

Radical analyticity and radical pro-drop scenarios of diachronic change in East and mainland Southeast Asia, West Africa and Pidgins and Creoles

Walter Bisang

University of Mainz & Zhejiang University

The paucity or absence of inflectional morphology (radical analyticity) and the omission of verbal arguments with no concomitant agreement (radical pro-drop) are well-known characteristics of East and mainland Southeast Asian languages (EMSEA). Both of them have a special status in typology and linguistic theory. Radical analyticity is known under the term of ‘morphological isolation’ and has recently been described as ‘diachronically anomalous’ (McWhorter 2016), while radical pro-drop is a theoretical challenge since Rizzi (1986). The present paper offers an alternative view on these characteristics based on data from EMSEA languages, radically analytic West African languages and pidgins and creoles. It develops diachronic evolutionary scenarios combining the specific properties of languages in their diachronic and geographic situations with two different notions of complexity (hidden vs. overt complexity) and factors which tend to block the development of inflectional morphological paradigms. From such a perspective, radical analyticity and radical pro-drop are by no means extraordinary. Given the enormous size of the task, the paper is a thought experiment based on available data and discussions on the above languages for stimulating further research.

Keywords: diachronic linguistics, frequency, grammaticalization, linguistic complexity, linguistic contact, morphological paradigms, radical pro-drop

1. Setting the stage: The basic idea of the paper

East and mainland Southeast Asian languages are generally known for their radically analytic morphology and for being radical pro-drop. Radical analyticity roughly corresponds to the more traditional notion of isolating morphology and

is defined by McWhorter (2016, p. 50) as the ‘*paucity or absence of inflectional (as opposed to derivational) marking indicated by affixation, tone, or vowel changes in quality or length*’. Grammars of languages with radical pro-drop allow the omission of verbal arguments without the presence of concomitant agreement morphology on the verb (Rizzi 1986; Ackema et al. 2006; Neeleman and Szendrői 2007 and many others). With this definition, radical pro-drop differs from “mild” pro-drop as it is found in Italian or Spanish, in which the person/number features of the subject are still expressed on the verb. Languages with radical pro-drop are a challenge for linguistic theory because they show that there is no necessary correlation between morphological richness and the licencing of pro-drop (also cf. Section 5.5 and footnote 12 in Section 4 on pidgins and creoles).

Radical analyticity is not limited to the languages of East and mainland Southeast Asia (EMSEA). Similar properties are found in some West African languages of the Niger-Congo family (Gbe languages, Yoruba, Nupe) and in some Austronesian languages spoken on the islands of Flores and Timor. The present paper will be focused on the comparison of EMSEA languages with isolating West African languages and with Pidgin and Creole languages. It starts out from the observation that the highly analytic languages of West Africa show no radical pro-drop, while the highly analytic languages of East and mainland Southeast Asia do so. As it will be shown, this difference is related to the question of whether the ancestor languages involved had inflectional morphological paradigms for the features of person, number and, in some languages, gender or noun class. East and mainland Southeast Asian languages did not have such paradigms in their ancestor languages and are overwhelmingly radical pro-drop, while the analytic languages of West Africa had them and show no radical pro-drop. The effects of such paradigms even show up after extreme situations of contact as they are found in pidgins and creoles. As will be shown, the pro-drop properties of pidgin and creole languages remarkably covary with the presence/absence of inflectional morphological paradigms in the contact languages involved. Moreover, the number of pidgin and creole languages with radical pro-drop is relatively small (for details, cf. Section 4). Finally, it may be added that other well-known radical pro-drop languages did not have inflectional paradigms with person/number marking. This is the case for the radically analytic languages of the Austronesian family¹ as well as for non-radically analytic languages like Japanese and Korean.²

1. Cf. Ross (2002) and Adelaar (2005). As one can see from Ross (2002: 52–53), Austronesian does have inflectional morphological paradigms but these paradigms typically do not include person/number marking.

2. These languages have inflectional morphology for various categories on the verb (tense, aspect, modality, evidentiality, negation, etc.) but no person/number marking.

The findings on the presence or absence of radical pro-drop in its correlation to the absence of inflectional morphological paradigms in EMSEA languages in analytic West African languages and in pidgin and creole languages will be modelled in diachronic evolutionary scenarios based on two types of complexity and two types of maturation (Bisang 2009, 2014b, 2015a) in combination with a few other factors which tend to block the emergence of incipient inflectional morphological paradigms in the radically analytic EMSEA languages.

From the perspective of radical pro-drop, the grammars of EMSEA languages have the potential to produce more simple surface structures than the majority of pidgins and creoles inasmuch as they can leave the information on person/number/(gender/noun class) features of verbal arguments to pragmatic inference. This shows that explanations in terms of intensive contact and the emergence of languages with minimal grammatical complexity is not necessarily sufficient. For that reason, the scenario presented in this paper needs a richer concept of complexity that integrates pragmatic inference and thus goes beyond the general discussion in linguistic typology (Kusters 2003; Dahl 2004; Sinnemäki 2011 and others). Such a concept is crucial for modeling the presence/absence of radical pro-drop in this paper. Its basic idea is that linguistic complexity is the result of the competing motivations of explicitness (overt encoding of grammatical categories) and economy (pragmatic inference of grammatical categories existing in a given language). This competition creates two types of complexity, explicitness-based overt complexity which manifests itself in linguistic form as it is discussed in typology and economy-based hidden complexity which leaves information which is part of the grammatical inventory of a language to pragmatic inference (Bisang 2009, 2014b, 2015a).

If a language has an inflectional morphological paradigm expressing person/number/ (gender/noun class) agreement on the verb, information on these features will generally be available for subject nouns (and maybe object nouns, if there is obligatory object agreement). Thus, one may argue that speakers are not used to this kind of pragmatic inference in languages which show overt complexity with regard to these features, while this is no problem to speakers who grow up in a linguistic environment of hidden complexity in which the frequency of these features is low because there are no inflectional morphological paradigms requiring their presence. This difference supports structures with no radical pro-drop in radically analytic West African languages and enhances radical pro-drop in EMSEA languages. In the case of pidgin and creole languages, both effects can be observed, depending on the properties of the input languages (also cf. Section 5.3).

The scenarios developed in this paper are understood as the result of a thought experiment based on representative data from the languages involved. Even though the paper does not offer a systematic and rigorous statistical analysis its aim is to contribute to the understanding of why languages show the properties

they have at a given time in a given area (cf. Bickel 2007). It particularly hopes to stimulate further research for improving the modeling of these questions (and maybe falsifying the scenarios developed here) by the production of (i) more detailed descriptions of relevant individual languages, (ii) more rigorous statistical analyses and (iii) more experimental studies on complexity and cognitive costs. For that purpose, it combines issues discussed in previous publications by the present author on linguistic complexity (Bisang 2009, 2014b, 2015a) and on grammaticalization in EMSEA languages (Bisang 2008, 2015b) with ideas on morphological paradigms and radical-pro-drop (Bisang 2014a). In addition to these topics, it will compare data on radical pro-drop from APiCS (*Atlas of Pidgin and Creole Language Structures*, feature 62; Michaelis et al. 2013) with the presence/absence of morphological paradigms in the input languages. Given this situation, some overlap with data from previous publications in this paper will be inevitable.

The definition of inflectional morphological paradigms adopted in the present paper follows Baerman and Corbett (2010), who understand morphological paradigms in terms of the three components of a form paradigm, a content paradigm and the mapping of these two paradigms onto each other. The interaction of these components creates patterns consisting of categories and subcategories (values) with individual cells that are filled by individual markers (e.g. the two categories of number and person with their respective values of singular and plural vs. 1st, 2nd and 3rd person).

The scenarios outlined so far will be discussed as follows: Radical pro-drop in EMSEA languages and its absence in radically analytic West African languages are presented in Section 2. The discussion of these two groups of languages is continued in Section 3 on the diachronic strength of inflectional morphological paradigms. Section 4 is on pidgin and creole languages and their pro-drop properties in the light of the presence/absence of inflectional morphological paradigms in the input languages. The scenarios for modeling these situations in the context of complexity/maturation and some factors that operate against incipient inflectional morphology (without being able to fully block it) will be developed in Section 5. Finally, Section 6 offers a short conclusion concerning the typological status of radical analyticity and radical pro-drop.

2. Radical pro-drop in East and mainland Southeast Asia and in West Africa

The area of East and mainland Southeast Asia extends over Peninsular Malaysia, the northwestern part of Myanmar, Thailand, Cambodia, Laos and Vietnam in the South and Sinitic in the north. The languages spoken in that area belong to five families – Sinitic, Mon-Khmer (a branch of Austroasiatic), Tai (a branch of

Tai-Kadai), Hmong-Mien (or Miao-Yao in Chinese) and Chamic (a branch of Austronesian spoken in Vietnam). In the vast majority of these languages, arguments do not agree with the verb and they may be omitted if they can be inferred from context. This is illustrated by the following two examples from Mandarin Chinese (1) and Thai (2):

(1) Mandarin Chinese:

你 买了 苹果 吗? – 买了。
Nǐ mǎi-le píngguǒ ma? – ø mǎi-le ø_i.
 you buy-PFV apple Q buy-PFV
 ‘Did you buy apples?’ – ‘[I] bought [some].’

(2) Thai:

Khun rúucàk aacaan_i máy? ø rúucàk ø_i.
 you know teacher Q know
 ‘Do you know the teacher_i?’ – ‘[I] know [him_i].’

Radical pro-drop is not a recent phenomenon in Chinese and most likely also in the languages from the other families spoken in East and mainland Southeast Asia. In the following example from classical Chinese (5th – 3rd centuries BC), the two protagonists, i.e. the king (*wáng* ‘king’) and the philosopher Mencius (*Mèng-zǐ* [Meng master] ‘Master Meng’), are introduced by overt nouns. After that, the arguments are left unexpressed in the discourse situation reported by the text and the distinction between 1st and 2nd person must be inferred from context:

(3) Classical Chinese (Mencius 2.B.10; Bisang 2013: 8–9):

王就見孟子曰，前日願見而不可得，得侍同朝甚喜。今又棄寡人而歸，不識可以繼此而得見乎。
 (孟子，公孫丑章句下)

Wáng_i jiù jiàn Mèng-zǐ_i ø_i yuē ‘qiánrì ø yuàn ø jiàn ér ø
 king get.close-to see Mencius-master say previously want see but
bù kě dé, ø dé ø shì tóngcháo ø shén xǐ. Jīn yòu ø qì
 NEG can succeed succeed serve court delighted now again abandon
guǎrén ér ø guī, ø bù shí ø kěyǐ jìcǐ ér ø dé ø jiàn ø hū?’
 I and go.home NEG know can hereafter and succeed see Q
 ‘The king_i went to see Mencius and ø_i said: “Previously, [I] wanted to see [you] but [I] did not succeed and when [I] succeeded in serving the same court [i.e. working together] [I] was highly delighted. Now, [you] abandon me again and [you] go home. [I] don’t know if it will be possible hereafter that [I] see [you].”’

Radical pro-drop has its effects in syntax. One of them shows up in relative-clause formation. Chinese relative clauses always end with the attributive marker *de* and precede their head nouns. The syntactic function and the semantic role of the head noun within the relative clause can be inferred from the fact that arguments of intransitive and monotransitive verbs must be zero-marked. Since pro-drop is generally possible in relative clauses, constructions consisting merely of a transitive verb are ambiguous (Bisang 2013). For that reason, the following relative clause, which consists only of the verb *zhǎo* ‘look for’ plus the relative marker *de*, can be interpreted either in terms of object coreference (4a) or subject coreference (4b):

(4) Mandarin Chinese:

找 的 人 还 没有 回来。

[*zhǎo de*] *rén* *hái méiyǒu huí-lái.*

look.for REL man/people still NEG:PST return-come

a. Object coreference:

‘The people [(we) were looking for] haven’t come back yet.’

b. Subject coreference:

‘The people [who looked (for us)] haven’t come back yet.’

In a similar way, a simple relative construction in Thai with the relative marker *thîi* followed by a monotransitive verb may stand for subject coreference (5a) or object coreference (5b):

(5) Thai:

a. Subject coreference:

khon [thîi kin]

man REL eat

‘the man who eats’

b. Object coreference:

mamûang [thîi kin]

mango REL eat

‘the mango which is eaten/which X eats’

In contrast to the EMSEA languages described so far, isolating languages of the West African type are not radical pro-drop. As can be seen from the following example from Yoruba (Niger-Congo: Atlantic-Congo: Volta-Congo: Benue-Congo: Defoid), arguments must be expressed overtly in independent declarative clauses irrespective of whether they are subjects or objects. For that reason, Examples (6a–d) with overt subject and object marking are grammatical, while Examples (6e–g) are ungrammatical because the subject, the object or both are not expressed overtly:

(6) Yoruba: On pro-drop with the verb *rà* ‘buy’:³

- a. *Ayọ̀`´` ràşọ.*
 Ayọ̀`´` rà aşọ
 Ayo HTS buy clothes
 ‘Ayo bought clothes.’
- b. *ó ràşọ.*
 ó rà aşọ.
 he buy clothes
 ‘He buys clothes.’
- c. *Ayọ̀`´` ràá.*
 Ayọ̀`´` rà-á.
 Ayo HTS buy-3.OBJ
 ‘Ayo buys it.’
- d. *ó ràá.*
 ó rà-á.
 he buy-3.OBJ
 ‘He buys it.’
- e. **ràşọ.*
 ø rà aşọ
 buy clothes
- f. **Ayọ̀`´` rà.*
 Ayọ̀`´` rà ø.
 Ayo HTS buy
- g. **rà.*
 ø rà ø.
 buy

In Fongbe (Niger-Congo: Atlantic-Congo: Volta-Congo: Kwa), the overt expression of arguments in simple independent clauses is obligatory (7). Only a few verbs like *cí* ‘seem, look like’ (8) and *d̥ɔ̀* ‘seem, resemble’ have optional subjects (Lefebvre and Brousseau 2002: 246, 276–277):⁴

3. HTS stands for ‘High Tone Syllable’. This tonal marker is a floating tone which occurs at the final syllable of subject NPs under certain semantic conditions (Bisang and Sonaiya (1999) and Déchaine (1993)).

4. Objects are also obligatory in Fongbe. For further information, cf. Lefebvre and Brousseau (2002: 247–248).

- (7) Fongbe (Lefebvre and Brousseau 2002: 276):

(É) cí dḡ Kòkú jè àzḡn.

3.SG seem COMP Koku be.sick

'It seems that Koku is sick.'

- (8) Lefebvre and Brousseau (2002: 277):

É vè-wú nú Kòkú ní yì.

3.SG be-difficult COMP Koku SUBORD leave

'It is difficult that Koku leaves.'

The syntactic effects of relative-clause formation are equally absent in Yoruba and Fongbe. In Yoruba, relative clauses are introduced by the relative marker *tí*. With subject coreference, the subject position within the relative clause must be filled by an obligatory pronoun which is coindexed with the head noun (9). In the case of object coreference, the object position of the verb in the relative clause must be empty (10):

- (9) Yoruba: Subject coreference (Bisang 2013: 70):

obínrin [t'ó maa ràá].

obínrin tí ó maa rà-á

woman REL 3.SG.SBJ TA buy-3.OBJ

'the woman who bought it'

- (10) Yoruba: Object coreference (Bisang 2013: 70):

ìyá [tí mo rà ø lánàá]

yam REL 1.SG buy yesterday

'the yam I bought yesterday'

In Yoruba, the grammar systematically blocks structures that leave coreference interpretation in relative clauses to pragmatic inference. Fongbe grammar does the same with different means. The head noun of Fongbe relative clauses 'is linked to a position within the relative clause through the lexical nominal operator, *dḡ* (Lefebvre and Brousseau 2002: 161). This operator attracts lexical material. In the case of relative clauses, it attracts resumptive pronouns by moving them from their initial position in the relative clause to the position immediately after *dḡ*. This process generates surface structures which clearly differ for each type of coreference. In the case of subject coreference, the nominal operator *dḡ* is followed by the high-tone 3rd singular subject pronoun *é* (11). In the case of object coreference, we find the low-tone 3rd singular object pronoun *è* (12) in the same position:

- (11) Fongbe: Subject coreference (Lefebvre and Brousseau 2002: 161):

súnû [dɛ́-é wá] ɔ́
 man OP-3.SG.SBJ come DEF
 'the man who came'

- (12) Object coreference (Lefebvre and Brousseau 2002: 161):

àsón_i [dɛ́-è_i Kòkú dũ ɔ_i] ɔ́.
 crab OP-3.SG.OBJ Koku eat DEF
 'the crab that Koku ate'

The absence of radical pro-drop in Yoruba and Fongbe reflects a general property of African languages. As was pointed out by Creissels (2005), we find the following three patterns of argument expression in Africa:

- i. Complementary distribution of pronominal marking on the verb vs. overt noun phrase.
- ii. Obligatory pronominal marking on the verb but optional expression of corresponding noun phrase.
- iii. Obligatory pronominal marking on the verb and obligatory corresponding noun phrase.

Based on this typology, Creissels (2005) presents the following conclusion:

In virtually all African languages, arguments encoded as subjects, if not represented by a noun phrase or free pronoun in subject function, must at least be represented by a subject marker. In the case of arguments encoded as objects, most of the time, their total omission triggers an indefinite rather than anaphoric interpretation. (Creissels 2005)

The above data from East and mainland Southeast Asia and from West Africa clearly show that the presence or absence of radical pro-drop (zero reference) is not related to morphological complexity in a simple way. Kibrik (2011: 113) points out this fact as follows:

However, it is not so easy to prove a relationship between zero reference and morphological scarcity. First, West Africa also is an area of isolating, analytic, and non-marking language, but zero reference is not typical there at all. Second, there are languages that are morphologically quite developed but still have zero reference. Both Yidiny and Japanese ... exemplify this combination.

As is argued in this paper, the correlation between inflectional morphology and radical pro-drop needs a look at the history of the languages involved.

3. The diachronic strength of inflectional morphological paradigms

3.1 Absence of inflectional morphological paradigms in EMSEA languages

Due to their long and intricately complex history of contact EMSEA languages share a lot of typological properties irrespective of their genetic affiliation. However, some of these properties seem to be independently existing in the families involved or may have been developed independently in each family and were then further enhanced and corroborated by contact with speakers whose languages already showed similar tendencies. One of them is the absence of verbal inflectional morphological paradigms marking the features of person, number (and gender/noun class). As will be argued in this subsection, inflectional morphology on the verb was completely absent or not very much developed in any of the language families involved in the last two millennia. This will be briefly illustrated by a look at Sinitic and Mon-Khmer. In the case of Tai-Kadai and Hmong-Mien, there is no evidence for the existence of inflectional morphological paradigms. In particular, there is no evidence for marking person, number (and gender/noun class) in historical-comparative work.⁵ In Proto-Austronesian, the features of person and number were not involved with verbal morphology even though there was an elaborate paradigm of diatheses and tense-aspect marking (cf. footnote 1).

The existence of bound morphology in Old Chinese (11th – 6th centuries BC) is uncontroversial but that morphology did not express inflectional categories like person, number or tense-aspect and it was not organized in morphological paradigms (cf. Downer 1959; Sagart 1999; Baxter and Sagart 2014). Moreover, most of the markers were multifunctional and their concrete function was determined lexically, i.e. a specific lexical item was associated with one specific function of the multifunctional marker. The following list (Bisang 2014a: 41) is a summary of the most important affixes in Old Chinese and their functions as they are presented in Sagart (1999):

- (13) *s-: Denominal verbs, causative, directive (acts/states directed toward external condition or other persons), maybe inchoatives;
- *m-: Controlled actions by volitional actors: (a) volitionality, (b) agentive nouns, (c) small animals;
- *k-: Actions/objects “that are well-delimited in time and space, and hence usually concrete and countable” (Sagart 1999: 107). With verbs: transient/attempted action, repeated action, continued action;

5. Cf. Li (1977), Edmondson and Solnit (1988), Luo (1997) on Tai and Ratliff (2010) on Hmong-Mien.

- **t*–: (a) stative verbs, (b) involuntary physiological actions, (c) some other intransitives, (d) some nouns;
- **r*–: General idea of intensification: (a) iteratives, (b) action taking place in two or more locations, (c) nouns denoting double/multiple object (e.g. eyeballs), (d) intense quality;
- **s*–: (a) denominal nouns, (b) exodirectional/endodirectional verbs (e.g. hear vs. ask).

Since none of these affixes seemed to be used productively and their concrete meaning was lexically determined for individual roots, Sinitic morphology as it can be reconstructed was not inflectional and it disappeared completely after classical Chinese, giving way to a number of new constructions that did not exist in classical Chinese (e.g. resultative constructions, directional verbs, coverbs/prepositions; Xu 2006; Bisang 2010).

The Austroasiatic family is characterized by a remarkable typological split between the Munda languages spoken across portions of central and eastern India and the other subfamilies spoken outside of that area in East and mainland Southeast Asia (Aslian, Khmeric, Vietic, Pearic, Katuic, Monic, Bahnaric, Khmuic, Palaungic, Mangic), the eastern part of Meghalaya State in India (Khasian) and the Nicobarese islands (Nicobarese). In Donegan and Stampe's (2004) description, Munda languages are synthetic, head-last (OV), agglutinative and have prefixes, suffixes and infixes, while the languages of the Mon-Khmer branch are analytic, head-first (VO), fusional and have prefixes and suffixes or are isolating. As for inflectional morphological paradigms, Munda languages show subject and object cross-reference on the verb as well as tense-aspect marking, transitive/intransitive marking and finiteness marking (Zide and Anderson 2001; Anderson 2007; Anderson 2016):

- (14) Juang (Matson 1964: 35; quoted from Anderson 2016: 390):

mɛ-dʒɔ-ki-n

2:SBJ-see-PRS.TR-1:OBJ

'you see me.'

- (15) Mundari (Cook 1965: 228; quoted from Anderson 2016: 391):⁶

hola Ranchi-te=ŋ sen-ke-n-a.

yesterday Ranchi-ALL-1 go-ASP-INTR-FIN

'Yesterday I went to Ranchi.'

6. Notice that the subject marker is preferably attached to the preceding word in Mundari and the other languages belonging to the Kherwarian subfamily of Munda.

In contrast, there are no verbal inflectional paradigms in the Mon-Khmer languages spoken in East and mainland Southeast Asia with the exception of Aslian languages (cf. subSection 5.4), where cross-reference morphology is a rather recent phenomenon. Outside of East and mainland Southeast Asia, Khasi spoken in north-eastern India has obligatory subject pronouns but no inflectional morphology. In this language, “[v]erbs are always preceded by a pronoun, either directly representing the subject or resumptively indexing the subject where it is already represented by a full noun, in which case it can be analyzed as obligatory subject person/number agreement” (Nagaraja 2016: 1171):

(16) Khasi (Nagaraja 2016: 1172):⁷

ki=briew ki la jadie ja ka=du:t.

PL=person 3.PL PST sell ACC F=milk

‘The men sold milk.’

Given the enormous cross-linguistic variation between languages with isolating morphology like Vietnamese and the Munda languages with their highly developed verbal inflectional morphology, it comes as no surprise that different hypotheses on the morphology of proto-Austroasiatic have been suggested in the literature. Sidwell (2009: 62–64) discusses three currently available scenarios:

Scenario (i): Western origin in northern India or in the vicinity of the Bay of Bengal. (Pinnow 1960; Zide and Anderson 2001; Van Driem 2001)

Scenario (ii): Northern origin in central or southern China.

(Donegan and Stampe 2004)

Scenario (iii): Central origin within Southeast Asia. (Sidwell 2008)

In scenario (i), proto-Austroasiatic verb morphology was similar to Munda morphology. Zide and Anderson (2001: 517) reconstruct the following verbal template for proto-Austroasiatic:

(17) SBJ-NEG-VERB.STEM-TRANS/TENSE-OBJ⁸

In scenario (ii), proto-Austroasiatic did not have inflectional morphology. Like in modern Khmer, it had a number of derivational prefixes and infixes mainly for marking nominalization (e.g. the infix *-b-* in *rə-b-am* ‘dance, n.’ from *rəəm* ‘to dance’ or *-m-* in *s-m-ò:m* ‘beggar’ from *sò:m* ‘ask for’), causativization/

7. In Khasi, case is obligatorily expressed by markers that precede the noun (e.g. *ja* for ACC). Moreover, the language has gender in the singular (masculine vs. feminine).

8. The verb stem itself maximally has the following structure: CAUS/RECIP-(REDUPLICATION)-ROOT-PASSIVE/INTRANSITIVE. The categories of transitivity and tense in (18) are expressed by fusional morphemes.

transitivization (e.g. *p-* in *p-dac* ‘break, separate, tr.’ from *dac* ‘break, be torn apart, intr.’ or *prə-* in *prə-kaət* ‘cause, bring about’ from *kaət* ‘be born, arise, happen’) and for specifying the reference of the simplex root (the infix *-vmn-/-vN-* in *l-ùm-ɔːŋ* ‘fine dust, pollen’ from *lɔːŋ* ‘dust’ or *p-* in *p-haəm* ‘swollen, pregnant’ from *haəm* ‘swollen’) (Jenner and Pou 1982; Bisang 1992, 2014a: 44–45).

If Zide and Anderson’s (2001) scenario (i) is right proto-Austroasiatic had rather complex inflectional paradigms that got lost through time when its speakers were settled in East and mainland Southeast Asia. If scenario (ii) of Donegan and Stampe (2004) is right the Austroasiatic languages spoken in East and mainland Southeast Asia directly reflect proto-Austroasiatic and the more complex morphological patterns of Munda developed later, most likely in contact with Indo-Aryan and Dravidian. For the EMSEA languages in question, this means that they had no or maybe almost no inflectional morphological paradigms when they got in contact with EMSEA languages from other families.

In scenario (iii), the number of affixes was smaller than in Khmer but the functions they expressed were similar and there was clearly no inflectional morphology expressing person, number or gender/noun class. This approach was further elaborated by a computational phylogenetic project (Sidwell and Blench 2011). As a result of his research that combines lexical, lexicostatistical, computational phylogenetic and phonological studies, Sidwell (2016) argues for a much flatter tree structure that is much less deeply nested than in many former approaches as for instance the one by Diffloth (2005) and does not follow the typologically motivated dichotomy between Munda and Mon-Khmer. Sidwell’s (2016) classification of Austroasiatic, he calls it a “provisional” classification, has eleven primary nodes:

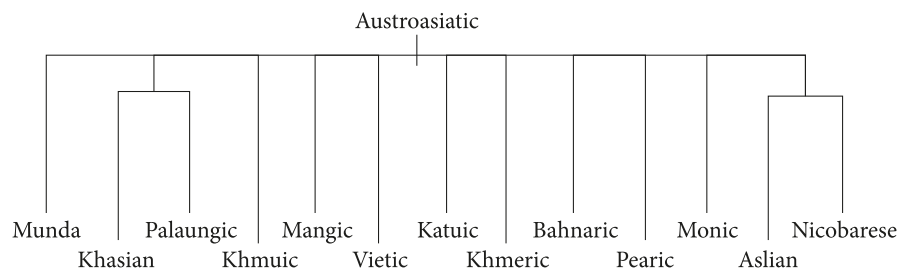


Figure 1. Classification of Austroasiatic (Sidwell 2016: 179)

In this scenario, proto-Austroasiatic is very old and remained very stable for millennia. Its speakers were situated ‘along the Mekong and surrounding uplands, perhaps as far afield as the Red River valley’ (Sidwell 2016: 180). Given the flatness of the tree structure in Figure 1, it is rather likely that there was ‘a phase of rapid

differentiation at some point in the history of the phylum' (Sidwell 2016:179). In the course of that rather short period of differentiation, the speakers of Austroasiatic took their relatively limited morphological inventory with its absence of inflectional properties to the areas where they are living up till now. The typological differences emerged through time by contact with speakers of the surrounding languages (e.g. Indo-Aryan and Dravidian in the case of Munda) or by language-internal changes. Thus, scenario (iii) further supports the assumption that the proto-Austroasiatic languages spoken in East and mainland Southeast Asia did not have inflectional morphological paradigms in their history.

3.2 The presence of morphological paradigms in Niger-Congo

In the Niger-Congo family with its approximately 1.400 languages, there is considerable variation between analytic structures with no morphology as illustrated for Yoruba (examples, (6), (9) and (10)) and Fongbe (Examples (7), (8), (11) and (12)) and synthetic morphology with sequences of affixes of considerable length as in (18) from Makwe (Bantu: Swahili group, spoken in Mozambique). As is characteristic of Bantu morpheme order, the preradical position in Makwe starts with a subject marker, followed by a tense/aspect marker and an object marker. In the post-radical position, we find derivational suffixes marking categories like causative, applicative and passive:

(18) Makwe (Devos 2008: 239):

mu-nandi u-nku-kom-ish-il-iw-a wa-ana.

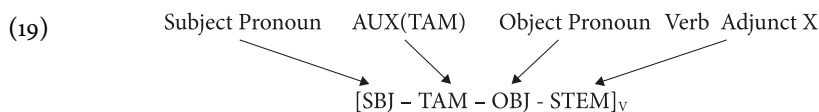
III-branch III-PROG-hit-CAUS-APPL-PAS-FV II-child

'The branch is used to make somebody hit the children.'

Given this impressive variation between analytic and synthetic structures, the question of how the morphology of proto-Niger-Congo (PNC) looked like has led to divergent hypotheses. Hyman (2004, 2010) argues that verb morphology was synthetic in proto-Niger-Congo, while Güldeman (2007, 2010) suggests that proto-Bantu (PB) and proto-Niger-Congo was basically analytic. Since this paper is about radical pro-drop and the expression of the features of person, number and noun class, the remainder of this subsection will focus on the preradical positions in which subject and object cross-reference is expressed.

Güldemann (2007, 2010) is based on the Macro-Sudan belt (MSB), a large area with languages from the Niger-Congo family as well as from the Sudanic family of Nilo-Saharan and the Chadic subfamily of Afroasiatic. The languages of that area were characterized by analytic structures. This also applied to proto-Bantu as long as its speakers were living in the Macro-Sudan belt. When they moved out of it their languages became more synthetic through processes of grammaticalization. In this

scenario, the morphological pattern of [SBJ-TAM-OBJ-STEM]_V as it was illustrated in (18) derived out of a former syntactic/analytic structure as illustrated in (19), consisting of a subject pronoun, an auxiliary verb expressing tense-aspect-mood and an object pronoun:



Hyman's (2010) argument against Güldemann's (2010) scenario is that the proto-Bantu dispersion out of the Macro-Sudan belt must have taken place some 4.000 to 5.000 years ago, while the time-depth for proto-Niger-Congo (PNC) must be about 10.000 to 12.000 years. Hyman (2010) agrees that the Bantu verb morphology must have developed out of more independent morphemes but this does not say much about the morphological properties of proto-Niger-Congo that was spoken several millennia before the proto-Bantu (PB) dispersion. In Hyman's (2010) analysis, the diachronic changes in the last 10.000 to 12.000 years that shaped the current typological variation within Niger-Congo are characterized by cyclical oscillations of building up and reducing inflectional morphology:

There has been plenty of time for PB (and even more time for PNC) to cycle back and forth, grammaticalizing full words as inflectional proclitics and prefixes, losing them, and creating them once more. The issue thus is not one of “diachronic typology”, to use G’s [Güldemann’s, W.B.] term, but of dating. This may not be easy to do, given the cyclicity. We all seem to agree that PB came from an earlier analytic stage – the question, however, is whether Basaa, Tunen etc. represent that unchanged stage, or whether they are completing the cycle: analytic > agglutinative > analytic. I maintain that the latter is the case. (Hyman 2010: 24)⁹

No matter which of the two scenarios is correct, what can be concluded from both of them is that the expression of number, person and noun class by a morphological paradigm of verbal cross-reference marking is a recurrent and very time-stable property of Niger-Congo languages throughout the whole history of that family.

9. Basaá and Tunen are two Bantu languages spoken in Cameroun that belong to Zone A (Basaá A43, Tunen A44). They are both characterized by comparatively high analyticity.

4. Radical pro-drop in creoles/pidgins and the relevance of inflectional morphological paradigms

The observation that there are similarities in the pro-drop properties of creoles and their contact languages can be seen from comparing the Fongbe data given in (7) and (8) with the data on the pro-drop properties of Haitian Creole. As in Fongbe, the subject position of the raising verb *genle* ‘seem’ must be empty (20a), while there is an optional pronoun in the subject position of the verb *difisil* ‘to be difficult’ (20b).

(20) Haitian Creole:

- a. DeGraff (1993: 72), (cf. Example (7)):

ø genle Jak damou.

seem Jack be.in.love

‘It seems that Jack is in love.’

- b. Déprez (1994) (cf. Example (8)):

**(Li) difisil pou pale ak Jan.*

3.SG difficult to speak with John

‘It is difficult to speak with John.’

Given the importance of Fongbe as a substrate language of Haitian Creole, such parallelisms are to be expected. The more general question is to what extent such correlations can be generalized, or more specifically in the context of this paper, to what extent there is covariation between the existence of morphological paradigms with the features of person, number (and gender/noun class) on the verb and the radical-pro-drop properties of creoles. This question can be studied more systematically since the publication of the *Atlas of Pidgin and Creole Language Structures* (APiCS; Michaelis et al. 2013). APiCS lists no less than 76 languages under feature 62 ‘Expression of pronominal subject’. Out of these 76 creoles/pidgins, 49 have ‘obligatory pronoun words’ and are thus non-pro-drop (64,5%). Only 18 are radical pro-drop (23,7%).¹⁰ The remaining nine creoles and pidgins have the features of ‘pronoun affixes’ (four languages), ‘subject pronouns in different positions’ (two languages) and ‘mixed behaviour of pronominal subjects’ (three languages).

The remainder of this section will focus on the creoles and pidgins with obligatory pronouns and with radical pro-drop. The creoles with English, Dutch and French as lexifier languages are mostly non pro-drop. As is summarized in

10. Michaelis et al. (2013) use the value ‘optional pronoun words’ for these languages. Since none of these languages show cross-reference marking on the verb, this value corresponds to radical pro-drop.

Table 1, 32 creoles/pidgins with obligatory arguments behave like their lexifier languages English, Dutch and French. 24 of these creoles/pidgins have developed out of contact situations with African languages.

Table 1. Number of non-pro-drop creoles/pidgins with non-pro-drop lexifier languages

<i>Lexifier language</i>	<i>Number of languages</i>
English-based (with African influence)	23 (17); cf. (21)
Dutch-based (with African influence)	3 (2); cf. (22)
French-based (with African influence)	6 (5); cf. (23)
Total	32 (24)

The details on the individual creoles/pidgins involved are listed in (21) to (23):

- (21) English-based creoles/pidgins (languages with African influence are marked by *):
23 languages: *Early Sranan, Sranan, *Saramaccan, *Nengee, *Creolese, *Trinidad English, *Vincentian Creole, *Jamaican, *Belizean Creole, *San Andres Creole, *Nicaraguan Creole, Bahamian Creole, *Gullah, *African American English, *Krio, *Ghanaian Pidgin English, *Nigerian Pidgin English, *Cameroon Pidgin English, *Pichi, Tok Pisin, Norf’k, Kriol, Hawai’i Creole.
- (22) Dutch-based creoles/pidgins (languages with African influence are marked by *):
3 languages: Negerhollands, *Berbice Dutch, *Afrikaans.
- (23) French-based creoles/pidgins (languages with African influence are marked by *):
6 languages: *Haitian Creole, *Guadeloupean Creole, *Martinican Creole, *Guyanais, *Louisiana Creole, Tayo.

In the case of creoles/pidgins based on Portuguese and Spanish, we find another ten languages with obligatory pronouns:

- (24) Non-pro-drop creoles/pidgins based on Portuguese (P) and Spanish (S) as their lexifier language (languages with African influence are marked by *):
*Cap Verdean Creole of Santiago (P), *Cap Verdean Creole of Brava (P), Cap Verdean Creole of São Vicente (P), *Guinea-Bissau Kriyol (P), *Casamances Creole (P), *Santome (P), *Angolar (P), *Principense (P), *Fa d’Ambô (P), *Palenquero (S).

In addition to the 42 creoles/pidgins discussed so far, the following seven languages are also listed as non-pro-drop:

- (25) Kikongo-Kituba: Major lexifier: Kikongo-Kimanyang (Bantu)
Other: many other Bantu languages
Fanakalo: Major lexifier: Zulu (Bantu)
Other: Afrikaans, English

Kinubi:	Major lexifier: Sudanese Arabic Other: Egyptian Arabic, Swahili, English
Juba Arabic:	Major lexifier: Sudanese Arabic Other: Egyptian Arabic, Bari, Pojulu, Dinka, Shilluk
Pidgin Hindustani:	Major lexifier: Fiji Hindi (Hindi spoken on Fiji island) Other: Fijian, English
Gurindji Kriol:	Major lexifier: Gurindji (Pama-Nyungan), Kriol Other: English
Chinook Wawa:	Major lexifier: Coastal Chinook Other: English, Nuuchahnulth (Wakash), French

Each of the creoles/pidgins in (25) have developed out of languages with inflectional morphological paradigms. The Bantu type of verbal morphological paradigms involved in Kikongo-Kituba and Fanakalo has been described in Subsection 3.2. As is well-known, Arabic and Hindi (cf. Kinubi, Juba Arabic; Pidgin Hindustani) have inflectional paradigms as well. Coastal Chinook as the lexifier language of Chinook Wawa is characterized by rich morphological templates. Finally, Gurindji Kriol is a mixed language (McConvell 2008) spoken in the Victoria River District in northern Australia. The languages involved are Gurindji (Pama-Nyungan with rich inflectional paradigms and Kriol, an non-pro-drop English-based creole (cf. (21)).¹¹

The discussion so far shows that each of the 49 non-pro-drop creoles/pidgins have emerged in the context of languages with inflectional morphological paradigms and thus confirms the correlation between the impossibility of omitting information on the subject (and the object) in creoles/pidgins and the presence of morphological paradigms in the contact languages involved. However, a look at the 18 creoles/pidgins with radical pro-drop shows that quite a few of them are related to Indo-European contact languages with inflectional morphology. The following Table lists no less than twelve such languages:

What is quite remarkable about these figures is that the ratio between non-pro-drop and radical pro-drop is much higher with the non-pro-drop languages English and French than with the “mild” pro-drop languages Portuguese and Spanish:

11. The two languages are combined as follows: “[B]asic verbs such as ‘go’ and ‘sit’, the tense-aspect-mood system and transitive morphology are derived from Kriol, whereas emphatic pronouns, possessive pronouns, case markers and nominal derivational morphology have been transplanted from Gurindji relatively intact, but with some innovations. Demonstratives, nouns, verbs and adpositions are adopted from both languages, however some generalisations can be made about their distribution. The coverbs of Gurindji compound verbs can appear in Gurindji Kriol as verbs” (McConvell 2008:189).

Table 2. List of radical pro-drop creoles/pidgins with Indo-European lexifier languages

Lexifier language	Number of creoles/pidgins	Creoles/pidgins concerned
English-based	2	Chinese Pidgin English, Singlish
French-based	2	Mauritian Creole, Seychelles Creole
Portuguese-based	5	Diu Indo-Portuguese, Korlai, Sri Lanka Portuguese, Papiá Kristang, Batavia Creole
Spanish-based	3	Ternate Chabacano, Cavite Chabacano, Zamboanga Chabacano

(26) Ratio of non-pro-drop/radical pro-drop:

- English: $23/2 = 11.5$
- French: $6/2 = 3.0$
- Portuguese: $9/5 = 1.8$
- Spanish: $1/3 = 0.3$

If one also takes into account the three Dutch-based creoles/pidgins, which are all non-pro-drop, one gets an overall ratio of $32/4 = 8.0$ for the non-pro-drop languages English, Dutch and French, while the ratio for Portuguese and Spanish is $10/8 = 1.24$.

In the case of English, the radical pro-drop properties of the two languages of Chinese Pidgin English and Singlish can be explained by the fact that the majority of the speakers of these languages have a Chinese background and thus apply the radical-pro-drop properties of Chinese to these two languages. Thus, the non-pro-drop properties of English are too weak to show up in Chinese Pidgin English and Singlish. In the case of the French-based Mauritian Creole, Adone (1994) points out that pro-drop is a recent phenomenon. In earlier times, there were no null pronouns in Mauritian Creole. Moreover, there are significant constraints on pro-drop in more recent Mauritian Creole. As Syea (1993) points out, pro-drop is only possible with nouns denoting humans, it depends on the presence of certain tense-aspect markers and it is ungrammatical with intransitive verbs. To what extent similar conditions apply to Seychelles Creoles needs more research.

In the case of the Portuguese-based and Spanish-based creoles/pidgins, nine out of the ten languages with obligatory pronouns have African substrates (cf. (24)). Thus, the importance of non-pro-drop in African languages as described in Subsection 3.2 seems to be reflected in these creoles/pidgins. In contrast, three of the five Portuguese-based radical pro-drop languages are spoken in India with Indo-Aryan or Dravidian substrates (Diu Indo-Portuguese, Korlai, Sri Lanka Portuguese). The substrates of both of these language families have inflectional morphological paradigms and both of them have properties of “mild” pro-drop. It may

be for the latter reason that the corresponding creoles/pidgins only have optional use of pronouns. Moreover, some of the person/number distinctions are often blurred in modern Indo-Aryan languages (cf. Masica 1990: 263).¹² The radical-pro-drop properties of the other two Portuguese-based creoles with radical pro-drop, i.e. Papiá Kristang and Batavia Creole (cf. Table 2), are most likely due to the influence of the pro-drop properties of their substrates, i.e. Malay and Hokkien in the case of Papiá Kristang and Malay and Javanese in the case of Batavia Creole. The three Spanish-based creoles with radical pro-drop (cf. Table 2) are all spoken in the Philippines with Philippine substrates (Austronesian), which have no inflectional morphological paradigms marking person and number.

Summarising the twelve creoles/pidgins with radical pro-drop based on English, French, Portuguese and Spanish, one can say that most of them have substrate languages with no inflectional morphological paradigms. However, there remain the three Portuguese-based radical pro-drop languages spoken in India for which one may expect absence of radical pro-drop because all or most contact languages have morphological paradigms.

The remaining six creoles/pidgins with radical pro-drop are listed in (27) together with their contact-languages:

- (27) Six additional creoles/pidgins with radical pro-drop:
- | | |
|-------------------------|---|
| Chinese Pidgin Russian: | Major lexifier: Russian |
| | Others: Chinese, Mongolian, Tungusic |
| Sri Lankan Malay: | Major lexifier: Malay |
| | Others: Shonam (Sri Lanka Muslim Tamil), Sinhala |
| Ambon Malay: | Major lexifier: Vehicular Malay |
| | Others: Portuguese, Dutch, vernacular languages of the Moluccas |
| Pidgin Hawaiian: | Major lexifier: Hawaiian (Austronesian) |
| | Others: English, Cantonese, Hakka, Portuguese |
| Yimas-Arafundi Pidgin: | Major lexifier: Yimas (Papua: Ramu phylum: Lower Sepik) |
| | Other: Arafundi (Papua: Piawi) |
| Sango: | Major lexifier: Ngbandi, Yakoma (Niger-Congo: Adamawa-Ubangi) |
| | Other: French |

12. Interestingly enough, none of the Romance languages developed into a radical-pro-drop language. This may be taken as another evidence of the strength of inflectional morphological paradigms. Only in situations of intensive contact as in the cases of Diu Indo-Portuguese, Korlai and Sri Lanka Portuguese was it possible to combine loss of inflectional morphology with pro-drop. Obviously, that option exists but, as is also pointed out at the end of this section, it is not an option that is frequently used even in the context of pidgins and creoles.

In the case of Chinese Pidgin Russian, radical pro-drop may be assigned to the influence of Chinese and Mongolian (Tungusic has cross-reference on the verb). In addition, Russian is a “mild” pro-drop language with subject cross-reference on the verb but no obligatory subject arguments. In Sri Lankan Malay and Ambon Malay, the lexifier language Malay is radical pro-drop. Similarly, the pro-drop properties of Hawaiian (Austronesian) can be said to be responsible for the radical pro-drop properties of Pidgin Hawaiian. The case of Yimas-Arafundi is different. Both Papuan languages involved are polysynthetic with bound pronouns within the verb. In this case, the explanation for radical pro-drop may be due to the fact that the pronominal elements in the verb are prefixes in Yimas and suffixes in Arafundi. This difference in morpheme order may have led to omitting morphological marking of cross-reference on the verb because the morphological system remained mutually opaque (McConvel 2008: 206). The case of radical pro-drop in Sango is of particular interest. Yakoma has no strict morphological paradigm but it has obligatory person, number and tense-aspect-mood marking, partly expressed by an individual lexeme that is described as a “person substitute” by Boyeldieu (1975: 40–71; *substitut personnel* in the French version). This person substitute expresses person and tense-aspect-mood, while number is marked partly on the substitute and partly on the verb by differences in tonality. Thus, Yakoma reflects an interesting stage between total loss of morphology and full syntacticization of person/number marking. In spite of this, Sango as a radical pro-drop language remains a counter example to the hypothesis of the strength of morphological paradigms in creoles because of the obligatoriness of expressing person/number features of verbal arguments.

Looking at the 49 non-pro-drop languages and the 18 radical pro-drop languages discussed above, all the 49 creoles/pidgins with obligatory pronouns are related to languages with verbal inflectional morphological paradigms marking person and number. Among the 18 radical pro-drop languages, only five (27,8%) are problematic. Seychelles Creole is radical pro-drop even though French has morphological paradigms and is non-pro-drop. The three Portuguese-based creoles Diu Indo-Portuguese, Korlai and Sri Lanka Portuguese are radical pro-drop even though the contact languages involved (Indo-Aryan, Dravidian, Portuguese) have morphological paradigms. Finally, one would expect Sango to have obligatory pronominals given the marking pattern described above.

Since pidgins and creoles are the result of extreme contact situation, one would have to expect a certain number of counterexamples. The fact that all 49 non-pro-drop languages confirm the hypothesis of morphological strength and only five out of the 18 radical-pro-drop languages disconfirm it, we get an overall confirmation rate of 92.5% (62/67) against 7.5% of counterexamples. This can

be taken as remarkable evidence for the effect of morphological paradigms on the presence/absence of radical pro-drop in pidgins and creoles.

5. Radical analyticity and radical pro-drop: Scenarios of variation

5.1 Introduction

The previous sections showed that the presence of inflectional morphological paradigms for person/number/(gender/noun class) as well as their absence can be very stable across time. Similarly, the presence and the absence of radical pro-drop show remarkable time stability. This section will show how this situation can be modelled. It crucially depends on the assumption of two types of complexity (hidden vs. overt complexity, cf. Section 5.2) plus a combination of factors which operate against the development of synthetic structures (Section 5.3). The properties that characterize incipient morphology if it develops in such environments is the topic of Section 5.4. Finally, Section 5.5 discusses the typological status of radical analyticity and radical pro-drop.

5.2 Hidden vs. overt complexity

The activity of speaking takes place between the two poles of explicitness and economy. Each utterance can be seen as a balance between the extent to which information needs to be phonologically encoded (explicitness) and to what extent it can be left to pragmatic inference (economy) for successful communication. The competition between these two poles is determined by what Levinson (2000) describes as the “articulatory bottleneck” in the context of information transmission. Human speech encoding is by far the slowest part of speech production and comprehension. Other processes like prearticulation, parsing and comprehension show a much higher transmission rate. This contrast creates a cost-related asymmetry between articulation and pragmatic inference in the sense that “[i]nference is cheap, articulation expensive” (Levinson 2000:29). Thus, every utterance can be seen as the result of the two competing motivations of economy and explicitness.

Economy and explicitness also have their impact on the grammar of individual languages – an observation that goes back to von der Gabelentz (1891: 251). Much later, the competition between these two forces was modeled in various theoretical approaches (e.g. economy vs. iconicity in Haiman (1983) or markedness vs. faithfulness in Optimality Theory (e.g. Kager 1999)). At the level of individual grammatical categories like tense, aspect, person, number or definiteness, the

competing motivations of economy vs. explicitness show their effects in two ways (for examples from EMSEA languages, cf. (28) and (29) below):

i. Obligatoriness:

– Explicitness [obligatoriness]:

The grammar of a language X forces the speaker to encode a certain grammatical category that exists in X in a given morphosyntactic environment. It forces the speaker to select a particular marker with a particular value out of the set of markers existing in X for expressing that category (examples: the selection of a number marker with count nouns or the selection of a tense marker on the verb).

– Economy [no obligatoriness]:

The grammar of a language X *does not* force the speaker to encode a certain grammatical category that exists in X in a given morphosyntactic environment (examples: X has a set of number markers but their use with count nouns depends on pragmatic inference / X has a set of tense markers but their use on the verb depends on pragmatic inference).

ii. Multifunctionality:

– Explicitness [no multifunctionality]:

A marker expresses a specific value of a single grammatical category.

– Economy [multifunctionality]:¹³

The functional range of a grammatical marker covers values belonging to more than one category or covering a broad range of functions within one and the same category.

The effects in (i) and (ii) have their impact on complexity. Explicitness leads to the type of complexity which is generally discussed in the rich typological literature and is described by its obligatoriness and its fine-grained semantic distinctions as well as by the presence of inflectional morphology (McWhorter 2001; Dahl 2004; Sinnemäki 2011 and others). Since this type of complexity is based on overt formal marking of clearly distinguishable grammatical categories, I called it “overt complexity” in Bisang (2009). In contrast, the economy-driven effects of non-obligatoriness and multifunctionality create morphosyntactic surface structures that look simple even though their interpretation creates additional costs because

13. Notice that this definition does not include cumulative morphology or *portemanteau* morphemes. Morphemes of this type combine functions from different categories but they select a specific value from each of the categories involved. Thus, the Latin suffix *-orum* in *amic-orum* [friend-GEN.PL.M] ‘of the friends’ marks the values genitive from the domain of case, plural from number and masculine from gender.

it needs additional inferential effort. This type of pragmatics-based complexity is called ‘hidden complexity’ in Bisang (2009, 2014b, 2015a). A good example of this type of complexity is radical pro-drop, since the relevant information concerning person, number (and gender/noun class) features of nominal arguments must be pragmatically inferred from context. Radical pro-drop is not simply a matter of omitting these features, it also affects pragmatic inference in terms of Generalized Conversational Implicatures (Levinson 2000) because it generates a three way Horn-scale of the type <zero, pronominal, lexical noun>, while non-pro-drop creates a two-way scale of the type <pronoun, lexical noun> (for the details, cf. Huang 1994 on Chinese).

In analogy to traditional approaches to complexity, languages can also be compared with regard to overt vs. hidden complexity. From our description of pro-drop properties in Sections 2 and 4, EMSEA languages clearly show a higher degree of hidden complexity with regard to that category than analytic West African languages and the majority of pidgins and creoles. Thus, the case of pro-drop shows that creole grammars are not necessarily the grammars that generate the most simple overt structures. Nor is it the case that older languages necessarily accumulate high degrees of overt complexity (Dahl 2004; McWhorter 2001, 2005) as is shown by the pro-drop properties of EMSEA languages. What drives the outcome of linguistic change in terms of complexity are the two competing motivations of economy and explicitness. The two motivations are competing in each individual case of linguistic change and lead to different types of maturation. If at a given stage of the grammar of a language X, explicitness wins for a given grammatical category, we get a type of complexity that is the result of explicitness-based maturation. This type of maturation is described by Dahl (2004). It generates phenomena like complex word structure (inflectional, derivational and incorporating morphology), lexical idiosyncrasies (grammatical gender, inflectional classes, idiosyncratic case marking), syntactic phenomena that depend on morphology (agreement), specific marking of subordinate clauses and morpheme and word-level features in phonology (for more, cf. Dahl 2004: 114–115). If economy wins, we get economy-based maturation, a type of maturation which favours grammars that create seemingly simple surface structures that must be pragmatically enriched for certain grammatical categories that exist in a given language. The two types of maturation happen individually for individual morphosyntactic expressions of grammatical categories. If economy wins frequently enough across many different domains of grammar, we get structures as we find them in EMSEA languages. The following examples provide some idea of the impressive number of grammatical domains which are characterized by economy-based maturation and hidden complexity (for more examples specifically from Chinese, also cf. Xing 2013, 2015):

- (28) Lack of obligatoriness:
- Radical pro-drop (Huang 1984; Bisang 2014a)
 - Non-obligatory marking of tense-aspect
(e.g. the perfective marker *-le*; cf. Li 2014; Bisang 2004, 2020)
 - Non-obligatoriness of number marking on the noun
(e.g. Mandarin: *-men* is obligatory only with personal pronouns)
 - Non-obligatory use of classifiers in the context of definiteness and indefiniteness (Li and Bisang 2012)
- (29) Multifunctionality:
- ‘get’-verbs / ‘come to have’-verbs (Enfield 2003; Bisang 1992)
 - ‘give’-verbs: benefactive, causativity, adverbial subordination (purpose), complementizer (Song 1997; Thepkanjana and Uehara 2008; Bisang 1992)
 - ‘be at’-verbs: durativity marker, locative preposition (Bisang 1992)
 - Classifiers as markers of individuation or atomization (numeral classifiers) and in the the context of identifiability (definiteness and indefiniteness) (Li and Bisang 2012)
 - Kinship terms in the function of pronouns
(marking 1st, 2nd and 3rd person in Vietnamese: Tuc 2003)

5.3 Factors against the development of extensive inflectional morphological paradigms

The cells in morphological paradigms as defined in Section 1 can either be obligatory or optional. Example (30a) from Turkish shows inflectional markers which express values of categories that are obligatory in an independent declarative clause (past, 1st singular), while (30b) additionally contains two markers specifying facultative categories (potential, negation):

- (30) Turkish:
- a. *gel-di-m.*
come-PST-1.SG
‘I came.’
 - b. *gel-e-me-di-m.*
come-POT-NEG-PST-1.SG
‘I was not able to come.’

Since the expression of the features which are relevant for the inflectional morphological paradigms associated with radical pro-drop (person, number, gender/

noun class) is typically obligatory,¹⁴ their occurrence in discourse is relatively frequent. As a consequence, one may argue that the frequency of these features in West African languages creates a certain expectation of their overt presence in discourse which is hard to overcome for speakers once they have acquired their mother tongue. If the morphological paradigm gets lost in processes of linguistic change, speakers who are not used to pragmatically infer this information start compensating that loss by using pronouns. In contrast, speakers of languages with no inflectional morphological paradigms of this type are used to infer the relevant information and do not need the use of pronouns or lexical nouns for retrieving it. This scenario generates languages of the EMSEA type with radical pro-drop. In the case of pidgins and creoles, both scenarios may apply. Creoles are often seen as the results of extreme contact situations in which speakers maximally limit their grammars to what is needed for communication (McWhorter 2001, 2005). The observation that the majority of pidgins and creoles do not show radical pro-drop (Section 4) indicates at least for this case that these languages do not necessarily reflect the maximal limitations of what is needed for communication. The extent to which grammars allow speakers to omit grammatical information (person/number but also tense or (in)definiteness, cf. Bisang 2015a) reveals that speakers are not fully free to what extent they can reduce grammatical information even in contexts of pidgin and creole formation. In the case of radical pro-drop, this means that the presence of inflectional morphological paradigms enhances the need of avoiding it by using pronouns, while the absence of such paradigms provides no obstacles in leaving that information to pragmatic inference and thus supports radical pro-drop.

A second property of morphological paradigms that even applies to facultative elements is the fact that their markers belong to clearly determined semantic domains like person, number, gender/noun class, tense, etc. Since the grammatical markers of EMSEA languages are characterized by the frequent absence of these two properties (cf. non-obligatoriness and multifunctionality in hidden complexity, Section 5.2), they lack essential preconditions for initiating processes of paradigm development. For that reason, the likelihood of the emergence of paradigms in EMSEA languages is relatively low.

In addition, processes of morphologization are generally affected by phonological and sociolinguistic factors. Phonologically, EMSEA languages are characterized by two rather strong constraints (Ansaldò and Lim 2004, Bisang 2008):

14. This works at least for subject agreement. In the case of object agreement, the selection of the agreement marker often depends on the referential status of the object (definiteness, specificity).

- The discreteness of syllable boundaries and the strong tendency to avoid sub-syllabic morphemes.
- Phonotactic restraints

These two constraints together affect the formal side of grammaticalization, which is characterized by phonetic erosion in terms of duration and vowel quality rather than by morphological reduction (Ansaldo and Lim 2004). Of additional importance is tonality. In languages with more than one tonal register like Cantonese and Hokkien, pitch tends not to be reduced in processes of grammaticalization because tonal contrast must be maintained for reasons of keeping up semantic contrasts. In Sinitic languages like Mandarin with only one tonal register, the distinctiveness of pitch is less important, and this allows loss of tone more easily (cf. perfective *-le* from *liǎo* ‘finish’ in Chinese).

The sociolinguistic factor that works against the development of morphology is linguistic contact. Even if there is morphology in an EMSEA-type language, the fact that the function expressed by it is marked by an independent word in the surrounding languages may favour the grammaticalization of new markers from the existing lexicon of that language. A good example is the formation of agent nouns in Khmer, which can be formed by the infix *-m-* as in *s<m>ò:m* ‘beggar’ (from *sò:m* ‘ask, ask a favour’) or *ch<m>am* ‘guard [n.]’ from *cam* ‘wait for, guard, keep’. The more productive strategy of agent noun formation in Khmer is identical to the strategies used in Thai, Vietnamese and Sinitic, i.e., the use of a head with the meaning of ‘person, man’ like Khmer *nèək* ‘person’. Thus, compounds of the type *nèək-daa* [person-walk] ‘pedestrian’ or *nèək-taey* [person-compose/write] ‘author, writer, composer’ are much more common in contemporary Khmer. If there is such a tendency in contact situations to avoid morphological strategies at the level of word formation, the probability of using affixation for expressing inflectional categories in Khmer and the vast majority of EMSEA languages in general is even less likely, since these categories are expressed by independent non-obligatory morphemes.

5.4 Morphological paradigms in EMSEA languages

Even though EMSEA languages are generally described as isolating or radically analytic this characterization is only partly true. There are some examples of morphological paradigms in the domain of the noun. In Northern Kam (Tai-Kadai), the classifier occurs in two forms for marking singular vs. plural. The most elaborate morphological paradigm of classifier inflection is found in Weining Ahmao, a Hmong-Mien language spoken in Southwest China (Wang 1957, Gerner and Bisang 2008, 2009). In this language, the classifier paradigm consists of twelve forms that combine the categories of number (singular vs. plural), definiteness (definite vs. indefinite) and size (augmentative, medial and diminutive). Examples like these provide important evidence of the presence of inflectional morphology in a linguistic area that is generally assumed to have no inflectional morphology.

There are also examples of inflectional morphological paradigms from the domain of the verb. It is quite remarkable that paradigms expressing person and number are comparatively rare. If they occur, they have similar properties as paradigms expressing other categories. Since these properties are characterized by hidden complexity, one may argue that they represent a specific type of EMSEA-related incipient morphology. To show this, this section presents some data from Sinitic on tense-aspect distinctions and from Semelai (Mon-Khmer: Aslian) on agreement marking on the verb.¹⁵ In Sinitic, verb agreement is unattested to the best of my knowledge. If so, this shows a remarkable diachronic continuity which reaches back at least to the time of Old Chinese (11th – 6th centuries BC). The only family in which there are some languages with person/number marking on the verb is the Mon-Khmer family (but more research will be needed for most branches, with Aslian being rather an exception).

Inflectional paradigms expressing mainly tense/aspect in Sinitic

The existence of inflectional morphology that goes beyond concatenative morphology is found in Sinitic varieties spoken in the provinces of Henan, Shandong, Shanxi and to some extent also in Shaanxi and Hebei. Most of these varieties are spoken in Mandarin and Jin languages. Since the descriptions of these languages are mostly in Chinese, this fact is made accessible to Western linguists only recently by Arcodia (2013, 2015). The morphology that is found in these languages is due to the fusion within sequences of two verbs [$V_1 V_2$], in which V_2 forms a single syllable together with V_1 , with V_2 being a resultative verb, a directional verb or a tense-aspect marker. As a result, verbs generally occur in a basic form and in a derived form, whose phonological structure is mostly predictable from the

15. This is a significantly shortened version of what is said in Section 5 from Bisang 2020.

rhyme structure of the basic form. The differences between the two forms manifest themselves in the use of different vowels, tonal morphology or the integration of a rhotaic consonant into the coda position of the stem syllable. To give an example from the dialect of Xunxian (浚县, Central Plains Mandarin, Henan), the verb which corresponds to Mandarin 改 *gǎi* ‘change’ is *kai⁵⁵* in its basic form and *kɛ⁵⁵* in its derived form. Even though these forms often look like binary *ablaut* paradigms (Arcodia 2015:17) for expressing two different values of a grammatical category, they differ in at least two ways from such systems as we know them from familiar Indo-European languages (cf. English *sing* vs. *sang*): (i) the derived form is often multifunctional and (ii) the degree of obligatoriness remains often unclear from the descriptions provided by the individual grammars. With these properties, the two morphological forms of the verb clearly show hidden complexity. Even though more research is needed in many of these languages, one may look at these patterns as instances of a specific type of “East and mainland South-east Asian inflectional morphological paradigms”.

Person/number marking on the verb in the Aslian family of Mon-Khmer

There are various Mon-Khmer languages which have at least some inflectional markers, but these markers must have developed in more recent times. Of particular interest are Aslian languages like Temiar (Benjamin 1976) and Semelai (Kruspe 2004) with their person markers on the verb. As one can see from Kruspe’s (2004:171) grammar of Semelai, these markers are proclitics whose forms are phonologically very similar to the pronouns of that language. This indicates that verbal cross-reference is a rather recent phenomenon. Moreover, the use of cross-reference markers is subject to various restrictions. Two of them are briefly mentioned here. (i) The proclitics mainly occur with transitive verbs and with two classes of intransitive verbs (Kruspe 2004:159). (ii) Their use in these contexts is not obligatory. It depends on pragmatic inference inasmuch as it is limited to individuated transitive events. Generic concepts are not marked on the verb. This further corroborates the conclusion that the inflectional morphology (person and number marking) of Semelai is a rather recent phenomenon that does not allow any conclusions concerning the presence of inflectional morphological paradigms in Proto-Mon-Khmer. In the case of their obligatoriness, their use depends on the pragmatic factor of concrete identifiability.

5.5 On the typological status of radical analyticity and radical pro-drop

Radical analyticity and radical pro-drop are treated as linguistically remarkable in various ways. McWhorter (2016: 49, 56, 84) describes radical analyticity as ‘diachronically anomalous’, arguing that languages with complete or near-complete absence of inflectional morphology can only be the result of incomplete non-native acquisition in situations of language contact. If the languages of East and mainland Southeast Asia are a hotbed of radical analyticity, this must be because of Chinese, whose contact-induced radical analyticity as it can be observed since Old Chinese spread over the whole area of EMSEA languages. In the scenario developed in this paper, radical analyticity and the absence of morphological paradigms is seen as the result of another type of maturation which does not depend on contact even though it can be kept alive and enhanced by contact (cf. the example of Khmer morphology in Section 5.3). As I tried to show, highly limited inflectional morphology is not only a property of Chinese, it is very likely that it was characteristic of Tai-Kadai, Mon-Khmer and Hmong-Mien as well, at least in the last two thousand years which are relevant for shaping the languages of East and mainland Southeast Asia with all their different types of cross-linguistic variation in the functions expressed by their grammatical markers (numeral classifiers, classifiers in the context of (in)definiteness, tense-aspect markers, directional verbs, coverbs/adpositions, causative verbs, etc.).¹⁶ Moreover, the strong tendency of being radically analytic does not necessarily depend on sociolinguistic and geographic criteria such as low-contact and high-contact (Trudgill 2011),¹⁷ since radical analyticity can be dominant in both conditions.

Hidden complexity and economy-based maturation are not limited to morphology. They also operate at the level of syntax, as in the case of radical pro-drop – a phenomenon that is a theoretical challenge in the generative model at least since Rizzi (1986). Interestingly enough, morphological paradigms are not only seen as crucial for radical pro-drop in the present paper. Rizzi (1986) argues that pro-drop is only possible in languages with rich agreement morphology for person and number on the verb. Jaeggli and Safir (1989) further specify this claim by explicitly mentioning inflectional paradigms and by arguing that pro-drop is only possible if each cell of a given paradigm is filled by an affix (morphological uniformity). Unfortunately, these approaches fail to account for radical pro-drop

16. In the case of Austronesian, the inflectional morphology that is reconstructed for that family does not cover the features associated with radical pro-drop (also cf. footnote 1).

17. Also cf. the distinction between “hill” cultures or low-contact situations and “valley” cultures and high-contact situations (on the concept of hill vs. valley cultures, cf. Burling 1965; Scott 2009).

because they take overt agreement marking on the verb as we find it in “mild” pro-drop for granted. Alternative explanations like topic drop (Huang 1984) or blocking by determiners (Speas 1994) are controversial as well. For that reason, Neeleman and Szendrői (2007) take their focus on paradigms again, this time on pronominal systems. In their theory, radical pro-drop is allowed only in languages with agglutinative pronominal systems as they are found, e.g., in Mandarin with its pronouns for 1st, 2nd and 3rd person that take the suffix *-men* for expressing plural. Since there are empirical counterexamples to this theory (Bisang 2014a: 33–35 on the pronominal system of Hmong and on older stages of Chinese) there is still quite a lot of room for looking at different explanations. Given the absence of uncontroversial UG-based accounts, I suggest another model which does not need the assumption of Universal Grammar.

6. Conclusion

The account presented in this paper explicitly takes a historical stance that starts out from an evolutionary perspective on how languages develop within families and across families in different situations of contact. Based on two types of diachronic maturation (explicitness-based vs. economy-based) with their effects on overt vs. hidden complexity, it argues that the presence of inflectional verbal morphological paradigms keeps the frequency of the features of person, number and gender/noun class as they are associated with nominal arguments relatively high. As a consequence, adult speakers are not used to pragmatically infer them from context and thus tend to keep the relevant information by using pronouns if the morphological paradigm is about to disappear as in the case of radically analytic West African languages. In East and mainland Southeast Asia, the languages from the relevant families did not have elaborate morphology in the last two thousand years. In particular, they did not have inflectional morphological paradigms with the features of person, number (and gender/noun class) nor had they strong constraints on the obligatory use of pronouns – a situation which further enhanced radical pro-drop. Other factors like contact and phonological properties operated against the development of morphological paradigms without completely suppressing their emergence. If they arise, they often have specific properties which are still associated with hidden complexity.

If these scenarios are correct, radical analyticity and radical pro-drop are by no means extraordinary. Given the right structural properties and the right situation of contact at the right time, they can develop as easily as any other cross-linguistically attested phenomena, among them complex inflectional morphological paradigms and absence of radical pro-drop.

Abbreviations

1, 2, 3	first, second, third person	OP	Operator (nominal)
II, III	Bantu noun classes	PAS	Passive marker
ACC	Accusative	PB	Proto-Bantu
ALL	Allative	PFV	Perfective
APPL	Applicative marker	PL	Plural
ASP	Aspect marker	PNC	Proto-Niger-Congo
AUX	Auxiliary	POT	Potential marker
CAUS	Causative marker	PROG	Progressive marker
COMP	Complementizer	PRS	Present
DEF	Definiteness marker	PST	Past
F	Feminine marker	Q	Question marker
FIN	Finiteness marker	REL	Relative-clause marker
FV	Final vowel	SBJ	Subject marker
HTS	High Tone Syllable	SG	Singular
INTR	Intransitive marker	SUBORD	Subordinator
NEG	Negation	TA	Tense-aspect marker
OBJ	Object pronoun	TAM	Tense-Aspect-Modality

References

- Ackema, P., Brandt, P., Schoorlemmer, M. and Weerman, F. (eds.). (2006). *Arguments and Agreement*. Oxford: Oxford University Press.
- Adelaar, A. (2005). The Austronesian languages of Asia and Madagascar: A historical perspective. In: A. Adelaar and N.P. Himmelmann (eds.), *The Austronesian Languages of Asia and Madagascar*. Oxford: Routledge.
- Adone, D. (1994). Creolization and language change in Mauritian Creole. In: D. Adone and I. Plag (eds.), *Creolization and Language Change* (pp. 24–43). Tübingen: Niemeyer. <https://doi.org/10.1515/97831133980123>
- Anderson, G.D.S. (2007). *The Munda Verb. A Typological Perspective*. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110924251>
- Anderson, G.D.S. (2016). Overview of the Munda languages, In: M. Jenny and P. Sidwell (eds.), *The Handbook of Austronesian Languages*, Vol. 1 (pp. 364–414). Leiden and Boston: Brill.
- Ansaldo, U. and Lim, L. (2004). Phonetic absence as syntactic prominence. Grammaticalization in isolating tonal languages. In: O. Fischer, M. Norde, and H. Perridon (eds.), *Up and Down the Cline – the Nature of Grammaticalization* (pp. 345–362). Amsterdam and Philadelphia: Benjamins. <https://doi.org/10.1075/tsl.59.18ans>
- Arcodia, G.F. (2013). Grammaticalization *with* coevolution of form and meaning in East Asia? Evidence from Sinitic. *Language Sciences* 40, 148–167. <https://doi.org/10.1016/j.langsci.2013.05.002>
- Arcodia, G.F. (2015). More on the morphological typology of Sinitic. *Bulletin of Chinese Linguistics* 8, 5–26.

- Baerman, M. and Corbett, G. (2010). Introduction: Defectiveness: Typology and diachrony. In: M. Barerman, G. Corbett and D. Brown (eds), *Defective Paradigms. Missing Forms and what they Tell us* (pp. 1–18). Oxford: Oxford University Press.
<https://doi.org/10.5871/bacad/9780197264607.003.0001>
- Baxter, W.A. and Sagart, L. (2014). *Old Chinese. A New Reconstruction*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199945375.001.0001>
- Benjamin, G. (1976). An outline of Temiar grammar. In: P.N. Jenner, L. C. Thompson and S. Starosta (eds.), *Austroasiatic Studies I and II* (pp. 129–188).
- Bickel, B. (2007). Typology in the 21st century: Major current developments. *Linguistic Typology* 11, 239–251. <https://doi.org/10.1515/LINGTY.2007.018>
- Bisang, W. (1992). *Das Verb im Chinesischen, Hmong, Vietnamesischen, Thai und Khmer*. Tübingen: Niemeyer.
- Bisang, W. (2004). Grammaticalization without coevolution of form and meaning: The case of tense-aspect-modality in East and mainland Southeast Asia. In: W. Bisang, N.P. Himmelmann and B. Wiemer (eds.), *What Makes Grammaticalization? – A Look from its Fringes and its Components* (pp. 109–138). Berlin: Mouton de Gruyter.
<https://doi.org/10.1515/9783110197440.2.109>
- Bisang, W. (2008). Grammaticalization and the areal factor: The perspective of East and mainland Southeast Asian languages. In: M. J. López-Couso and E. Seoane (eds.), *Rethinking Grammaticalization. New Perspectives* (pp. 15–35). Amsterdam and Philadelphia: John Benjamins. <https://doi.org/10.1075/tsl.76.04bis>
- Bisang, W. (2009). On the evolution of complexity – sometimes less is more in East and mainland Southeast Asia. In: G. Sampson, D. Gil and P. Trudgill (eds.), *Language Complexity as an Evolving Variable* (pp. 34–49). Oxford: Oxford University Press.
- Bisang, W. (2010). Grammaticalization in Chinese – a construction based account. In: E. C. Traugott and G. Trousdale (eds), *Gradience, Gradualness, and Grammaticalization* (pp. 245–277). Amsterdam and Philadelphia: Benjamins. <https://doi.org/10.1075/tsl.90.13bis>
- Bisang, W. (2013). Language contact between geographic and mental space. In: Auer, P., M. Hilpert, A. Stukenbrock and B. Szendrői, *Linguistic Perspectives on Space: Geography, Interaction, and Cognition* (pp. 61–100). Berlin: Mouton de Gruyter (Series *Linguae et Litterae*). <https://doi.org/10.1515/9783110312027.61>
- Bisang, W. (2014a). On the strength of morphological paradigms – a historical account of radical pro-drop. In: M. Robbeets and W. Bisang. (eds.), *Paradigm Change in Historical Reconstruction: The Transeurasian Languages and Beyond* (pp. 23–60). Amsterdam and Philadelphia: Benjamins.
- Bisang, W. (2014b). Overt and hidden complexity – two types of complexity and their implications. *Poznan Studies in Contemporary Linguistics* 50 (2), 127–143.
<https://doi.org/10.1515/psicl-2014-0009>
- Bisang, W. (2015a). Hidden complexity – the neglected side of complexity and its implications. *Linguistics Vanguard*. <https://doi.org/10.1515/lingvan-2014-1014>
- Bisang, W. (2015b). Problems with primary vs. secondary grammaticalization: the case of East and mainland Southeast Asian languages. *Language Sciences* 47, 132–147.
<https://doi.org/10.1016/j.langsci.2014.05.007>
- Bisang, W. 2020. Grammaticalization in Chinese – a cross-linguistic perspective”, in: J. Xingn (ed.), *A Typological Approach to Grammaticalization and Lexicalization: East Meets West*, (pp. 17–54). Berlin: Mouton de Gruyter.

- Bisang, W. and Sonaiya, R. (1999). The function of the High Tone Syllable in Yoruba. *The Journal of African Languages and Linguistics (JALL)* 20, 1–19. <https://doi.org/10.1515/jall.1999.20.1.1>
- Boyeldieu, P. (1975). *Études Yakoma, langue du groupe Oubangien (R. C. A.) Morphologie-synthématique. Notes linguistiques*. Paris, Éditions SELAF.
- Burling, R. (1965). *Hill Farms and Padi Fields: Life in Mainland Southeast Asia*. New Jersey: Prentice Hall. Inc.
- Cook, W. (1965). A Descriptive Grammar of Mundari. PhD dissertation, Georgetown University.
- Creissels, D. (2005). A typology of subject marker and object marker systems in African languages. In: E. F. K. Voeltz (ed.), *Studies in African Linguistic Typology* (pp. 445–459). Amsterdam and Philadelphia: John Benjamins.
- Dahl, Ö. (2004). *The Growth and Maintenance of Linguistic Complexity*. Amsterdam and Philadelphia: John Benjamins. <https://doi.org/10.1075/slcs.71>
- Déchaine, R.-M. (1993). Predicates across categories: towards a category-neutral syntax. Amherst, Mass. GLSA. (Doctoral dissertation, University of Massachusetts).
- DeGraff, M. F. (1993). Is Haitian a Pro-Drop language? In: F. Byrne and J. Holm (eds.), *Atlantic Meets Pacific: A Global View of Pidginization and Creolization* (pp. 71–90). Amsterdam and Philadelphia: Benjamins.
- Déprez, V. (1994). Haitian Creole: A Pro-Drop language? *Journal of Pidgin and Creole Languages* 9, 1–24. <https://doi.org/10.1075/jpcl.9.1.02dep>
- Devos, M. (2008). *A Grammar of Makwe*. München: LINCOM Europa.
- Diffloth, G. (2005). The contribution of linguistic paleontology to the homeland of Austroasiatic. In: L. Sagart, R. Blench and A. Sanchez-Mazas (eds.), *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics* (pp. 77–80). London: RoutledgeCurzon. https://doi.org/10.4324/9780203343685_chapter_5
- Donegan, P. and Stampe, D. (2004). Rhythm and the synthetic drift of Munda. In: R. Singh (ed.), *The Yearbook of South Asian Languages and Linguistics 2004* (pp. 3–36). Berlin: Mouton de Gruyter.
- Downer, G. B. (1959). Derivation by tone-change in classical Chinese. *Bulletin of the School of Oriental and African Studies* 22, 258–290. <https://doi.org/10.1017/S0041977X00068701>
- Edmondson, J. and Solnit, D. (1988). Introduction. In: J. Edmondson and D. Solnit (eds.), *Comparative Kadai: Linguistic Studies beyond Tai* (pp. 1–26). Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.
- Enfield, N. J. (2003). *Linguistic Epidemiology. Semantics and Grammar of Language Contact in Mainland Southeast Asia*. London: RoutledgeCurzon.
- Gabelentz, Georg von der. (1891). *Die Sprachwissenschaft, ihre Aufgaben, Methoden und bisherigen Ergebnisse*. Leipzig: T.O. Weigel Nachfolger.
- Gerner, M. and Bisang, W. (2008). Inflectional speaker-role classifiers in Weining Ahmao. *Journal of Pragmatics* 40, 719–731. <https://doi.org/10.1016/j.pragma.2007.11.004>
- Gerner, M. and Bisang, W. (2009). Inflectional classifiers in Weining Ahmao: Mirror of the history of a people. *Folia Linguistica Historica* 30, 183–218.
- Güldemann, T. (2007). Bantu in its macro-areal context of Africa and implications for the early typology of Bantu and Niger-Congo. Paper read at the International Conference on Bantu Languages, Gothenburg: Oct 4–6, 2007.

- Güldemann, T. (2010). Proto-Bantu and Proto-Niger-Congo: Macro-areal typology and linguistic reconstruction. In: C. König and O. Hieda (eds), *International Symposium of the Center of Corpus-Based Linguistics and Language Education (CbLLE)*. Amsterdam and Philadelphia: John Benjamins.
- Haiman, J. (1983). Iconic and economic motivation. *Language* 59: 781–819.
<https://doi.org/10.2307/413373>
- Huang, C.-T.J. (1984). On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15, 531–574.
- Huang, Y. (1994). *The Syntax and Pragmatics of Anaphora. A Study with Special Reference to Anaphora*. Cambridge: Cambridge University Press.
<https://doi.org/10.1017/CBO9780511554292>
- Hyman, L. M. (2004). How to become a Kwa verb. *Journal of West African Languages* 30, 69–88.
- Hyman, L. M. (2010). The Macro-Sudan Belt and Niger-Congo reconstruction. September 04, 2011: linguistics.berkeley.edu/~hyman/Hyman_Niger_Congo_Dynamics.pdf
- Jaeggli, O. and Safir, K. (1989). The null subject parameter in language acquisition. In: O. Jaeggli and K. Safir (eds.), *The Null Subject Parameter* (pp. 215–238). Dordrecht: Foris.
https://doi.org/10.1007/978-94-009-2540-3_1
- Jenner, Philipp N. and Pou, Saveros. (1982). *A Lexicon of Khmer Morphology*. [Mon-Khmer Studies IX-X]. Honolulu: University of Hawaii Press.
- Kager, R. (1999). *Optimality Theory*. Cambridge: Cambridge University Press.
<https://doi.org/10.1017/CBO9780511812408>
- Kibrik, A. A. (2011). *Reference in Discourse*. Oxford: Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780199215805.001.0001>
- Kruspe, N. (2004). *A Grammar of Semelai*. Cambridge: Cambridge University Press.
<https://doi.org/10.1017/CBO9780511550713>
- Kusters, W. (2003). *Linguistic Complexity: The Influence of Social Change on Verbal Inflection*. Utrecht: Landelijke Onderzoekschool Taalwetenschap (Netherlands Graduate School of Linguistics).
- Lefebvre, C. and Brousseau, A.-M. (2002). *A Grammar of Fongbe*. Berlin: Mouton de Gruyter.
<https://doi.org/10.1515/9783110880182>
- Levinson, S. C. (2000). *Presumptive Meanings. The Theory of Generalized Conversational Implicatures*, Cambridge, Mass.: The MIT Press.
- Li, F.-K. (1977). *A Handbook of Comparative Tai*. Honolulu: University Press of Hawaii.
- Li, W. (2014). Perfectivity and grounding in Chinese. *Studies in Language* 38.1, 127–168.
<https://doi.org/10.1075/sl.38.1.o4li>
- Li, X. and Bisang, W. 2012. Classifiers in Sinitic languages: From individuation to definiteness-marking. *Lingua* 122, 335–355. <https://doi.org/10.1016/j.lingua.2011.12.002>
- Luo, Y. (1997). *The Subgroup Structure of the Tai Languages: A Historical-Comparative Study*. Berkeley: University of California. [Journal of Chinese Linguistics Monograph Series 12].
- Masica, C. P. (1990). *The Indo-Aryan Languages*. Cambridge: Cambridge University Press.
- Matson, D. (1964). A Grammatical Sketch of Juang. PhD dissertation, University of Wisconsin.
- McConvell, P. (2008). Mixed languages as outcomes of code-switching: Recent examples from Australia and their implications. *Journal of Language Contact* 2, 187–212.
<https://doi.org/10.1163/000000008792525327>
- McWhorter, J. H. (2001). The world's simplest grammars are creole grammars. *Linguistic Typology* 5, 125–166. <https://doi.org/10.1515/lity.2001.001>

- McWhorter, J. H. (2005). *Defining Creole*. Oxford: Oxford University Press.
- McWhorter, J. H. (2016). Is radical analyticity normal? Implications of Niger-Congo and Southeast Asia for typology and diachronic theory. In: Van Gelderen, Elly (ed.), *Cyclical Change Continued* (pp. 49–91). Amsterdam and Philadelphia: John Benjamins.
<https://doi.org/10.1075/la.227.03mcw>
- Michaelis, S. M., Maurer, P., Haspelmath, M. and Huber, M. (eds.) (2013). *Atlas of Pidgin and Creole Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://apics-online.info>, Accessed on 2016-11-05.)
- Nagaraja, K. S. (2016). Standard Khasi. In: M. Jenny and P. Sidwell (eds.), *The Handbook of Austronesian Languages*, Vol. 2 (pp. 1145–1185). Leiden and Boston: Brill.
- Neeleman, A. and Szendro, K. (2007). Radical pro drop and the morphology of pronouns. *Linguistic Inquiry* 38, 671–714. <https://doi.org/10.1162/ling.2007.38.4.671>
- Pinnow, H.-J. (1960). Über den Ursprung der voneinander abweichenden Strukturen der Munda- und Khmer-Nikobar Sprachen. *Indo-Iranian Journal*, 4.1.
- Ratliff, M. (2010). *Hmong-Mien Language History*. Canberra: Pacific Linguistics.
- Rizzi, L. (1986). Null objects in Italian and the theory of pro. *Linguistic Inquiry* 17, 501–557.
- Ross, M. (2002). The history and transitivity of western Austronesian voice and voice-marking. In: F. Wouk and M. Ross (eds.), *The History and Typology of Western Austronesian Voice Systems* (pp. 17–62). Canberra: The Australian National University.
- Sagart, L. (1999). *The Roots of Old Chinese*. Amsterdam and Philadelphia: Benjamins.
<https://doi.org/10.1075/cilt.184>
- Scott, J. (2009). *The Art of Not Being Governed*. New Haven: Yale University Press.
- Sidwell, P. (2008). Issues in the morphological reconstruction of Proto-Mon-Khmer. In: *Morphology and Language History. In Honour of Harold Koch* (pp. 251–265). Amsterdam and Philadelphia: Benjamins. <https://doi.org/10.1075/cilt.298.22sid>
- Sidwell, P. (2009). *Classifying the Austroasiatic Languages: History and State of the Art*. München: LINCOM EUROPA.
- Sidwell, P. (2016). Austroasiatic classification. In: M. Jenny and P. Sidwell (eds.), *The Handbook of Austronesian Languages*, Vol. 1 (pp. 144–220). Leiden and Boston: Brill.
- Sidwell, P. and Blench, R. (2011). The Austroasiatic Urheimat: the Southeastern Riverine Hypothesis. In: N. Enfield (ed.), *Dynamics of Human Diversity*. Canberra: Pacific Linguistics (pp. 315–344).
- Sinnemäki, K. (2011). Language universals and linguistic complexity. PhD dissertation at the University of Helsinki, Department of Modern Languages.
- Song, J. J. (1997). On the development of MANNER from GIVE. In: J. Newman (ed.), *The Linguistics of Giving* (pp. 327–348). Amsterdam and Philadelphia: John Benjamins.
- Speas, M. (1994). Null arguments in a theory of economy of projections. *University of Massachusetts Occasional Papers in Linguistics* 17, 179–208.
- Syca, A. (1993). Null subject in Mauritian Creole and the Pro-Drop Parameter. In: Byrne, Francis and Holm, John. (eds.), *Atlantic Meets Pacific* (pp. 91–102). Amsterdam and Philadelphia: John Benjamins.
- Thepkanjana, K. and Uehara, S. (2008). The verb of giving in Thai and Mandarin Chinese as a case of polysemy: A comparative study. *Language Science* 30, 621–651.
<https://doi.org/10.1016/j.langsci.2007.04.001>
- Trudgill, P. (2011). *Sociolinguistic Typology: Social Determinants of Linguistic Structure and Complexity*. Oxford: Oxford University Press.

- Tuc, H.-D. (2003). *Vietnamese-English Bilingualism. Patterns of Code-Switching*. London and New York: Routledge Curzon.
- Van Driem, G. (2001). *Languages of the Himalayas*. Vol. 1. Leiden: Brill.
- Wang, Fushi (王辅世). (1957). 贵州威宁苗语量词 Guizhou Weining Miaoyu liangci [The classifier in the Weining dialect of the Miao language in Guizhou]. *Yuyan Yanjiu* 1957: 75–121.
- Xing, J. Z. (2013). Semantic reanalysis in grammaticalization in Chinese. In: Zh. Jing-Schmidt (ed.), *Increased Empiricism: New Advances in Chinese Linguistics* (pp. 223–246). Amsterdam and Philadelphia: John Benjamins. <https://doi.org/10.1075/scld.2.11xin>
- Xing, J. Z. (2015). A comparative study of semantic change in grammaticalization and lexicalization in Chinese and Germanic languages. *Studies in Language* 39.3, 594–634. <https://doi.org/10.1075/sl.39.3.03xin>
- Xu, D. (2006). *Typological Change in Chinese Syntax*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199297566.001.0001>
- Zide, N. H. and Anderson, G. D. S. (2001). The Proto-Munda verb system and some connections with Mon-Khmer. In: B. Rao and K. V. Subbarao (eds.), *Yearbook of South Asian Linguistics* (pp. 517–540). Berlin: Mouton de Gruyter.

Address for correspondence

Walter Bisang
 Johannes Gutenberg-University
 FB 05, Department of English and Linguistics
 Jakob-Welder-Weg 18
 D 55128, Mainz
 Germany
wbisang@uni-mainz.de