The areal typology of grammaticalization: the case of northern China

Giorgio Francesco Arcodia
Ca' Foscari University of Venice
giorgio.arcodia@unive.it

Abstract
Chinese/Sinitic is often seen as a textbook example of isolating typology, with little or no inflection, stable morpheme boundaries, no cumulative exponence, and no allomorphy or suppletion. From the diachronic point of view, the isolating nature of Chinese, as well as other typological features (e.g. lack of obligatory categories), are said to be associated with grammaticalization without formal evolution (see e.g. Bybee, Perkins and Pagliuca 1994; Bisang 2004). In this paper, we will discuss the typology of Sinitic in its genetic and areal context. We will then focus on how grammaticalization works in languages of the East and Mainland Southeast Asian area (EMSEA), and we shall discuss possible exceptions to this general trend in some Northern Sinitic languages. We will show that the typological features traditionally attributed to EMSEA languages do seem to prevent the establishment of morphological paradigms, but secondary grammaticalization (in the sense of Traugott 2002) may still occur, as a morphophonological phenomenon connected with frequency of cooccurrence and with specific prosodic patterns. We shall also discuss the implication of this for the typology of Sinitic, and for grammar-based cross-linguistic research.

1. Introduction[1]

The term 'Chinese' is mostly used in English to refer to the official language of the People's Republic of China, i.e. Modern Standard (Mandarin) Chinese (henceforth: MSC). MSC is by far the best described language of China, and much typological

[1] The glosses follow the general guidelines of the *Leipzig Glossing Rules*. Additional glosses include: ADD = additive; CONT = continuous aspect; COS = change of state; DIR = directional; EXP = experiential; FRUS = frustrative; GOAL = goal marker; INCP = incipient; INV = inverse; LTR = low transivity; QTAG = question tag; SPON = spontaneous; TENT = tentative. Simplified Chinese characters and the *Hanyu Pinyin* transcription have been used as a default for Modern Standard Chinese; for all other varieties, we use the transcriptions provided by the sources. When no transcription is available, we use smallcaps toneless *Pinyin* following the Modern Standard Chinese pronunciation.

research refers mainly or only to MSC data for Sinitic.[2] However, Sinitic is a highly diverse major branch of the Sino-Tibetan family, with as many as 7 (or 10, according to another classification; see Li 1985) sub-branches, each of which contains hundreds of languages, i.e. the so-called Chinese 'dialects': as pointed out e.g. by Chappell (2006), from the typological point of view, MSC is not always representative of Sinitic as a whole. This seems to be true also for the features of grammaticalization.

An often-quoted textbook definition of grammaticalization is "that part of the study of language change that is concerned with such questions as how lexical items and constructions come in certain linguistic contexts to serve grammatical functions or how grammatical items develop new grammatical functions" (Hopper & Traugott 2003: 1). This simple definition focusses on the semantic/functional side of grammaticalization, i.e. the development of grammatical functions. However, a recurring argument in grammaticalization studies is that formal evolution goes hand in hand with the semantic/functional evolution of a linguistic sign: see, for instance, the notion of "cline of grammaticalization" (content item > grammatical word > clitic > inflectional affix) in Hopper & Traugott (2003: 6), or Lehmann's (2015) notion of "autonomy" which, in his view, decreases with grammaticalization. However, the two aspects of grammaticalization, i.e. the semantic/function and the formal one, should be kept separate: this is apparent in Traugott's (2002) distinction between 'primary grammaticalization' (i.e. the development of functional meaning) and 'secondary grammaticalization' (i.e. the development of morphological bonding, phonetic erosion, etc.). This distinction is particularly relevant for Sinitic: in a number of papers, Bisang (1996, 2004, 2008; see also Ansaldo, Bisang and Szeto 2018) argues that, in languages of the East and Mainland Southeast Asian (EMSEA) *Sprachbund*, including Chinese, primary grammaticalization mostly does not involve secondary grammaticalization.

In this paper, we will show that the above-mentioned typological constraints on grammaticalization do not apply in the same way to the Sinitic family as a whole. Just as for several other phenomena (see e.g. Chappell 2015b for some examples), there appear to be exceptions to these general trends, especially in Northern Sinitic.[3] Particularly, in languages of the (broadly understood) Central Plain (中原 *Zhōngyuán*) area, we see several instances of secondary grammaticalization (Arcodia 2013, 2015; Lamarre 2015), while in Northern Sinitic languages of the Qinghai-Gansu (or Amdo) *Sprachbund* we see the development of Altaic-type agglutinative morphology. We shall also discuss the implication of this for the typology of Sinitic, and for grammar-based cross-linguistic research.

This paper is structured as follows. Firstly, we shall briefly present the main typological features of Sinitic in its genetic and areal context, highlighting the role of contact in shaping the different profiles which are found within this language group (section 2). Secondly, we shall devote some space to a general overview of the salient areal features of the East and Mainland Southeast Asian (EMSEA) *Sprachbund*, and we

[2] See, for instance, the *World Atlas of Language Structures* (https://wals.info/; accessed 12/2/2020): only ten Sinitic languages have been considered, and only five of those are included in maps concerning grammatical topics. The remaining five Sinitic languages have been considered only for phonological (or lexical) features (Arcodia 2015).

[3] Here we use 'Northern Sinitic' to broadly indicate Mandarin and Jin dialects. However, Mandarin dialects are actually spoken in Central and Southwestern China, and it has been pointed out that there is a rather strong divide between 'Northern Mandarin' and 'Southern Mandarin' (i.e. the Jiang-Huai and Southwestern Mandarin subgroups), and that Southern Mandarin dialects tend to be more diverse than Northern Mandarin dialects (arguably due to contact with non-Mandarin dialects; Szeto, Ansaldo & Matthews 2018).

shall illustrate how grammaticalization works in the languages of this area (section 3). We shall then move to the discussion of some data from Northern Sinitic languages in which grammaticalization (and, more generally, the expression of grammatical meaning) does not seem to conform to the EMSEA model (section 4). Lastly, we summarize the main points of this paper, and we propose some concluding remarks and hints for further research (section 5).

2. Chinese/Sinitic in its genetic and areal context

As mentioned in the Introduction, the term 'Chinese' is commonly found in non-specialistic language to refer to a single, standardized Sinitic variety, namely MSC, although the term in itself could refer just to any language belonging to the Sinitic branch of the Sino-Tibetan family (see Norman 2003). In one of the most influential works in the history of Chinese linguistics, namely Chao Yuen Ren's *A Grammar of Spoken Chinese*, we read that Chinese 'dialects' all share a "universal Chinese grammar" (Chao 1968: 13): differences among varieties would thus be mostly limited to phonology and the lexicon. However, research on the grammar of Sinitic languages other than MSC has revealed that they do vary considerably at all levels, just as e.g. any major branch of the Indo-European family, often with (very) limited mutual intelligibility (Norman 2003; Chappell 2015a). In fact, 'Sinitic languages' is often used instead of 'Chinese dialects' in works written in European languages, to emphasize the fact that they should be regarded as sister languages of MSC, rather than as simple regional varieties. Nevertheless, the use of the term 'dialect' (the norm in the linguistic literature written in Chinese) is appropriate from the sociolinguistic point of view: Sinitic languages other than MSC have no official recognition, they are not standardized, and no standardized orthography exists for them (with limited exceptions). Thus, from the point of view of their status, they are indeed 'dialects' (as opposed to MSC), in the sense that they lack most of the features usually attributed to a standard language (as e.g. codification, a written standard, etc.): we may also refer to them as 'non-standardized' Sinitic languages.

Within the Sino-Tibetan family, Sinitic languages stand out for being morphologically 'simpler' (for lack of a better term) than many other languages in this genetic grouping, and for having verb-medial order, as opposed to verb-final for most other Sino-Tibetan languages. Sinitic languages are often seen as prototypically isolating, having little or no inflection, stable morpheme boundaries, no cumulative exponence, and no allomorphy or suppletion (see Packard 2006). While e.g. Karen and Lizu share the isolating typology of Sinitic, agglutinative and even polysynthetic morphology is easily found in Tibeto-Burman languages.[4] There are two main typological patterns for verbal morphology within Tibeto-Burman (DeLancey 2015):

a. transparent and regular agglutinative morphology (e.g. Lolo-Burmese, Tibetic, and Boro-Garo)

[4] The validity of 'Tibeto-Burman' as a homogeneous subgrouping of Sino-Tibetan is indeed questionable. Nevertheless, the use of this term to refer to non-Sinitic Sino-Tibetan languages is indeed extremely common in the literature. For the sake of simplicity, here we choose to follow the dominant practice, and consistently use 'Tibeto-Burman' for all non-Sinitic Sino-Tibetan languages, without any implication as to their status as a branch of the family (see Jacques 2017).

b. complex verb morphology, with elaborate argument indexation and transitivity management systems; this includes conservative languages with substantial archaic morphology (e.g. Rgyalrongic and Kiranti), and languages with innovative paradigms (e.g. Kuki-Chin)

Thus, in the Sino-Tibetan family one may indeed find complex verb morphology, as in the two following examples:

(1) Minyong (Post & Sun 2017: 330)
ami *ə=kom* *gok-ta-ki-ram-hi-kaa-to=î.*
person IND=ADD call-INCP-TENT-FRUS-REFL-EXP-PFV=QTAG
'The guy also tried in vain to have a go at calling, eh.'

(2) Rgyalrong (Sun 2014: 634)
tʰɐ-kə-o-nə-ɟɐ-sɐ-ʁⁿduʔ *tɐlŋaʔ=nəʔ*
CONT:LTR-NMLZ:SBJ-INV-SPON-REFL-CAUS-beat child=DET
'The child who is getting himself beaten'

In Tibeto-Burman, complexity is seen also at the paradigmatic level, especially in Rgyalrongic and Kiranti languages. See, for instance, the transitive verbal paradigm of Camling, a Kiranti language (adapted from DeLancey 2015: 67; Σ = verb stem):

<table>
<tr><td colspan="2">O
A</td><td colspan="3">1</td><td colspan="2">INCL</td><td colspan="3">2</td><td colspan="2">3</td></tr>
<tr><td colspan="2"></td><td>SG</td><td>DU</td><td>PL</td><td>DU</td><td>PL</td><td>SG</td><td>DU</td><td>PL</td><td>SG</td><td>NSG</td></tr>
<tr><td rowspan="3">1</td><td>SG</td><td colspan="5" rowspan="4"></td><td rowspan="3">Σ-na</td><td rowspan="3">Σ-na-ci</td><td rowspan="3">Σ-na-ni</td><td>Σ-uŋa</td><td>Σ-uŋ-c-uŋa</td></tr>
<tr><td>DU</td><td colspan="2">Σ-c-ka</td></tr>
<tr><td>PL</td><td>Σ-um-ka</td><td>Σ-um-c-um-ka</td></tr>
<tr><td rowspan="2">I</td><td>DU</td><td rowspan="2" colspan="3"></td><td colspan="2">Σ-ci</td></tr>
<tr><td>PL</td><td>Σ-um</td><td>Σ-um-c-um</td></tr>
<tr><td rowspan="3">2</td><td>SG</td><td>ta-Σ-uŋa</td><td rowspan="3">ta-Σ-c-ka</td><td rowspan="3">ta-Σ-i-ka</td><td colspan="5" rowspan="3"></td><td>ta-Σ-u</td><td>ta-Σ-u-cy-u</td></tr>
<tr><td>DU</td><td>ta-Σ-ci</td><td colspan="2">ta-Σ-ci</td></tr>
<tr><td>PL</td><td>ta-Σ-i</td><td>ta-Σ-um</td><td>ta-Σ-um-c-u-m</td></tr>
<tr><td rowspan="3">3</td><td>SG</td><td rowspan="3">pa-Σ-uŋa</td><td rowspan="3">pa-Σ-c-ka</td><td rowspan="3">pa-Σ-i-ka</td><td rowspan="3">pa-Σ-ci</td><td rowspan="3">pa-Σ-i</td><td rowspan="3">ta-Σ-a</td><td rowspan="3">ta-Σ-ci</td><td rowspan="3">ta-Σ-i</td><td>Σ-u</td><td>Σ-u-cy-u</td></tr>
<tr><td>DU</td><td colspan="2">pa-Σ-ci</td></tr>
<tr><td>PL</td><td>pa-Σ</td><td>Σ-u-cy-u</td></tr>
</table>

Table 1. Transitive Paradigm in Camling.

As we can see in Table 1, Camling has a very complex system of verb agreement with

both the subject and the object argument by means of prefixes and affixes, which also involves the inclusive/exclusive distinction.

Lastly, morphological complexity is expressed in some Tibeto-Burman languages in systems of stem alternation. Interestingly, while stem alternation is often found in the conservative Rgyalrongic and Kiranti languages, it is attested also as an innovative phenomenon in Kuki-Chin. Thus, for instance, in Hakha Lai the verb 'buy' generally has the stem form I (*tsòo*) in main clauses, and the stem form II (*tsook*) in subordinate clauses; some verbs also have a third stem form (VanBik 2009: 12-16; see the source for an overview of the functional correlates of stem alternation).

What is most interesting about the distribution of morphological complexity within the Sino-Tibetan family is the interaction between genealogy and language contact, i.e. between vertical and horizontal transmission of linguistic features. The type of complex verb morphology which we find e.g. in Rgyalrong and Kiranti languages is claimed to be an archaic feature of the family (Bickel & Nichols 2013). This can be inferred from the fact that the most complex morphology is found in languages spoken in isolated areas: Rgyalrongic in the mountains of Sichuan, Nungic in the most inaccessible mountain valleys of northern Myanmar, and Kiranti and Kham-Magar languages in the mountain valleys of Nepal, i.e. prototypical 'residual zones' (following the definition in Nichols 1992; DeLancey 2015: 63-64). The languages which underwent the most radical simplification of the (supposed) original Sino-Tibetan model, i.e. Sinitic, Tibetic and Burmese, were all used as *lingua francas* of vast empires; other 'simplified' Tibeto-Burman languages, as e.g. Tani, despite having developed in relative isolation, "show apparent evidence of intensive contact and creolization" (DeLancey 2015: 63). However, innovative complexity is actually attested in languages spoken in isolated areas, as the above-mentioned Kuki-Chin languages spoken in remote areas of Myanmar and Northeast India: it thus appears that while simplification is significantly correlated with contact, complexification is largely independent from it. Also, the Kuki-Chin case shows that a trend towards complexification may 'oppose' competing tendencies towards decomplexification (DeLancey 2015: 76).

Thus, Sinitic is arguably the branch of Sino-Tibetan which has diverged most radically from the rest of the family, from the point of view of morphological typology, and this divergence is normally explained as the product of language contact. At least since the epoch of Qin imperial unification (221–207 BCE), there have been repeated waves of migration of Chinese-speaking people to southern China, which was then inhabited by speakers of Hmong-Mien, Tai-Kadai, and Austroasiatic languages, i.e. typical EMSEA languages. The resulting admixture of languages led to convergence between Sinitic and those languages, which are also very good representatives of the isolating morphological type, as we shall see in greater detail in the next section. Perhaps unsurprisingly, the convergence towards the strongly isolating EMSEA type is more advanced in Southern Sinitic languages, while Northern Sinitic languages share more features with languages of northern Asia. Indeed, while Southern Sinitic languages developed in closer contact with EMSEA languages, Sinitic-speaking people in northern China came into contact with speakers of Mongolic, Turkic, and Tungusic languages (LaPolla 2001; Enfield 2005; Ansaldo 2010). Consequently, Hashimoto (1986) speaks of the 'Altaicization' of Northern Chinese, and of the 'Taization' of Sourthern Chinese. In Table 2, we list some of the main differences between Northern and Southern Sinitic (adapted from Chappell 2015b: 17).

Northern Sinitic	Southern Sinitic
Stress-based and fewer tones	More tones
Higher proportion of polysyllabic words	Higher proportion of monosyllabic words
Simpler syllable structure	More complex syllable structure
Smaller inventory of classifiers	Larger inventory of classifiers
Preverbal adverbs	Possibility of postverbal or clause-final adverbs
Marker-standard-adjective order in the comparative construction	Adjective-marker-standard order in the comparative construction

Table 2. Some differences between Northern and Southern Sinitic.

The differences at the phonological and prosodic level between Northern and Southern Sinitic are particularly relevant for secondary grammaticalization. Ansaldo & Lim (2004) point out that in Southern varieties such as Cantonese and Hokkien, grammaticalized items may show signs of phonetic erosion, which is expressed in terms of shorter duration and changes in vowel quality: however, stronger reduction does not seem to occur, due to due the discreteness of syllable boundaries. This is also related to the fact that Cantonese and Hokkien are syllable-timed languages which have more than one tonal register, and they lack the neutral tone option: hence, a reduction in pitch height may be misinterpreted as, for instance, a mid level tone becoming a low tone, rather than as a sign of erosion. However, SMC and, generally speaking, Northern Chinese dialects are dominated by stress, a feature which might derive from contact with 'Altaic' and Tibetic languages (Lamarre 2015), and have neutral tone (weakly stressed) syllables (Ansaldo & Lim 2004). Indeed, in Northern Sinitic languages some grammatical morphemes lost their tone values, and even underwent segmental reduction to some extent, as e.g. the MSC aspect markers 了 *-le* (perfective, < 了 *liǎo* 'finish'), 着 *-zhe* (durative, < 着 *zhuó* 'touch'), and 过 *-guo* (experiential, < 过 *guò* 'pass, cross'; see Bisang 2008).

Interestingly, these neutral tone grammatical morphemes all follow the lexical morpheme they are related to (for aspect markers, a verb): this reduction happens in a specific prosodic environment, namely a trochaic (strong-weak) stress pattern, in which grammatical morphemes are thus in a prosodically weak position. This pattern tends to favour the cliticisation of post-head elements, which arguably led to phonetic erosion in the case of the MSC aspect markers seen above (and in some other cases, e.g. for some postpositions): the steps in this process of formal evolution become more evident if we compare related dialects (see Jiang 1999). Also, note that neutral tone syllables "have variable realizations, i.e. they can coarticulate with adjacent tones as much as they want": their realization is hence context-dependent, as "different neutralized morphemes tend to be realized similarly in the same context" (Ansaldo & Lim 2004: 347). This paves the way for allomorphy, and further formal evolution of these grammaticalized items, as we shall see below (Section 4).

Thus, the differences between Northern and Southern Chinese (as a whole) may (partly) explain why grammaticalization has different formal correlates in languages belonging to these two major groupings of Sinitic. However, as mentioned above, strong reduction of grammaticalized signs is seen only in a subset of Northern Sinitic languages, spread over an area whose centre lies in the Central Plain: indeed, in most Northern Chinese dialects, as well as MSC, erosion of grammaticalized items (as the aspect

markers seen above) is the exception, rather than the norm. Besides, an entirely different morphological profile is found in the Northern Sinitic languages of the Qinghai-Gansu *Sprachbund*, which developed agglutinative, suffixal grammatical morphology marking categories proper of Tibetic and Mongolic languages (as e.g. case; see e.g. Sandman 2016): this is widely recognized to be a contact-induced pattern of areal convergence (Slater 2003). Due to space constraints, we shall discuss in detail only data of reduced morphology in languages of the Central Plain area: the reader is referred e.g. to Xu (2017) for an overview of the Qinghai-Gansu *Sprachbund*.

Let us now move to a presentation of the features of EMSEA languages which are most relevant for grammaticalization, before discussing the Northern Sinitic data alluded to above.

3. EMSEA languages and grammaticalization

The EMSEA area is a typical example of a *Sprachbund*, being a region in which languages from different families developed shared features due to a history of contact (Matisoff 2001; Enfield 2005; Goddard 2005). In the narrow definition, EMSEA is defined as "the region encompassing Vietnam, Laos, Cambodia, and Thailand, with some extension west into Burma, south into Peninsular Malaysia, and north into southern China" (Enfield 2005: 182): northern China would thus be excluded from this area. However, a broader definition, including the whole of China (thus, also Northern Sinitic), is often (implicitly or declaredly) used in typological works: crucially, this is the case for Bisang's work on grammaticalization in EMESA languages (see e.g. Bisang 2004).[5] In the broad definition, the EMSEA area includes languages belonging to the Sino-Tibetan, Tai-Kadai, Hmong-Mien, Austroasiatic (specifically, Mon-Khmer), and Austronesian (Cham) families (Goddard 2005). In Table 3, we list some properties which are often attributed to EMSEA languages (see Matisoff 2001; Enfield 2005; Goddard 2005; Ansaldo 2010).

Tendency towards monosyllabism
Isolating/analytic morphology
Lack of agreement for number, case, etc.
Lack of obligatory arguments (zero anaphora)
Topic-prominent syntax
Use of lexical morphemes with grammatical functions
Use of serial verb constructions
Verb-medial, head-modifier order, use of prepositions
Use of lexical (and grammatical) tone
Use of (modal) sentence-final particles
Use of classifiers
Prominence of aspect over tense
Rich vowel inventories

Table 3. Some salient areal features of EMSEA languages.

[5] In Bisang (2008: 31), it is stated the EMSEA areal type of grammaticalization encompasses Mon-Khmer, Tai (rather than the whole of Tai-Kadai), Sinitic, and Hmong-Mien.

Several among the features presented in Table 3 are relevant for grammaticalization. First and foremost, the combination of isolating/analytic morphology, lack of agreement, and zero anaphora, as well as topic-prominent syntax (which entails that word order relies heavily on information structure, rather than on subjecthood and/or agentivity), results in one of the most prominent characteristics of EMSEA languages, namely 'indeterminateness' (Bisang 2004).

Indeterminatedness, or the lack of obligatory categories, entails that "arguments can be omitted without concomitant agreement morphology on the verb" (known as 'radical pro-drop'; Bisang 2015: 135). This is nicely summarized by Enfield (2005: 188) as follows:

> In no MSEA language are clausal heads or dependents morphologically marked for argument structure relations − i.e., there is neither case-marking nor agreement. Although it is often presumed that in isolating languages the functions of such morphological marking are performed by constituent order, there is considerable within-language constituent order variability. The typical MSEA language combines widespread noun phrase ellipsis (of definite arguments) with noun phrase movement (into clause-external positions like topic), resulting in great indeterminacy of surface sequences.

Also, even grammaticalized categories are rarely (if ever) obligatorily expressed. Bisang (2004: 111-112; no Chinese characters in the source) proposes the following two MSC examples to illustrate the notion of indeterminateness:

(3) (*tā*) *lái*
3SG come
'(S/he) comes / has come / is coming / will come / etc.'

(4) *tā* *mǎi* *bào*
3SG buy newspaper
'S/he bought a newspaper / newspapers / the newspaper / the newspapers'

As pointed out by Bisang (2004: 111), these sentences "are perfectly acceptable in a context in which no particular information beyond the concept denoted by the verb [(3)] or the noun [(4)] is needed". This leads to the situation described by Enfield (2001: 259), namely that "[n]ormal utterances are often impossible to interpret properly outside the contexts in which they actually occur". According to Bisang (2004), indeterminateness is one of the main reasons for the peculiar features of grammaticalization in EMSEA languages, together with the weak correlation between the lexicon and morphosyntax. Here, 'weak correlation of lexicon and morphosyntax' means that the association between word classes (noun, verb, etc.) and syntactic slots is not as rigid as e.g. in the Indo-European languages of Europe. This, we may add, is related to another feature attributed to EMSEA languages: namely, the use of lexical morphemes with grammatical functions (see Table 3).

How does this translate into primary grammaticalization without secondary grammaticalization? One characteristic of EMSEA languages is that processes of grammaticalization do not follow (unidirectional) clines (as the above-mentioned cline content item > grammatical word > clitic > inflectional affix; see Section 1): instead of a

step-by-step evolution from ‘less grammatical’ to ‘more grammatical’, or from ‘less bound’ to ‘more bound’, grammaticalized items often retain different interpretations, which are all accessible. While the role of inference is often acknowledged in grammaticalization studies, it is said to be mostly involved in the early stages of the process of grammaticalization, with inferential freedom decreasing with increased grammaticalization (Bybee, Perkins & Pagliuca 1994; Hopper & Traugott 2003; see Bisang 2004: 116; Bisang 2008: 29-30). In EMSEA languages, on the other hand, “[o]ne can see the step from lexical item to grammaticalized item but it is often hard to clearly distinguish between more and less grammaticalized items” (Bisang 2008: 23), and “one and the same marker may express different grammatical concepts in different situations or in different constructions” (Bisang 2008:16). Word order is the main indicator of grammaticalization (as e.g. when a verb occupies the syntactic slot of an adposition; see below, Ex. 7), but even the very same surface word order may be open to different interpretation through pragmatic inference, as e.g. the Khmer verb *ʔaoy* ‘give’ in the following example (Bisang 2015: 139):

(5) *ʔo:pùk sɔŋ phtɛ̀əh ʔaoy ko:n nɤ̀u*
father build house give child live/stay
a. Father builds a house for his children to live in (*ʔaoy* = coverb)
b. Father builds a house for making his children to live there (*ʔaoy* = causative verb)
c. Father builds a house with the purpose that his children live there (*ʔaoy* = adverbial subordinator)

The verb *ʔaoy* ‘give’ grammaticalized into a coverb that marks benefactives (a.), into a causative verb (b.), and into an adverbial subordinator (purpose or manner; c.), as well as into a complementizer. While the different ‘identities’ of *ʔaoy* are associated with different constructions, there are indeed cases in which more than one interpretation is available, as shown in (5).

The above-mentioned weak correlation between lexicon and morphosyntax is obviously involved here, as it permits the use of a single item in different syntactic environments (or construction; Bisang 2004: 116-117). This polysemy, again, is connected with the indeterminateness of EMSEA languages, and with the lack of coevolution of meaning and form, as pointed out in Bisang (1996: 535):

> [i]n a language in which almost every grammatical category almost always can be inferred from the context, i.e., in a language where there is almost no obligatory grammatical category, even a highly grammaticalized linguistic item shows a higher degree of informative value than in a language showing a lower degree of indeterminateness. This higher degree of informative value is reflected by the fundamental phonological stability of a linguistic sign even in a context of high grammaticalization.

A very good example of the above-described characteristics of grammaticalized signs in EMSEA languages is the MSC item 在 *zài*. The meaning of 在 *zài* as an open-class lexical item is ‘be at, be located’ (6); however, 在 *zài* is used also as a locative adposition (7) and as an aspect (progressive) marker (8; Bisang 2004: 117):

(6) *tā zài túshūguǎn*
3SG be.at library

‘S/he is at the library’

(7) *tā zài yīyuàn sǐ-le*
3SG at hospital die-PFV
‘S/he died at the hospital’

(8) *tā zài chuān pí-xié*
3SG PROG put.on leather-shoe
‘S/he is putting on her/his leather shoes’ (qtd. from Li & Thompson 1981: 221)

There seem to be no significant formal differences between these three ‘identities’ of 在 *zài* (but cf. Ansaldo & Lim 2004: 346), i.e. no secondary grammaticalization: each identity can (and must) be recovered through pragmatic inference, just as for (5) above. The above-mentioned ‘flexibility’ in the use of items in different constructions “supports the reanalysis of morphemes in different functions, and thus enhances the probability of processes of grammaticalization to take place” (Bisang 2004: 117).

Lastly, the lack of obligatory categories and the polysemy of grammaticalized items are closely related to another factor which motivates the lack of coevolution of meaning and form in grammaticalization, and inhibits the rise of morphological paradigms: namely, the relatively low frequency of grammatical markers, if compared to languages with obligatory grammatical categories. In Bisang’s words (2008: 33),

> [...] morphological paradigms develop from categories which are frequently used. Frequency, in turn, is enhanced by semantic generality, which grants its compatibility with a wide range of lexical items. If a marker is semantically general enough to be coextensive with a basic grammatical entity, like noun or verb, its occurrence may become obligatory with that entity. As a consequence, it becomes even more frequent. [...]
> In East and Mainland Southeast Asian languages, the emergence of a situation in which grammatical markers are frequent and homogeneous enough to become part of a coherent paradigm is systematically undermined by the high degree of indeterminateness [...] and the broad functional spectrum of markers. Thus, on the one hand, despite their highly generalized meanings, grammatical markers are not so frequent as, for example, tense markers in English or German, because they are optional. On the other hand, the meaning of grammatical markers depends on pragmatics at all levels of grammaticalization, and therefore their functional range is not homogeneous, not limited to a single clearly determined semantic domain. Thus, the emergence of a paradigm is rather unlikely for both reasons, low frequency and low degree of semantic homogeneity.

The above-mentioned MSC perfective marker 了 *-le* is a case in point (Bisang 2004). While it is normally described as an aspect marker, it has more ‘fuzzy’ semantics: it is said to have “a component of relative past as part of its meaning” (Lin 2006: 19, Fn. 18), and thus is not a ‘pure’ aspect marker. Moreover, 了 *-le* is not actually required in perfective contexts, and its absence does not entail that the predicate is imperfective: in point of fact, there are even typical perfective contexts (as e.g. when “one action is correlated to another ongoing action”; Bisang 2004: 128) in which the use of 了 *-le* is unacceptable.

To sum up, according to Bisang’s account, in EMSEA languages grammaticalization

does take place, but it is characterized by some peculiar features which are connected with the EMSEA typology sketeched above. Specifically, the lack of obligatory categories, the polysemy of grammaticalized items, the predominance of pragmatic inference, and the relatively weak correlation between lexicon and morphosyntax, as well as other factors which we did not mention here due to space constraints (the reader is referred to Bisang 1996 for an extensive discussion), lead to primary grammaticalization without secondary grammaticalization, and inhibit the rise of morphological paradigms. In the next section, we shall present some possible counterexamples to these generalizations.

4. A possible exception: grammaticalization in Northern Sinitic

As mentioned earlier (Section 2 and 3), the development of phonetic erosion in grammaticalization is uncommon in Chinese. While the general typological constraints on grammaticalization are expected to apply more or less equally to Sinitic as a whole, prosodic differences between Northern and Southern Sinitic suggest that secondary grammaticalization is more likely to occur in the former, rather than in the latter. However, even though some degree of morphological bonding and phonetic erosion (especially, loss of tone and coarticulation) is indeed visible in grammatical markers in MSC and other Northern Chinese dialects, this is limited to very few items, and morpheme boundaries are normally well preserved. In other words, while there are some preconditions for secondary grammaticalization in Northern Chinese which are missing in Southern Chinese, their impact is very modest.

The most significant counterexamples to this generalization may be found in an area spanning over Henan, Hebei, Shaanxi, Shanxi, and Shandong, i.e. in and around the historical Central Plain region (see above, Section 1). In this area we find dialects belonging to the Northern Sinitic Jin and Mandarin groups (specifically, to the Central Plain Mandarin, Ji-Lu Mandarin, and Jiaoliao Mandarin subgroups; see Arcodia 2015) in which morphological bonding and phonetic erosion does occur in grammaticalization to a much higher degree than most other regions of China, even those in which dialects belonging to the same subbranches of Mandarin are spoken.[6] In the languages of this area, secondary grammaticalization may be expressed by strong reduction and allomorphy of suffixes (9b), tone change (with or without vowel lengthening) (9c, 10b), rhotacization (11b), and ablaut (in the sense of Bickel & Nichols 2007), also known in Chinese linguistics as 'rhyme change' (变韵 *biànyùn*; 12b; see Lamarre 2009, 2015; Arcodia 2013, 2015):

(9) Boshan (adapted from Qian 1993: 18)

a. 吃了饭，出了门，来了客

tʂʻʅ^{21}- *liɔ*	*fɑ̃*^{31}	*tʂʻu*^{214}- *liɔ*	*mə̃*^{55}	*lɛ*^{55}-*liɔ*	*kʻə*^{214}
eat-PFV	food	exit-PFV	door	come-PFV	guest

b. 吃 ə 饭，出 ə 门，来 ə 客

[6] While phenomena of strong reduction of grammaticalized signs have been reported in the Chinese dialectological literature since the late '50s of the 20th century, the very few systematic typological studies of the phenomenon appeared only relatively recently (Lamarre 2009, 2015; Arcodia 2013, 2015; see Lamarre 2015: 278 for an explanation for this gap in the literature).

tʂʼʅ^{21}*-ɤ* *fɑ̃*^{31} *tʂʼu*^{214}*-ɤ* *mə̃*^{55} *lɛ*^{55}*-ɛ* *kʼə*^{214}
eat-PFV food exit-PFV door come-PFV guest
'(I, she,etc.) ate, went out, and a guest arrived'

c. 换一双鞋
xuɑ̃:^{214} YI SHUANG XIE
change.PFV one pair shoe
'(I, she, etc.) changed a pair of shoes' (Chen 2006: 320)

(10) Nanhe (adapted from Zhang 2011: 20)

a. 她编个篮子
TA *pia*^{44}*-la* GE LANZI
3SG.F weave-PFV CLF basket

b. 她编个篮子
TA *pia:*^{443} GE LANZI
3SG.F weave.PFV CLF basket
'She weaved a basket'

(11) Qixia (Zhang & Li 2007: 98)

a. 我问了老师
WO *uən*^{41}*-lə* LAOSHI
1SG ask-PFV teacher

b. 我问老师
WO *uər*^{41} LAOSHI
1SG ask.PFV teacher
'I asked the teacher'

(12) Xunxian (Xin 2006: 58)

a. 买一斤盐
mai^{55} *i*^{42} *tɕin*^{24} *ian*^{42}
buy one jin salt
'(I, she,etc.) will / am going to buy one *jin* of salt'

b. 买一斤盐
mɛ^{55} *i*^{42} *tɕin*^{24} *ian*^{42}
buy.PFV one jin salt
'(I, she,etc.) bought one *jin* of salt'

If we compare Exx. (9a-11a) with Exx. (9b-11b), we see different exponents for the same grammatical meaning (here, perfective aspect), which differ in terms of bondedness, integration with the root, and degree of erosion. In Boshan, a (Ji-Lu) Mandarin dialect of Shandong, perfective aspect may be expressed by means of the suffix 了 *-liɔ*, an obvious cognate of MSC 了 *-le* (9a), and also by a 'shorter' version, consisting of a single vowel: while we follow the source and indicate it with a schwa, the actual phonetic shape of this single-vowel suffix depends from the shape of the rhyme of the lexical item it attaches to (here, *-ɤ* and *-ɛ*), according to Qian's (1993: 24-25) account. In a more recent description for this dialect (Chen 2006), we read that the suffix may be dropped altogether, and the same grammatical meaning may be expressed by tone change and lengthening of the nucleus vowel (9c). In Nanhe, a Jin dialect of Hebei, the perfective marker cognate to

MSC −了 *le* has the allomorphs −*la*, −*a*, or −*a:* (10a): the addition of this suffix often causes, again, tone change and lengthening in the verb root, and the suffix itself may coarticulate with the root −*a*. The suffix may then be dropped, and the 'burden' of conveying its meaning lies on the modified verb root (10b; Zhang 2011). In Qixia, a (Jiaoliao) Mandarin dialect of Shandong, the perfective aspect suffix 了 *-lə* (11a) may be substituted by 'rhotacization' (known as 儿化 *érhuà* in Chinese linguistics), i.e. the addition (or substitution) of a rhotic coda to the rhyme of the lexical item (11b). Lastly, in Xunxian, a (Central Plain) Mandarin dialect of Henan, we see the expression of perfective aspect by means of segmental ablaut: the base form of the verb 买 *mai*^{55} (12a) is substituted by *mɛ*^{55} (12b). However, differently from the other cases seen here, there is no segmental suffix with the same function: ablaut is the only true marker of perfective aspect (Xin 2006: 168).

Despite the different outcomes, all of the grammatical markers which undergo reduction described here are invariably found in the postverbal position, and they always appear to be the product of the 'integration' of a concatenative exponent in a lexical root (see e.g. Lamarre 2009, 2015). While there are (almost) no diachronic data on these phenomena available, as there is no significant written tradition for the dialects of this area (and, anyway, these phenomena are unlikely to be recorded in writing; see below, Section 5), for some varieties the evidence for a gradual process of reduction may be drawn by comparing competing strategies for marking the same grammatical meaning: Boshan (9a-c) is a case in point. Another very good example comes from the comparison of constructions for marking the attainment of a goal in Jizhou, a Jin dialect of Hebei (Lamarre 2009: 154):

(13) a. 拿唠屋里去
NA^{55}-*laɔ* WU-LI QU
take-GOAL house-inside go
b. 拿唠屋里去
NA^{55}-*ɔ* WU-LI QU
take-GOAL house-inside go
c. 拿屋里去
NA^{55} WU-LI QU
take.GOAL house-inside go
d. *拿屋里去
NA^{53} WU-LI QU
take house-inside go
'Take [it] inside the house'

In Jizhou, the addition of the goal marker 唠 *-laɔ* causes tone change (tone sandhi) in the verb: here, 拿 NA^{53} 'to hold' becomes NA^{55} (13a). The marker 唠 *-laɔ* may be reduced to *-ɔ* (13b), and may even be dropped (13c): similarly to some cases of *Umlaut* in Germanic languages (e.g. Eng. *man* vs. *men*), once the segmental marker disappears, the changed tone becomes the only exponent of the grammatical meaning at issue (here, goal). Indeed, the verb in its basic tone (拿 NA^{53}) cannot be used in this context, as shown by the ungrammaticality of (13d; see Lamarre 2015 for further examples).

Moreover, the history of the evolution of these markers may be gleaned from the comparison of cognate forms in related dialects. Xin (2006: 85) proposes a comparison

among the forms of perfective aspect markers in the above-mentioned Xunxian and in nine other dialects of the same area, reproduced in Table 4:

Anyang	Tangyin	Hebi	Weihuang	Neihuang
lœʔ / *œ* / *nœ* / *lɐn* / *ɐn*	*lɛʔ* / *ɛ* / *nɛ* / *lɐn* / *ɐn*	*ləʔ* / *ə* / *ɐ*	*ləʔ* / ablaut	*o* / *ə* / ablaut
Puyang	Qixian	Xunxian	Huaxian	Yanjin
lə / ablaut	ablaut	ablaut	ablaut	ablaut

Table 4. Markers of perfective aspect in ten dialects of Henan province (Xin 2006: 85).

As we can see in Table 4, markers cognate to MSC 了 *-le* have 'heavier' and 'lighter' (one-vowel) allomorphs in dialects as Anyang, Tangyin, Hebi, and Neihuang; they alternate with ablaut in Weihuang, Neihuang, and Puyang; in Qixian, Xunxian (see 12b), Huaxian, and Yanjin, the suffix disappeared, leaving ablaut as the only exponent of perfective aspect. This may be interpreted as an approximation of how the exponence of perfective aspect likely evolved in these dialects: even when the original segmental marker is no longer present in a dialect, the steps leading to its 'demise' may be inferred from comparative data.

Based on the analysis of a sample of 26 Northern Sinitic languages from the area under consideration here, Arcodia (2013: 154) proposes the following cline of grammaticalization for perfective markers cognate to MSC 了 *-le*, i.e. deriving from the verb 了 *liǎo* 'finish' (see Section 2):

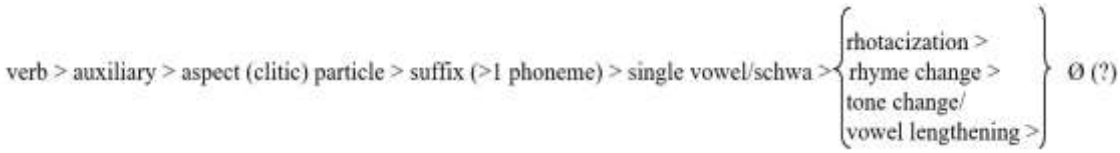

Figure 1. Cline of grammaticalization for perfective markers in Northern Sinitic

This cline is arranged on the basis of the relative 'weight' of the exponent, from heavier to lighter: note also that a crucial step in the process of erosion of grammatical exponents is loss of tone, a precondition for coarticulation and further reduction (see above, Section 2; Ansaldo & Lim 2004; Zhang & Li 2007), which generally occurs at the (clitic) particle stage in Northern Sinitic. The (pre-)terminal stages exemplified above, namely rhotacization, rhyme change, and tone change / vowel lengthening are placed between curly brackets here because, arguably, they "are not steps that all necessarily occur", and while "it does not seem likely that tone change occurs *before* rhyme change in a given language", the existence of rhyme change (i.e. ablaut) "cannot be taken as sufficient evidence for reconstructing an earlier stage at which tone change was used instead" (Arcodia 2013: 155; emphasis in the original). In fact, it might well be the case that ablaut and tone change are generated by different mechanisms. Lamarre (2015: 283) points out that tone change (and/or) vowel lengthening are caused by a common sandhi phenomenon of Northern Chinese, namely "tone change before toneless syllables", the toneless syllable here being the suffix: this is what we saw before for the goal marker in Jizhou (13a-d). In the case of rhyme change, however, the suffix becomes fused with the verb

root (see e.g. Lamarre 2009): tone sandhi is not necessary, as shown above (12a-b).[7] The same goes for rhotacization: while this is also regarded as the product of the influence of a suffix on the verb root, again there generally is no tone change (Zhang & Li 2007: 98), and segmental change affects only or mostly the coda, rather than the whole rhyme of the root.[8]

The final stage in the cline, i.e. zero exponence (loss of exponence), is not attested in Arcodia's sample (hence the question mark), but it has been reported e.g. in Lamarre's (2009) work on goal markers. The cline in Figure 1 actually involves both primary and secondary grammaticalization, but in different ways: while semantic evolution must come into play in the initial stages, in the shift from an open-class item (verb) to a closed-class item (auxiliary, then clitic), the following steps in the cline appear to be independent of functional change, i.e. they are purely a matter of secondary grammaticalization (Arcodia 2013; for some possible counterexamples, see Jiang 1999; Chen 2005, 2007).

Arcodia (2013: 154) argues that the evolution of markers of perfective aspect in the Northern Chinese dialects he considered indeed involves several of the indices of secondary grammaticalization proposed by Bybee, Perkins and Pagliuca (1994: 107-114):

> (a) phonetic reduction (including: loss of stress, reduction to a neutral tone, shortening and reduction of vowels, loss of vowels and/or consonants)
>
> (b) increased dependence (including: development of non-purely phonetically conditioned allomorphs, suprasegmental reduction)
>
> (c) fusion (including: no open class intervening between the gram[9] and the verb, stem-conditioned allomorphy, conditioning of stem allomorphy by the gram)

Specifically, 'stem-conditioned allomorphy' is clearly visible e.g. in the Boshan 'schwa suffix' (9b), while ablaut and tone change may be interpreted as 'conditioning of stem allomorphy by the gram'. As said above, the starting point for reduction is tone neutralization, which may be followed by loss of segments (often, the syllable onset), and also centralization of the nucleus vowels (Li 2002; Zhang & Li 2007). With further integration between the lexical root and the grammatical marker, the above-mentioned stem-conditioned allomorphy, as well as (segmental and/or suprasegmental) changes in the root due to the influence of the marker, may both occur.

Indeed, the pathways of evolution sketched above are hardly unusual, in cross-linguistic perspective. All of the items which undergo reduction here are postverbal, toneless morphemes, with nothing intervening between them and the lexical root they attach to. Bybee (2003: 617) points out that when a word and a morpheme often occur together, they "come to be stored and processed in one chunk": sequences of units which

[7] Note, however, that tone change may be involved in rhyme change too. For instance, Xingyang (a Central Plain dialect of Henan; Wang 1998) has a system of grammatical ablaut very similar to that of Xunxian (12a-b), but in Xingyang some verbs also change their tone, and not only their rhyme. Nevertheless, segmental ablaut is the only systematic alteration which may express grammatical meaning in this variety.

[8] Note also that rhotacization as a morphophonological phenomenon is very common in Northern Sinitic (and beyond), including MSC (Lamarre 2015). Its use as a grammatical exponent, however, appears to be mostly limited to some dialects from the area under consideration here (especially, Shandong).

[9] 'Gram' is used here as the short form of 'grammatical morpheme' which, in Bybee, Perkins and Pagliuca's (1994: 2) understanding, include "affixes, stem changes, reduplication, auxiliaries, particles, or complex constructions such as English *be going to*".

often co-occur may come to be processed as a single unit, their "gestural representation" changes, and the multiple gestures involved in their articulation are reorganized into single gestures, which causes reduction and an "increased overlap of gestures"; namely, coarticulation. As highlighted in Arcodia (2015: 19), "coarticulation may lead both to allomorphy [...], erosion and fusion, and even nonlinear exponence". Also, as mentioned earlier (Section 2), within Northern Sinitic reduction of grammatical morphemes generally occurs in a specific syntactic and prosodic context: namely, within a syntactically or semantically tight phrasal unit, normally in a weak prosodic position, such as next to or between stressed content morphemes. MSC 了 *−le* and its cognates in other Sinitic languages form a trochaic foot with the verb: the aspect marker is thus found in a weaker prosodic position with respect to the root, and becomes tightly associated with it (Jiang 1999; Li 2002). Indeed, in some dialects, ablaut and/or tonal morphology is restricted to single syllable verbs (Liu 2006; Arcodia 2015), arguably because the combination of a disyllabic verb and a suffixed marker may not fit into the prosodic template described above.[10]

However, as discussed in the preceding section, it has been claimed that grammatical morphemes in EMSEA languages, including MSC 了 *−le* (which – we stress again – did undergo erosion to some degree), have comparatively low frequency. This view is challenged e.g. in Arcodia (2013), and more forcefully in Lamarre (2015): both Arcodia and Lamarre provide examples of contexts in which, in some dialects, the so-called 'inflected form' of the verb (i.e the 'changed' form of the verb root after ablaut / tone change / rhotacization; Lamarre 2015: 277) are obligatorily used (see e.g. 13d), just as inflectional exponents, a fact which obviously contradicts Bisang's generalizations outlined above (Section 3). With obligatorification, at least in some contexts, frequency necessarily increases, providing fertile terrain for the processes of coalescence between verb and grammatical marker. Nevertheless, both Arcodia (2013) and Lamarre (2015) acknowledge that the evidence for the obligatory use of reduced morphology is still limited, and that the issue requires further research.

How does this relate to paradigm creation? Having an opposition between 'base' and 'inflected' forms of a verb, especially when the inflected form is obligatory in some contexts, might be argued to be (proto-)paradigmatic, in a sense. In Xi'an (a Central Plain dialect of Shaanxi) we do find what looks like a paradigm of suprasegmental morphology, with 'inflectional classes' associated with different tone contours of the verb. Based on Sun's (2007) description of this dialect, this 'paradigm' may be represented as follows (Arcodia 2015: 18; data from Sun 2007: 190-193; 'VL' = vowel lengthening):

Tone category	Progressive/continuous	Perfective	Goal/degree
31	313 and VL	42	24 or VL
35	242 and VL	242	242 or VL
51	VL	31	VL
55	51	553	53

Table 5. Tonal paradigm in Xi'an.

[10] Note that the issue of stress in disyllabic words is very controversial in Chinese linguistics (see Duanmu 2007, 2014 for an overview). However, there seems to be general agreement on the point that loss of tone and segmental reduction generally occurs on righthand constituents, both in complex words and in (some) phrases: see e.g. MSC 妈妈 *māma* [ma:^{55}.mə] 'mum'.

However, in Sun (2007) we find no clear indication as to the obligatoriness of tone change, which is what we would expect for a morphological paradigm, normally associated with inflectional (obligatory) categories. Again, only a more accurate analysis of obligatorification of reduced morphology could tell us whether a true system of paradigmatic opposition has developed.

Moreover, as said in the preceding section, paradigm formation is not only inhibited by the low frequency of grammatical markers, but also by their "low degree of semantic homogeneity". In many of the cases discussed here, this is exacerbated by the fact that, due to parallel process of reduction of grammatical markers, differently from what we saw in Table 5 for Xi'an, unrelated markers end up being conveyed by exactly the same exponents. For instance, while in (12b) we showed that ablaut is used in Xunxian to convey perfective aspect, the very same patterns of ablaut are also used to convey continuous aspect (14) and the attainment of a goal (15; Xin 2006: 58-59):

(14) 俩人睡一个床
lia^{55} *ʐən*^{42} *ʂɛ*^{213} *i*^{42} *kə* *tʂ'uaŋ*^{42}
two person sleep.CONT one CLF bed
'Two people are sleeping in one bed'

(15) 会改明个了
xuei^{213} *kɛ*^{55} *mɛ*^{42}*kə* *lə*
meeting change.GOAL tomorrow COS
'The meeting has been postponed to tomorrow'

Most often, the meaning conveyed by the inflected form of the verb may be understood from the syntactic context: thus, as predicted by Bisang's model, inference is still necessary. There are also cases in which more than one interpretation is available, as in the following Xingyang (see Fn. 7) example (Wang 1998: 277):

(16) 他背袋儿面
TA *pɛ*^{13} DAIR MIAN
3SG.M carry.on.the.back.PFV/CONT bag flour
a. 'He shouldered a bag of flour on the back'
b. 'He is carrying a bag of flour on the back'

As shown by the translations provided, both a perfective (a.) and a continuous (b.) interpretation are available in this case (but cf. Lamarre 2015: 292).

Thus, on the whole, the data seem to suggest that even though secondary grammaticalization does occur to a significant extent in the Northern Sinitic languages considered here, 'true' paradigms do not really arise: Bisang (2014: 53) argues that the systems of proto-paradigmatic organization which we find in these dialects might be referred to as "East Asian paradigms", which "are characterized by their ability to combine multifunctionality with paradigm formation". Bisang further argues that "the emergence of this type of paradigms in which loss of semantic distinctiveness did not take influence on morphophonology was due to the relative frequency of multifunctional markers in the grammatical system as a whole". In other words, the dialects considered here only partly contradict Bisang's areal model of grammaticalization for EMSEA

languages: unexpected processes of strong reduction and fusion do occur, indeed, but this evolution, more often than not, is a morphophonological process, rather than 'coevolution of meaning and form' (Arcodia 2013); multifunctionality is anyway prominent, and pragmatic inference still seems to play a significant role.

The final question we would like to address concerns the skewed distribution of reduced morphology. Arcodia (2013, 2015) and Lamarre (2015), based on their own survey, paint a largely overlapping picture of the areal distribution of the phenomena at issue. While, as said earlier, reduced morphology is spread over a continuous area, the type of reduction we find is not homogeneous, and there appear to be four main areal clusters:

a. Northern Henan (along the border with Shanxi, Hebei, Shandong), the area around Zhengzhou and Kaifeng, and Southern Hebei, characterized by ablaut morphology (both grammatical and derivational)

b. Central-Southern Shaanxi (e.g. Xi'an, Shangzhou, Fengxiang), characterized by tonal morphology and, to a lesser extent, ablaut (mostly, grammatical)

c. Central-Eastern Shandong (especially the area around Zibo and the Jiaodong Peninsula), characterized by rhotacization and tonal morphology (grammatical and derivational)

d. Shanxi, characterized by (mostly) ablaut morphology, derivational or used to mark number in pronouns (but not for verbal categories)

We already discussed earlier the prosodic preconditions for reduction in Northern Sinitic, which however cannot explain why strong reduction happens almost only in these regions. Lamarre (2015: 284) highlights that the phenomenon of tone sandhi before toneless syllables (which, as said above, leads to the morphologization of tone change after the loss of the suffix) is not attested in MSC and in the Beijing dialect (the main contributor to MSC), which might explain the skewed distribution of this type of suprasegmental exponence for grammatical categories. However, this does not tell us much about the distribution of ablaut and rhotacization. Moreover, this type of grammatical (and derivational) tone change is attested also in Southern Sinitic, e.g. in a very typical syllable-timed language with no lexical stress as Cantonese (ex. from Matthews & Yip 2011:31; for other cases in Yue dialects, see Gan 2010):

(17) a. 食咗飯未呀?
sihk-jó-faahn meih a
eat-PFV-rice not.have Q
b. 食飯未呀?
sík-faahn meih a
eat.PFV-rice not.have Q
'Have you eaten?'

In (17a-b), we see that the perfective suffix 咗 *-jó*, with a high rising tone (thus, as expected, not toneless; see above, Section 2) causes tone change in the verb root, which becomes the only exponent of perfective aspect as the suffix disappears. It must be

pointed out, though, that the process behind this type of grammatical tone change may be different, and looks more like contraction (fusion of the verb and the suffix), rather than tone sandhi (see Yu 2007).

What about language contact? Lamarre (2015: 300) suggests that reduced morphology may be "linked with Northern Mandarin as an innovative area", and that the distribution of this type of exponents "is consistent with the distribution of other innovations that appeared in the same area". However, she also mentions that contact with Mongolic, Tungusic and Tibetic languages may have played a role in creating the phonetic and prosodic preconditions for the erosion of grammatical morphemes (see above, Section 3). Arcodia (2015) highlights that in the descriptions of many dialects with reduced morphology it is mentioned that the area where these dialects are currently spoken were populated by masses of immigrants from (present-day) Shanxi, starting at least from the Ming dynasty (see e.g. Xin 2006; Zhang 2011; Ai 2012): the influence of Shanxi (i.e. Jin) phonology may be found even in dialects of Shandong (Qiao 2008). Some authors (Wang 1999; Xin 2006) explicitly claim that reduced morphology originated in Shanxi and then spread eastwards with the dialects spoken by those immigrants: however, they specifically refer to derivational ablaut, i.e. rhyme change (mostly, of nouns) with derivational, rather than grammatical functions (known as 子变韵 *zi biànyùn* in Chinese linguistics). Indeed, as hinted at above, conservative dialects from Shanxi do not seem to make use of grammatical ablaut, or anyway of any type of reduced morphology with grammatical meaning (but, crucially, this is not true for Jin dialects in Henan and Hebei). Note, however, that in the field of Chinese dialectology there is much more published research on derivational ablaut than on verbal morphology (Lamarre 2015), and hence grammatical (verbal) morphology is more likely to be underreported.

Another possible explanation is that the clusters of reduced morphology described here are the product of more limited areal patterns of convergence, again connected with the innovative profile of Northern Sinitic (especially, Northern Mandarin). An area which deserves particular attention here is Henan (and Southern Hebei), where we find the highest concentration of grammatical ablaut: indeed, while tonal morphology is attested elsewhere, including southern China (see 17a-b), grammatical ablaut is clearly concentrated in this (sub-)area. Arcodia (2019) points out two more features whose distribution largely overlaps with reduced morphology: namely, object markers based on speech act verbs (e.g. Yexian 叫 *tɕiau*^{312} 'call'; Zhang 2005), and structural particles with an *l-* initial (e.g. Heshun 哩 *lei*^{31}; Liu 2013). Arcodia stresses the fact that these two features show a higher concentration in Henan, compared to the rest of the area of northern China at issue. Thus, Henan stands out as a cluster of typological features both within China and, to a lesser extent, within northern China.

What all of the hypotheses sketched here have in common is that they do not attribute the development of reduced morphology to contact with non-Sinitic languages (with the exception of the possible influence of 'Altaic' and Tibetic phonology), but rather to internal developments and, possibly, language contact within Sinitic. This is obviously different for all the other cases discussed earlier (Section 2), for which the diversity of Sinitic is explained by contact with non-Sinitic languages.

5. Concluding remarks

In this paper, we showed that the high degree of variation within Sinitic languages involves processes of grammaticalization too: while Chinese as a whole is generally seen as part of the EMSEA area, and it is claimed that grammaticalization in Sinitic follows the EMSEA areal pattern, this does not necessarily apply to the whole family. Indeed, the characteristics of grammaticalization in the Northern Sinitic languages we considered here represent another model, (partly) different from the EMSEA pattern, but also from the Indo-European model. It is very likely that this pattern was not generated under the influence of neighbouring non-Sinitic langauges, but is rather an internal development, arguably favoured by intrafamilial diffusion.

The study of patterns of secondary grammaticalization contributes to highlighting some methodological issues in Chinese dialectology and, more generally, in typology at large. As mentioned in the preceding section, the phenomena at issue here are likely to be underreported, due to some problematic assumptions in the research on Chinese dialects. Lamarre (2015: 278) summarizes these methodological issues as follows:

> A proper description of coalesced verb suffixes requires investigating full utterances, not only word lists, i.e. grammar, not only lexicon. On the other hand, a full training in dialectology is needed to account for tone sandhi patterns and various phonetic adaptation phenomena, which often differ according to the rime [...]. Eventually, one also needs to overcome the bias of character writing (one syllable = one character) and the widespread dogma that Chinese has no morphology

The influence on data collection of the writing system, in which there is an (almost) perfect correspondence between syllables and characters (and, thus, no grapheme for marking subsyllabic sounds and/or suprasegmental features) is explicitly mentioned e.g. in Wu & Han's description of the phenomenon of 'partial rhyme reduplication' (a type of ablaut) in the Qishan dialect. According to Wu & Han (2016: 225; my translation), these phenomena are often 'discarded' in descriptions because "there is no way to transcribe or record them using Chinese characters". They also point out that reduced morphology may not be easy to perceive for fieldworkers, and may thus escape their attention: this is highlighted also e.g. by Ai (2012) in her work on grammatical tone change in Changshan. They suggest that reduced morphology is likely to be more common than what can be gleaned from published descriptions. Again, further data collection is needed to properly assess the diffusion of these phenomena.

As pointed out in Section 1, much typological research concerning Sinitic is mainly or exclusively based on data of MSC (and/or Cantonese, the best described non-standardized Sinitic variety). However, MSC is a language 'created', in a sense, as "an imagined standard language for a growing modern nation in the 1950s" (Zhou 2012: 3). The relevance of dialectal data for generalisations about language (sub-)families and areas has already been pointed out, among others, by Lass (2000) and Traugott (2002), in reference to English, and by Laitinen (2004), with specific reference to grammaticalization:

> We need to find out more about non-standard varieties not only to complete the record but to test the worry expressed by Lass (2000) and others that many of the changes we observe might be artifacts of standardization and the pedagogical and discursive practices associated with it (Traugott 2002: 38)

> Grammaticalization phenomena were at first investigated in languages with large text corpora extending over centuries, but today spoken varieties and oral languages are also being studied within the same framework. The standardizing processes have, however, influenced different languages and language varieties in different ways. I argue that we should take more seriously into consideration the data we use when analyzing linguistic changes. [...] when studying languages that have gone through standardization we analyze languages that are at least partly artificial (Laitinen 2004: 247-248)

The concerns expressed by Traugott and Laitinen are indeed very relevant for the research on Sinitic languages, especially as far as grammaticalization is concerned.

Lastly, the unexpected patterns of secondary grammaticalization discussed here may also be seen as an aspect of internal differentiation within the Mandarin group: this is in line with recent research arguing that, despite the relatively shallow historical depth, Mandarin dialects are way less homogeneous than what is usually assumed (see Szeto, Ansaldo & Matthews 2018).

References

Ai, Hongjuan (艾, 红娟). 2012. *Shandong Changshan fangyan yanjiu* (山东长山方言研究). Beijing: Yuwen Chubanshe.

Ansaldo, Umberto & Lim, Lisa 2004. Phonetic absence as syntactic prominence. Grammaticalization in isolating tonal languages. In Fischer, Olga, Norde, Muriel & Perridon, Harry (a cura di), *Up and down the cline: The nature of grammaticalization*, 345–361. John Benjamins: Amsterdam-Philadelphia.

Ansaldo, Umberto, Bisang, Walter & Szeto, Pui Yiu. 2018. Grammaticalization in isolating languages and the notion of complexity. In Narrog, Heiko & Heine, Bernd (a cura di), *Grammaticalization from a typological perspective*, 219–234. Oxford: Oxford University Press.

Ansaldo, Umberto. 2010. Surpass comparatives in Sinitic and beyond: typology and grammaticalization. *Linguistics* 48(4). 919-950.

Arcodia, Giorgio F. 2013. Grammaticalisation *with* coevolution of form and meaning in East Asia? Evidence from Sinitic. *Language Sciences* 40. 148–167.

Arcodia, Giorgio F. 2015. More on the morphological typology of Sinitic. *Bulletin of Chinese linguistics* 8(1). 5–35.

Arcodia, Giorgio F. 2019. Diversity in Northern Sinitic languages. (Paper presented at the 32nd Paris Meeting on East Asian Linguistics, 27–29 June 2019).

Bickel, Balthasar & Nichols, Joanna. 2007. Inflectional morphology. In Shopen, Timothy (a cura di), *Language typology and syntactic description. Volume III: Grammatical categories and the lexicon* (second edition), 169–240. Cambridge: Cambridge University Press.

Bickel, Balthasar & Nichols, Joanna. 2013. Fusion of selected inflectional formatives. In Dryer, Matthews S. & Haspelmath Martin (a cura di), *The World Atlas of Language Structures Online*, chapter 20. Leipzig: Max Planck Institute for Evolutionary Anthropology (https://wals.info/chapter/20).

Bisang, Walter. 1996. Areal typology and grammaticalization: Processes of grammaticalization based on nouns and verbs in East and mainland South East Asian languages. *Studies in language* 20(3). 519–597.
Bisang, Walter 2004. Grammaticalization without coevolution of form and meaning: the case of tense-aspect-modality in East and mainland Southeast Asia. In Bisang, Walter, Himmelmann, Nikolaus P. & Wiemer, Björn (a cura di), *What makes grammaticalization? A look from its fringes and its components*, 109–138. Berlin-New York: Mouton de Gruyter.
Bisang, Walter 2008. Grammaticalization and the areal factor: The perspective of East and Mainland South East Asian languages. In López-Couso, Maria J. & Seoane, Elena (a cura di), *Rethinking grammaticalization*, 13–35. Amsterdam-Philadelphia: John Benjamins.
Bisang, Walter 2014. On the strength of morphological paradigms: A historical account of radical pro-drop. In Robbeets, Martine and Bisang, Walter (a cura di), *Paradigm change: In the Transeurasian languages and beyond*, 23–60. Amsterdam-Philadelphia: John Benjamins.
Bisang, Walter 2015. Problems with primary vs. secondary grammaticalization: the case of East and mainland Southeast Asian languages. *Language sciences* 47. 132–147.
Bybee, Joan 2003. Mechanisms of change in grammaticization: The role of frequency. In Joseph, Brian D. & Janda, Richard D. (a cura di), *The handbook of historical linguistics*, 602–623. Oxford: Blackwell.
Bybee, Joan, Perkins, Revere & Pagliuca, William. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago: Chicago University Press.
Chao, Yuen Ren. 1968. *A grammar of Spoken Chinese*. Berkeley: University of California Press.
Chappell, Hilary. 2006. Language contact and areal diffusion in Sinitic languages. In Aikhenvald, Alexandra Y. & Dixon, Robert M.W. (a cura di), *Areal diffusion and genetic inheritance*, 328–357. Oxford: Oxford University Press.
Chappell, Hilary. 2015a. Introduction: Ways of tackling diversity in Sinitic languages. In Chappell, Hilary (a cura di), *Diversity in Sinitic languages*, 3–12. Oxford: Oxford University Press.
Chappell, Hilary. 2015b. Linguistic areas in China for differential object marking, passive, and comparative constructions. In Chappell, Hilary (a cura di), *Diversity in Sinitic languages*, 13–52. Oxford: Oxford University Press.
Chen, Ning (陈, 宁). 2006. Shandong Boshan fangyan de zi bianyun ji xiangguan wenti (山东博山方言的子变韵及相关问题). *Fangyan* 4. 316–322.
Chen, Pengfei (陈, 鹏飞). 2005. Linzhou fangyan “le” de yuyin yanbian jiqi yuyi fengong (林州方言“了”的语音演变及其语义分工). *Nankai Yuyanxuekan* 1. 76–80.
Chen, Pengfei (陈, 鹏飞). 2007. Zuhe gongneng bianhua yu “le” yufahua de yuyin biaoxian (组合功能变化与 “了” 语法化的语音表现). *Henan Shehui Kexue* 15(2). 138–140.
DeLancey, Scott. 2015. The historical dynamics of morphological complexity in Trans-Himalayan. *Linguistic Discovery 13*(2). 60-79.
Duanmu, San. 2007. *The phonology of Standard Chinese (second edition)*. Oxford: Oxford University Press.

Duanmu, San. 2014. Syllabe structure and stress. In Huang, James C.-T., Li, Audrey Y.-H. & Simpson, Andrew (a cura di), *The handbook of Chinese linguistics*, 422–442. Oxford: Wiley Blackwell.

Enfield, Nick. 2005. Areal linguistics and Mainland Southeast Asia. *Annual Review of Anthropology* 34. 181–206.

Gan, Yu'en (甘, 于恩). 2010. *Guangdong Siyi fangyan yufa yanjiu* 广东四邑方言语法研究. Guangzhou: Jinan Daxue Chubanshe.

Goddard, Cliff. 2005. *The languages of East and Southeast Asia*. Oxford: Oxford University Press.

Hashimoto, Mantaro. 1986. The Altaicization of Northern Chinese. In McCoy, John & Light, Timothy (a cura di), *Contributions to Sino-Tibetan studies*, 76–97. Leiden: Brill.

Hopper, Paul J. & Traugott, Elizabeth C. 2003. *Grammaticalization (second edition)*. Cambridge: Cambridge University Press.

Jacques, Guillaume. 2017. The genetic position of Chinese. In Sybesma, Rint *et al.* (a cura di), *Encyclopedia of Chinese language and linguistics*, Vol. 2, 297–306. Leiden: Brill.

Jiang, Lansheng (江, 藍生). 1999. Yufahua chengdu de yuyin biaoxian (语法化程度的语音表现). In Shi, Feng (石， 锋) & Pan, Wuyun (潘, 悟云) (a cura di), *Zhongguo Yuyanxue de Xin Tuozhan* (中国语言学的新拓展), 195–204. Hong Kong: Hong Kong City University Press.

Laitinen, Lea. 2004. Grammaticalization and standardization. In Fischer, Olga, Norde, Muriel & Perridon, Harry (a cura di), *Up and down the cline*: *The nature of grammaticalization*, 247–262. John Benjamins: Amsterdam-Philadelphia.

Lamarre, Christine (柯理思). 2009. Lun beifang fangyan zhong weiyi zhongdian biaoji de yufahua he juweiyi de zuoyong (论北方方言中位移重点标记的语法化和句位义的作用). In Wu, Fuxiang (吴, 福祥) & Cui, Xiliang (崔， 希亮) (a cura di), *Yufahua yu yufa yanjiu - si* (语法化与语法研究−四), 145–187. Beijing: Shangwu Yinshuguan.

Lamarre, Christine. 2015. The morphologization of verb suffixes in Northern Chinese. In Cao, Guangshun, Djamouri, Redouane & Peyraube, Alain (a cura di), *Language contact in north China: Historical and synchronic studies*, 277–308. Paris: EHESS.

LaPolla, Randy J. 2001. The role of migration and language contact in the development of the Sino-Tibetan family. In Aikhenvald, Alexandra Y. & Dixon, Robert M.W. (a cura di), *Areal diffusion and genetic inheritance*, 225–254. Oxford: Oxford University Press.

Lass, Roger. 2000. Remarks on (uni)directionality. In Fischer, Olga, Rosenbach, Annette & Stein, Dieter (a cura di), *Pathways of change. Grammaticalization in English*, 207–227. Amsterdam-Philadelphia: John Benjamins.

Lehmann, Christian. 2015. *Thoughts on grammaticalization*. 3ª ed. Berlin: Language Science Press.

Li, Charles N. & Thompson, Sandra A. 1981. *Mandarin Chinese. A functional reference grammar*. Berkeley: University of California Press.

Li, Rong (李, 荣). 1985. Guanhua fangyan de fenqu (官话方言的分区). *Fangyan* 1. 2–5.

Li, Rulong (李, 如龙). 2002. Lun Hanyu fangyan de yuliu yinbian (论汉语方言的语流音变). *Xiamen Daxue xuebao* 6. 43-50.

Lin, Jo-Wang 2006. Time in a language without tense: The case of Chinese. *Journal of Semantics* 23(1). 1–53.
Liu, Lin (柳, 琳). 2013. Qian tan Shanxi Jinzhong Heshun fangyan de tese zhuci “li” (浅谈山西晋中和顺方言的特色助词“哩”). *Sheke xuelun* 2. 158–159.
Liu, Xuexia (刘, 雪霞). 2006. *Henan fangyan yuyin de yanbian yu cengci* (河南方言语音的演变与层次). Shanghai: Fudan Daxue (PhD dissertation).
Matisoff, James A. 2001. Genetic versus contact relationship: Prosodic diffusibility in South-East Asian Languages. In Aikhenvald, Alexandra Y. & Dixon, Robert M.W. (a cura di), *Areal diffusion and genetic inheritance*, 291–326. Oxford: Oxford University Press.
Matthews, Stephen & Yip, Virginia. 2011. *Cantonese: A Comprehensive Grammar (second edition)*. London: Routledge.
Nichols, Johanna. 1992. *Linguistic diversity in space and time*. Chicago: University of Chicago Press.
Norman, Jerry. 2003. The Chinese dialects: phonology. In Thurgood, Graham & LaPolla, Randy J. (a cura di), *The Sino-Tibetan languages*, 72–83. London: Routledge.
Packard, Jerome. 2006. Chinese as an isolating language. In Brown, Keith *et al.* (a cura di), *Encyclopedia of language and linguistics* (second edition), 355–359. Amsterdam: Elsevier.
Post, Mark & Sun, Jackson T.-S. 2017. Tani languages. In Thurgood, Graham & LaPolla, Randy J. (a cura di), *The Sino-Tibetan languages (second edition)*, 322–337. London: Routledge.
Qian, Cengyi (钱, 曾怡). 1993. *Boshan fangyan yanjiu* (博山方言研究). Beijing: Shehui kexue wenxian chubanshe.
Qiao, Quansheng (乔, 全生). 2008. Jin fangyan xiangwai de jici kuosan (晋方言向外的几次扩散). *Yuwen Yanjiu* 106(1). 45–48.
Sandman, Erika. 2016. *A grammar of Wutun*. Helsinki: University of Helsinki (PhD dissertation).
Slater, Keith W. 2003. *A grammar of Mangghuer*. London: Routledge.
Sun, Jackson T.-S. 2014. Sino-Tibetan: Rgyalrong. In Lieber, Rochelle & Štekauer, Pavol (a cura di), *The Oxford handbook of derivational morphology*, 630–650. Oxford: Oxford University Press.
Sun, Lixin (孙, 立新). 2007. *Xi’an fangyan yanjiu* (西安方言研究). Xi’an: Xi’an Chubanshe.
Szeto, Pui Yiu, Ansaldo, Umberto & Matthews, Stephen. 2018. Typological variation across Mandarin dialects: An areal perspective with a quantitative approach. *Linguistic Typology* 22(2). 233–275.
Traugott, Elizabeth C. 2002. From etymology to historical pragmatics. In Minkova, Donka & Stockwell, Robert (a cura di), *Studies in the history of the English language*, 19–49. Berlin-New York: Mouton de Gruyter.
VanBik, Kenneth. 2009. *Proto-Kuki-Chin: A reconstructed ancestor of the Kuki-Chin languages*. Berkeley: Sino-Tibetan Etymological Dictionary and Thesaurus Project, University of California.
Wang, Futang (王, 福堂). 1999. *Hanyu fangyan yuyin de yanbian yu cengci* (汉语方言语音的演变和层次). Beijing: Yuwen Chubashe.

Wang, Sen (王, 森). 1998. Zhengzhou Xingyang (Guangwu) fangyan de bianyun (郑州荥阳（广武）方言的变韵). *Zhongguo Yuwen* 4. 275–283.
Wu, Yuan (吴, 媛) & Han, Baoyu (韩, 宝育). 2016. *Qishan fangyan diaocha yanjiu* (岐山方言调查研究). Beijing: Zhonghua Shuju.
Xin, Yongfen (辛, 永芬). 2006. *Shunxian fangyan yufa yanjiu* (浚县方言语法研究). Beijing: Zhonghua Shuju.
Xu, Dan. 2017. *The Tangwang language: An interdisciplinary case study in Northwest China*. Cham: Springer.
Yu, Alan C.L. 1997. Understanding near mergers: the case of morphological tone in Cantonese. *Phonology* 24. 187-214.
Zhang, Li (张, 丽). 2011. *Hebei Nanhe fangyan yinbian diaocha yanjiu* (河北南和方言音变调查研究). Baoding: Hebei Daxue (MA dissertation).
Zhang, Xueping (张, 雪平). 2005. Henan Yexianhua de “jiao” zi ju (河南叶县话的“叫”字句. *Fangyan* 4. 301–305.
Zhang, Zhanshan (张, 占山) & Li, Rulong (李, 如龙). 2007. Xuhua de zhongji: heyin (虚化的终极：合音). *Ludong Daxue xuebao* 24(2). 95–100.
Zhou, Minglang. 2012. Introduction: The contact between Putonghua (Modern Standard Chinese) and minority languages in China. *International Journal of the Sociology of Language* 215. 1–17.