade, S. 2016a. The emergence of argument marking. In: S.G. Roberts, C. Cuskley, L. McCrohon, L. Barceló-Coblijn, O. Fehér, and T. Verhoef (eds) *The Evolution of Language: Proceedings of the 11th International Conference (EVOLANG11)*.
- Lestrade, S. 2016b. The place of Place. In: L. Ströbel (ed.), *Sensory Motor Concepts in Language & Cognition*, 131–146. Düsseldorf: Düsseldorf University Press.
- Lestrade, S. *to appear*. The emergence of differential case marking.
- Lestrade, S. & H. de Hoop. 2016. On case and tense. *The Linguistic Review* 33(3).
- Lestrade, S., G. van Bergen, & P. de Swart. 2016. On the origin of constraints. In Legendre et al. (eds.) *Optimality-Theoretic Syntax, Semantics, Pragmatics*, 179-199. Oxford: Oxford University Press.
- Levelt, W.J.M. 1983. Monitoring and self-repair in speech. *Cognition* 14: 41–104.
- LaPolla, Randy J. 1992. Anti-ergative marking in Tibeto-Burman. *Linguistics of the Tibeto-Burman Area* 15(1): 1–9.
- LaPolla, Randy J. 1994. Parallel grammaticalization in Tibeto-Burman. *Linguistics of the Tibeto-Burman Area* 16(1): 61–80.
- LaPolla, Randy J. 1995. ‘Ergative’ marking in Tibeto-Burman. In Nishi, Yoshio, Matisoff, James A. & Nagano, Yasuhiko (eds.) *New horizons in Tibeto-Burman morpho-syntax (Senri Ethnological Studies 41)*, 189–228. Osaka: National Museum of Ethnology.
- LaPolla, Randy J. 2001. The role of migration and language contact in the development of the Sino-Tibetan language family. In Dixon, R.M.W. & A.Y. Aikhenvald (eds.) *Areal diffusion and genetic inheritance: case studies in language change*, 225–254. Oxford and New York: Oxford University Press.
- LaPolla, Randy J. 2003. *A grammar of Qiang, with annotated texts and glossary*. Berlin/New York: Mouton de Gruyter.
- LaPolla, Randy J. 2004. On nominal relational morphology in Tibeto-Burman. In Ying-chin Lin, Fang-min Hsu, Chun-chih Lee, Jackson T.-S. Sun, Hsiu-fang Yang and Dah-an Ho (eds.) *Studies of Sino-Tibetan languages: Papers in Honor of Professor Hwang-chenng Gong on His Seventieth Birthday*, 43–73. Taipei: Institute of Linguistics, Academia Sinica.
- Lidz, Liberty A. 2011. Agentive marking in Yongning Na (Mosuo) *Linguistics of the Tibeto-Burman Area* 34(2): 49–72.

- Malchukov, A. 2006. Transitivity parameters and transitivity alternations: Constraining co-variation. In P. de Swart et al. (eds.), *Case, valency and transitivity*, pp. 329-358. Amsterdam/Philadelphia: Benjamins.
- Matisoff, James A. 1991. Sino-Tibetan linguistics: Present state and future prospects. *Annual Review of Anthropology* 20: 469–504.
- Matisoff, James A. 2003. *Handbook of Proto-Tibeto-Burman: System and philosophy of Sino-Tibetan reconstruction*. Berkeley: University of California Press.
- McGregor, William B. 2009. Typology of ergativity. *Language and Linguistics Compass* 3(1): 480–508.
- Miller, Roy Andrew. 1996. Review of David Bennett, et. al, *Subject, voice and ergativity: Selected essays*. *Journal of the American Oriental Society* 116(4): 755-757.
- Morey, Stephen. 2012. The Singpho agentive – functions and meanings. *Linguistics of the Tibeto-Burman Area* 35.1: 1–14.
- Nettle, D. 1999. *Linguistic diversity*. New York: Oxford University Press.
- Noonan, Michael. 2009. Patterns of development, patterns of syncretism of relational morphology in the Bodic languages. In Barðdal, Jóhanna and Shobhana L. Chelliah (eds.), 261–282.
- Peterson, David A. 2011. Core participant marking in Khumi. *Linguistics of the Tibeto-Burman Area* 34(2): 73–100.
- Reinöhl, Uta. 2016. *Grammaticalization and the Rise of Configurationality in Indo-Aryan*. Oxford: Oxford University Press.
- Sapir, Edward. 1921. *Language: An introduction to the study of speech*. New York: Harcourt, Brace & World.
- Siewierska, Anna & Dik Bakker. 2009. Case and alternative strategies: Word order and agreement marking. In Malchukov, Andrej & Andrew Spencer (eds.) *The Oxford handbook of case*, 290–303. Oxford: Oxford University Press.
- Silverstein, Michael. 1976. Hierarchy of features and ergativity. In Dixon, R.M.W. (ed.) *Grammatical categories in Australian languages*, 112–171. Canberra: Australian Institute of Aboriginal Studies.
- Sinnemäki, K. 2014. A typological perspective on differential object marking. *Linguistics*, 52(2): 281-314.
- Smith, A. D. 2014. Models of language evolution and change. *WIREs Cogn Sci* 5: 281-293. (doi:10.1002/wcs.1285)
- Steels, L. 1997. Constructing and sharing perceptual distinctions. *Machine Learning: ECML-97*, 4–13. Springer.
- Steels, L. 2003. Language re-entrance and the inner voice. *Journal of Consciousness Studies* 10(4-5): 173–185.
- de Swart, P. 2007. *Cross-linguistic variation in object marking*. Ph.D. thesis, Radboud University Nijmegen.
- Teo, Amos. 2012. Sumi agentive and topic markers: *no* and *ye*. *Linguistics of the Tibeto-Burman Area* 35(1): 49–74.
- Tomasello, M. 2003. *Constructing a language: a usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- Touradre, Nicolas. 1991. Rhetorical use of the Tibetan ergative. *Linguistics of the Tibeto-Burman Area* 14(1): 93–107.
- Van Valin, Robert D. 1999. Generalized semantic roles and the syntax-semantics interface. In: F. Corblin, C. Dobrovie-Sorin & J.-M. Marandin (eds.), *Empirical issues in formal syntax and semantics* 2, 373–389. The Hague: Thesus.

- Vogels, J., E. Krahmer, & A. Maes. 2015. How Cognitive Load Influences Speakers' Choice of Referring Expression. *Cognitive Science* 39 (6): 1396–1418.
- Vosse, Th. & G. Kempen 2000. Syntactic structure assembly in human parsing: A computational model based on competitive inhibition and a lexicalist grammar. *Cognition* 75: 105-143.
- Wierzbicka, A. 1996. *Semantics: Primes and universals*. Oxford University Press.
- Willis, Christina. 2011. Optional case marking in Darma (Tibeto-Burman). *Linguistics of the Tibeto-Burman Area* 34(2): 101–131.
- Zeevat, H. 2007. Simulating recruitment in evolution. In: G. Bouma, I. Kraemer & J. Zwarts (eds.), *Cognitive Foundations of Interpretation*, 175–194. Amsterdam: Royal Netherlands Academy of Arts and Sciences.