HAL
open science

LANGUAGE CONTACT AND AREAL DIFFUSION IN SINITIC LANGUAGES

Hilary M. Chappell

To cite this version:

Hilary M. Chappell. LANGUAGE CONTACT AND AREAL DIFFUSION IN SINITIC LANGUAGES. Alexandra Aikhenvald & RMW Dixon. Areal diffusion and genetic inheritance: problems for typology and genetic affiliation, Oxford University Press, pp.328-357, 2006, 0199283087. hal-00850205

HAL Id: hal-00850205

https://hal.science/hal-00850205v1

Submitted on 5 Aug 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

LANGUAGE CONTACT AND AREAL DIFFUSION IN SINITIC LANGUAGES

(pre-publication version)

Hilary Chappell

This analysis includes a description of language contact phenomena such as stratification, hybridization and convergence for Sinitic languages. It also presents typologically unusual grammatical features for Sinitic such as double patient constructions, negative existential constructions and agentive adversative passives, while tracing the development of complementizers and diminutives and demarcating the extent of their use across Sinitic and the Sinospheric zone. Both these kinds of data are then used to explore the issue of the adequacy of the comparative method to model linguistic relationships inside and outside of the Sinitic family. It is argued that any adequate explanation of language family formation and development needs to take into account these different kinds of evidence (or counter-evidence) in modeling genetic relationships.

In §1 the application of the comparative method to Chinese is reviewed, closely followed by a brief description of the typological features of Sinitic languages in §2. The main body of this chapter is contained in two final sections: §3 discusses three main outcomes of language contact, while §4 investigates morphosyntactic features that evoke either the North-South divide in Sinitic or areal diffusion of certain features in Southeast and East Asia as opposed to grammaticalization pathways that are crosslinguistically common.[i]

1. The comparative method and reconstruction of Sinitic

In Chinese historical phonology, various methods have been applied with relative success to the Sinitic family in the reconstruction of both stages of Middle and Old Chinese. In *Etudes sur la phonologie chinoise* (1915-26), Karlgren published his ground-breaking reconstruction of Middle Chinese according to three main sources: an analysis of rhyme tables based on the early seventh century dictionary *Qièyùn* (601

AD), Sinoxenic readings from Japanese and Vietnamese, and data from nineteen dialects which he collected while carrying out fieldwork in China from 1910 to 1912. Strictly speaking, he did not apply the comparative method to these dialect data but determined the phonological system of Middle Chinese on the basis of the *Qièyùn*, interpreting and assigning phonetic values to the rhyme categories.[ii] Note that the *Qièyùn* dictionary was compiled as a guide to the correct pronunciation for the recitation of the classics. Hence, its precise relation to the spoken language of its time is not transparent. Many scholars believe that it is based on several different spoken dialects of the time and not just that of the capital, Chang'an (present-day Xi'an), while others believe it reflects educated speech from the 6th century CE, that is, the end of the Nanbeichao dynasty (Northern and Southern dynasties, 420-589 CE).

Karlgren later worked on the reconstruction of Old Chinese based on his Middle Chinese reconstruction in conjunction with an analysis of the rhyme categories of the *Shījīng* (Book of Odes) and the information which could be deduced from the phonetic components present in most Chinese written characters. Old Chinese hypothetically reflects the elevated speech of the late Zhou period of fifth to third centuries BCE, in the view of some scholars, or the even earlier period of the Western Zhou in the view of others (roughly the first half of the first millennium BCE). These are not, however, uncontroversial issues, for which a fuller discussion may be found in Sagart (1999) or for a contrary view, in Baxter (1992).

The *Shījīng* is an anthology of poems from 1000 – 500 BC, compiled in the sixth century BCE. An early observation made by scholars in China was that characters which rhymed in it generally contained the same phonetic element. Karlgren's contribution was similarly to interpret and assign values to the categories of initials and finals in the *Shījīng* (Book of Odes) which would obey regular phonetic laws for development into those he had earlier posited for Middle Chinese. Karlgren's second reconstruction was published in 1940 as *Grammata Serica* with a revised version appearing in 1957. Given the lack of records of real dialect materials from the late Nanbeichao and Sui periods to which Middle Chinese roughly corresponds, the reconstruction of Old Chinese could not avoid being the more hypothetical of the two. Karlgren's postulation of these two earlier stages of the Chinese language inspired further work by sinologists resulting in revisions and new breakthroughs, and provided indisputable evidence for the genetic relationship of Sinitic languages, though be it mainly on the basis of phonology and the lexicon.[iii]

Nonetheless, the focus on phonetic laws and the use of the neo-grammarian approach with its assumption of homogeneous data in Chinese linguistic reconstruction was early criticized by Grootaers (1943) and Serruys (1943) as the sole means of relating dialects to Old and Middle Chinese. In particular, they both objected to Karlgren's use of character lists for elicitation and dialect dictionaries based on the reading of standard Chinese characters. The reading lists not only required literate language informants but could also hardly avoid producing the literary pronunciations which by definition hold a close relationship to the standard language, Mandarin, and thus neatly supported his reconstruction (see also §3.2.1 on stratification). In many cases, these pronunciations represented morphemes not used at all in the local *patois* which belong to the purely colloquial level.

In the same study, Grootaers (1943) shows how methods in geographical linguistics can be successfully applied to capturing dialect isoglosses in Northern Chinese for both the innovation and extent of use of phonetic and lexical features, based on 'real' colloquial items. Similarly, Hashimoto (1992) pioneered the use of *Wellentheorie* (wave theory) in Chinese linguistics to account for the spread of tonal categories and phonetic features such as retention or loss of voicing in Chinese dialects. The use of lexical and morphological data has also been incorporated in various handbooks produced by Beijing University in the 1960s such as *Hànyǔ fāngyán cíhuì* [A lexical list for Chinese dialects] and *Hànyǔfāngyán gàiyào* [An outline of Chinese dialects] compiled by Yuan (1960 [1989]) which includes syntactic data. More recently the inutility of the family tree model to explain how languages develop in a relatively stable environment is raised by Hashimoto (1992: 32) for Hakka and by Dixon (1997) for the general case.

In sections three and four which follow, it is argued that the family tree model, used alone, is inadequate to capture the complexities of linguistic phenomena created during the course of evolution and geographical distribution of a language family: the comparative method and the family tree model simply cannot account for all the facets associated with language change and development and to be fair were never intended to do so. They need to be used in conjunction with other methods to account for the effects of language contact such as stratification, hybridization and convergence, not to mention other possible outcomes such as mixed languages and language obsolescence.

2. Typological features of Sinitic

Sinitic languages form a sister group with the Tibeto-Burman languages of the Sino-Tibetan language family located in East and Southeast Asia. As a language family, Sinitic languages are as diverse as the Romance or Germanic languages within the Indo-European family. The spoken forms of Chinese languages are not mutually intelligible: a speaker of Suzhouese, a Wu dialect, will not understand a compatriot from Quanzhou, who speaks a Southern Min dialect. Even within dialect groups such as Min or Yue there is a high degree of mutual unintelligibility between subdivisions such as Coastal versus Inland Min, or one of the Guangxi Yue dialects versus Hong Kong Cantonese Yue.

Typologically, Sinitic languages are tonal languages which show analytic or isolating features, though in some Min languages, for example, the development of case markers and complementizers from lexical verbs, and the use of a range of nominal suffixes, has moved further along the path of grammaticalization than for Mandarin. Complex allomorphy is also widespread in Min dialects, exemplified by the many forms for each negative marker in Fuzhouese (Northeastern Min) and for the diminutive suffix in Southern Min.

Tone sandhi (or tone change) can be used to code morphological functions in Chinese languages. For example, in Toishan Cantonese, aspectual distinctions such as the perfective and the plural form of pronouns can be signalled in this way. Tone sandhi phenomena are, however, most conspicuous in the Min and Wu dialect groups where citation or juncture forms for each syllable differ from contextualized forms. Although Sinitic languages have SVO basic word order, object preposing is a common contrastive device and postverbal intransitive subjects are common in presentative constructions. The modifier generally precedes the modified element. This means that subordinate or backgrounding clauses typically precede main clauses while attributives precede head nouns and adverbs precede verbs. Well-known exceptions to this rule are presented by the case of gender affixes on animal terms and certain semantic classes of nominal compounds and adverbs in many Southern Sinitic languages.

The ten major Sinitic languages (or Chinese dialect groups) that are generally recognized are listed below:

I.	Northern Chinese (Mandarin)	北方话
II.	Xiang	湘
III.	Gan	赣
IV.	Wu	吴
V.	Min	闽
VI.	Kejia or Hakka	客家
VII.	Yue dialects	粤
VIII.	Jin dialects	晋
IX.	Hui dialects	徽
X.	Pinghua	平话

Mandarin covers the largest expanse of territory from Manchuria in the northeast of China to Yunnan and Sichuan provinces in the southwest. Apart from the Jin dialects, the eight other dialect groups fall neatly into almost complementary geographical distribution with Mandarin, covering the east and southeast of China: Xiang dialects are largely concentrated in Hunan province; Gan in Jiangxi; Wu in southern Jiangsu and Zhejiang provinces; Hui dialects in southern Anhui and adjacent areas of Jiangxi and western Zhejiang provinces; Min in Fujian; Yue in both Guangdong and Guangxi provinces; Kejia in northeastern Guangdong, southwestern Fujian and parts of Jiangxi and Sichuan provinces and the Pinghua dialects in Guangxi. The Jin dialects in Shanxi province and Inner Mongolia represent the only non-migrant dialect group to be found in Northern China, apart from Mandarin.

2.1 *A note on Chinese dialect history*

According to Bellwood (this volume), archaeological evidence points to Neolithic settlements in two areas of modern China -- the middle and lower Huang He (Yellow River) and the Yangzi River valleys. These can be dated to around 7000 BCE. However, reconstruction of Proto-Chinese, based on the diversity found in modern dialects, cannot hope to reach much further back than the first millennium BCE (see §1).

Overall, the development of Sinitic languages over the last two and a half millennia can be aptly modeled in terms of its history of imperialist unification and

expansion accompanied by ensuing periods of relative equilibrium. These were in turn regularly punctuated by periods of disunity and temporary fragmentation of the Chinese empire. During the formation time of the Sinitic group, the major migrations of the Han Chinese took place from northern China to various regions in the south, for which a detailed coverage of population movements in China over the last several millennia is provided in LaPolla (this volume) while a brief history of Chinese dialects is given in Chappell (in press [d]) and thus not recapitulated here.

The general consensus regarding the approximate time of diversification of Chinese into the present-day dialect groups is around the time of Medieval Chinese during the Sui (581-618) and early Tang dynasties (618-907) for Yue, Xiang, and Gan but earlier, during the transitional period for the Han dynasty (206 BC - 220 AD) for the ancestral language(s) of Wu and Min. Sagart (1988, in press) and You (1992: 97) claim that Wu, Xiang, Yue and Gan developed directly from earlier stages of Northern Chinese whereas Min was probably a secondary development from a Southern Sinitic language such as Wu (or proto Wu-Min), and Hakka, similarly, a secondary development from Southern Gan during the Tang period. Ting (1983) and Norman (1988: 189) do not entirely concur with this view regarding Min, holding that there is a strong demarcation line between Wu and Min linguistic territory, but agree on the early split. The larger dialect picture for Sinitic languages was thus essentially in place by the end of the Southern Song (1127-1279), apart from the later formation of the Hui dialects by the early Ming dynasty (1368-1644).

Sagart aptly describes dialect groups as 'fuzzy entities that owe (as) much (of) their make-up to contact as opposed to vertical inheritance' (1997: 298-9). He further argues for the difficulty of using isoglosses to determine dialect boundaries given that innovations may be obliterated or reversed through contact, with the result that the family tree model is only strictly applicable to rarer situations where diversification and loss of contact co-occur, as for Austronesian, concurring with Dixon (1997). The history of Sinitic languages certainly presents a case in point, exemplifying the difficulties that could arise if the family tree model and comparative method were exclusively used to represent genetic relationships. The implication is that a fuller description of the evolution of Sinitic languages necessarily involves modeling genetic relatedness as well as the characteristics of *Mischsprachen*, 'mixed languages', (see Heine & Kuteva this volume) combining substratum or superstratum features of 'step-parent' contact languages (Dixon 1997: 71). These, in their turn, can

be either genetically related or unrelated which has further typological ramifications. Next I consider some aspects of areal diffusion in the Southeast and East Asian region before beginning on the main discussion.

2.2 *Areal diffusion*

Mantaro Hashimoto has convincingly argued for a North-South divide for Chinese languages on the basis of phonological, lexical and syntactic evidence (see Hashimoto 1974, 1976a, 1976b, 1986). His thesis essentially has the following argumentation: Chinese languages are sandwiched between Altaic languages in the north and Tai languages in the south, with the typogeographical consequence of Altaicization of northern Chinese varieties and Taiization of Southern Sinitic. Furthermore, he observes that the north-south opposition can be clearly perceived in features such as the increasing number of classifiers, tones and consonantal endings to syllables, not to mention the monosyllabic nature of morphemes as one moves southwards. He notes that some varieties of Northern Chinese show agglutinative tendencies, witnessed in the existence of a postposition for accusative/dative case in Qinghai Mandarin, stress-accent dominance over tone, and adoption of O-V structures as in Northwestern Mandarin dialects spoken in Qinghai and Gansu provinces. Other broad divisions are the typically MODIFIER-MODIFIED word order in the north versus MODIFIED-MODIFIER order for some structures in the south; different comparative strategies; different word orders for the 'double object' or ditransitive construction; and aspect and tense distinctions maintained in the south while merged in the north.

To this could be added the more limited use of patient-marking or disposal constructions where the direct object is positioned before the main verb and preceded by a special marker, for example, the extensively researched *bǎ* 把 construction in Mandarin: S – *bǎ*– O – V. In its canonical form, it codes a highly transitive event that affects a referential object with a specifiable effect or result state. Cheung (1992) has shown that Cantonese, which uses the Medieval Chinese exponent *jeung*[1] 将[*jiāng*], is restricted to transitive verbs, whereas Mandarin also allows its use with intransitive verbs provided there is a causative interpretation (see Chappell 1992a). Furthermore, the use of *jeung*[1] is more a feature of formal discourse than colloquial Cantonese,

evidence of Mandarin influence. Similarly, Hakka also reportedly uses this construction much less frequently than Mandarin (Yuan 1989).

Bisang (1996) presents a typology of classifiers according to their functions in Southeast and East Asian languages, showing a similar set of geographical correlations with respect to enumeration, referentialization and other parameters. In Cantonese, for example, classifiers may also be used as possessive and relative clause markers, thus showing a greater alliance with Tai languages as opposed to Northern Chinese which does not permit this function.

With regard to Northern Chinese, Hashimoto (1986: 95) suggests that a pidgin Chinese developed when Altaic peoples became sinicized, and that while they adopted Chinese lexicon and morphology they retained the syntax of Altaic, and possibly its phonetic system as well. This must be a two-step process however: presumably what is meant by Altaicization follows on as the next step after cultural sinicization, whereby the superstrate Altaic syntactic structures slowly diffuse into the different varieties of Northern Chinese and then gradually southwards into other Sinitic languages by virtue of the prestige of Mandarin. He observes that this is not unique to northern Chinese: the Ong-Bê language of southwestern China, a Tai language, has undergone the same process of sinicization (1986: 95), as too pre-war Korea with respect to the effect of Japanese on Korean.

Matisoff (1991: 386; this volume) refines Hashimoto's basic classification by dividing the larger Southeast Asian zone into two main areas: the Sinospheric and the non-Sinospheric. The Sinospheric area includes Southern Sinitic (basically Sinitic languages south of the Yangzi) and the language families which have been in close cultural contact with China such as Hmong-Mien, Tai-Kadai, Vietnamese in the Mon-Khmer branch of Austroasiatic, and certain branches of Tibeto-Burman such as Lolo-Burmese. The non-Sinospheric languages include Austronesian languages, many Mon-Khmer languages, and Tibeto-Burman languages, for example those found in Northeastern India and Nepal.

According to Matisoff (1991) some of the broad grammatical features which unify the Southeast Asian area into a linguistic zone are the following:

1) development of modal verbs > desiderative markers, 'be likely to'
2) development of verbs meaning 'to dwell' > progressive aspect markers
3) development of verbs meaning 'to finish' > perfective aspect markers

4) development of verbs meaning ‘to get, obtain’ > ‘manage’, ‘able to’, ‘have to’
5) development of verbs of giving > causative and benefactive markers
6) development of verbs of saying > complementizers, topic and conditional markers
7) formation of resultative and directional compound verbs through verb concatenation

With respect to Sinitic, all of these pathways of grammaticalization apply to Northern Chinese as well, with the exception of a ‘say’ verb developing into a complementizer and the limited use of ‘give’ with a causative meaning. Both these paths of grammaticalization are treated in §4 for Southern Sinitic languages while other pathways, such as for *get* verbs, are analysed in depth in Enfield (this volume). Next, I discuss some linguistic phenomena that are the result of language contact, illustrating some of the potential difficulties for modeling the outcomes of language contact including stratification, metatypy, hybridization and convergence.

3. Language contact: stratification, hybridization and convergence

Synchronically, there are three main outcomes of language contact situations for Sinitic languages: stratification, hybridization and convergence. Examples of all three outcomes are discussed in this main section. Stratification and hybridization of syntactic and morphosyntactic forms are a widespread phenomenon in Sinitic languages.

3.1 *Stratification*

Stratification has resulted from the systematic introduction of certain features of the prestige language in China for the purposes of reciting classical texts; or as forms borrowed from this standard language (different varieties of Mandarin). Moreover, this has occurred more than once in the historical development of several of the major Chinese dialect groups such as Min which has three such layerings from Northern Chinese: the Han dynasty stratum (206 BC – 220 AD); the Nanbeichao stratum (420-581 AD) and the late Tang stratum (8th – 10th centuries). The degree of stratification varies along a continuum from minor phonological differences, as in Hakka, to major stratification of the lexicon and a marked contrast between the literary and colloquial pronunciations as in Southern Min. The differences in pronunciation are known as

wén-bái yì-dú 文白异读 in Chinese linguistics. The *bái* or vernacular pronunciation for each syllable in a given dialect represents the native morpheme which may or may not have a *wén* or reading doublet whose pronunciation has been adopted from Northern Chinese.

For example, in the Xiamen or Amoy dialect of Southern Min words in the reading pronunciation which end in a velar nasal often have a nasalized vowel in the cognate colloquial form: the character for 'name',名, has the literary form *bêng* versus colloquial *miâ*^{n}. In other cases the relationship is not so straightforward: the preposition 'to, with' written as 共 has *kā* as its colloquial pronunciation but *kiōng* as its reading pronunciation, with the latter closer to the modern standard Mandarin /*kuŋ*/[1] in form. Similarly, the possessive morpheme 其 has *ê* for its colloquial pronunciation but *kî* for its literary one, closer to Mandarin /*tɕ'i*/[2]. In many cases, it first needs to be established whether there is any cognacy at all. There clearly is none for the suppletive relationship between these possessive morphemes, nor for the two readings of the diminutive suffix 仔 which has *á* for the colloquial as opposed to *tsú* for the literary. Again, the reading form resembles modern standard Mandarin very closely, which is /*tsɹ*/[3]. As argued below, the diminutive suffix has evolved from another morpheme for 'son' in Min: *kiá*^{n}.

Most non-Mandarin Sinitic languages show this kind of phonological and lexical stratification as a result of different periods of intense contact with Mandarin, particularly with the emergence of an official court language in the mid- to late- Tang period (eighth to tenth centuries AD), a koine based on the language of the capital, Chang'an, where a northwestern dialect of Northern Chinese waɪɪɪs spoken. This was brought to southern regions during the migrations of the later Tang dynasty and is the basis of the reading or literary pronunciation in most Southern Sinitic languages, as noted above. In some dialect groups, a second overlay of a more eastern variety of Northern Chinese occurred after the establishment of the Liao (916-1125 CE), Jin (1115-1234 CE) and Yuan dynasties (1271-1368 CE) in northern China, whose capitals were located in the region of Beijing. It is significant that both koines are associated with flourishing vernacular literatures (Norman 1988) and the strong tendency to standardize language use that accompanies the consolidation of an imperial system of government. More traditional research has mainly concentrated on describing the phonological correspondences between the reading and colloquial

pronunciations of characters. Recent pioneering work on syntax by Zhu Dexi (1990) and Anne Yue-Hashimoto (1991) has uncovered several different strata for the syntax of interrogative forms in Southern Sinitic (see §3.2.3). For the purposes of any kind of comparative work, the native stratum must first be clearly separated from the imported stratum.

3.2 *Lexical and morphological stratification*

Lien's study of morphological change in Taiwanese Southern Min (to appear) shows that this historical process of layering has resulted in different kinds of stratificational distinctions in the lexicon for the native colloquial morphemes versus the 'alien' literary forms. Taiwanese Southern Min belongs to the subdivision of Coastal Min and is closely related to the Xiamen (Amoy), Quanzhou and Zhangzhou dialects spoken on the south coast of Fujian province. It is the first language of over seventy three per cent of the population in Taiwan, despite the fact that Mandarin is the official language.

As Lien observes, since this variation is present in everyday colloquial language, it cannot simply be explained as the existence of separate registers resulting from the impact of Mandarin on Southern Min during the Tang period. He discusses cases of morphological competition which have been synchronically resolved in favour of either the colloquial or literary stratum and concludes on the basis of his data that the diffusion is clearly bidirectional.

For example, the morpheme *lâng* 儂'person' represents the first type where this colloquial form is in the ascendant over the literary and unproductive bound morpheme *jîn* 人 which also means 'person' but was borrowed from the Tang Northern Chinese koine. It is not cognate with *lâng*. Couplets thus exist, such as *toā lâng* 大儂 'adult' versus *tāi jîn* 大人'police officer'(a polite vocative akin to 'Sir'), where both are formed with morphemes for 'big' + 'person'. This is indicative, Lien argues, of *jîn* developing a special idiomatic meaning in many of its compounds. The literary morpheme *jîn* generally occurs with less frequency as a suffix than *lâng,* according to a statistical count made by Lien. It is much less likely to occur affixed to disyllabic stems, and never with those from the colloquial stratum. Furthermore, in coining new words, he notes, the younger generation prefers the native morpheme *lâng.*

Similarly, for numerals, the colloquial forms are used for cardinal numbers while the literary forms are used for giving telephone numbers and for calendar years in the Gregorian or western calendar. Lien observes, however, that in the case of ordinal numbers, the colloquial forms are winning out from the lexeme 'third' upwards. He attributes this outcome to the lack of literacy in the native language, Taiwanese Southern Min, as opposed to high literacy in the official language, Mandarin: it is nowadays rare for younger generation first-language speakers of Taiwanese to be instructed in the reading pronunciations and forms of Southern Min.

The second type, where the literary form is more productive than the colloquial form, is represented by suffixes which are in complementary distribution such as colloquial *ke* versus literary *ka* (which share the etymon for 'family'家). These are used as agentive suffixes or nominalizers but, significantly, in different semantic fields: the first, colloquial form *ke* shows a broader application as it is used not only for family relationships but also for those pertaining to the old agrarian society such as *head-servant* and *master* and names for relatives in-law while the second, literary form *ka* applies to higher status professions of the new industrialized society such as *writer, connoisseur, diplomat, statesperson.* Nonetheless, colloquial *ke* has become 'inert' and unproductive.

Similarly, colloquial *sai-hū* versus literary *su* act as agentive suffixes, the first referring to trades and crafts that require manual labour, while the second refers to professions that require intellectual skills. This is shown in the following two tables reproduced from Lien (in press):

Table 1: Derivatives with colloquial suffix *sai-hū* 师傅 in Southern Min

Agent noun	Gloss	Translation
thô͘-chúi sai-hū 涂水师傅	mud-water-master	bricklayer
chúi-tiàn sai-hū 水电师傅	water-electricity-master	electrician/plumber
iû-chhat sai-hū 油漆师傅	oil-paint-master	painter
bak-chhiū^{n} sai-hū 木匠师傅	wood-wright-master	carpenter

Table 2: Derivatives with literary suffix *su* 师

Agent noun	Gloss	Translation
i-su 医师	treat medically-master	doctor
kàu-su 教师	teach-master	teacher
ūi-su 画师	draw-master	artist
káng-su 讲师	talk-master	instructor

Both these cases contrast with the outcome for the competition between morphemes for 'person' in that the literary form is very productive, and a clear semantic division of labour is apparent. Lien characterizes the colloquial stratum as typified by basic and popular vocabulary, versus the technical and cultural vocabulary representative of the literary stratum. Despite this mixing and integration of the literary stratum into everyday language, convergence of the two strata is not likely, particularly where the semantic specialization of the two sets has occurred, as for *ke* and *ka* and *sai-hū* and *su*. Lien concludes that only a bidirectional diffusion of features can explain the continuing co-existence of these strata.

3.3 *Syntactic stratification: preverbal interrogative markers*

Zhu (1990) and Yue-Hashimoto (1991) discuss the complementary distribution in Sinitic languages of neutral interrogative constructions using the Northern Sinitic strategy of VP-NEG-VP as opposed to Southern Sinitic constructions using either a preverbal interrogative adverb (ADV-VP) or a VP-NEG-(PARTICLE) form for this type of Yes/No question. These interrogatives are described as neutral in terms of any presupposition concerning the response. The type which uses the ADV-VP form is found in some Southern Min and Wu dialects but also in certain Southwestern and Lower Yangtze Mandarin dialects of Anhui province, while the VP-NEG-(PARTICLE) form is characteristic of Hakka and Yue dialects.

Yue-Hashimoto is able to pinpoint different strata for these interrogative structures by comparing several colloquial Southern Min texts from the Ming and Qing dynasties (dating from the sixteenth century onwards) written in the Chaozhou

and Quanzhou dialects. Her analysis of these texts enables her to resolve apparent counterexamples where certain Min dialects possess all three strategies described above and thus seem to belie this basic Northern versus Southern distinction. She argues that the ADV-VP form using the adverbial interrogatives *kě* 可 or *qǐ* 岂 belongs to a residual premodern colloquial stratum found in certain Southern Min dialects such as Yilan in Taiwan and Shantou (Swatow) in Northeastern Guangdong province, China. This contrasts with the form of VP-NEG-(PARTICLE) which has been in use over many centuries and represents a standard and native Southern Min stratum, while VP-NEG-VP represents the non-native stratum which has been borrowed from Northern Chinese. Further comparisons with non-Sinitic languages are made: the ADV-VP form is commonly found in Tibeto-Burman while the VP-NEG form is typical of Kam-Tai, though languages in both families show use of the VP-NEG-VP strategy which overall appears to have the widest distribution in Sino-Tibetan, presumably through diffusion.

3.4 *Syntactic hybrids and metatypy*

Another consequence of language contact is the mixing or *hybridization* of syntactic forms. There are many clearcut cases of this in Sinitic languages where native and borrowed syntactic strategies are eclectically combined into the one new form. This is quite distinct from the situation known as metatypy (Ross 1996) where the syntactic configuration for a construction is borrowed from the prestige language entailing the calquing of its grammatical exponents by the appropriate morphemes. When metatypy occurs, it may replace the native strategy (if there is one - see §4.3 on complementizers below) or it may be used side-by-side with this native form, possibly in different speech levels or registers. Hong Kong Cantonese shows an unusual case of retention of the native form, in combination with metatypy and hybridization for the relative clause construction which I next examine.

Matthews and Yip (in press) have coined the useful term of *ditaxia* which refers to the parallel use of two syntactic structures in different registers. This lays the basis for analysing a third peculiar construction for the relative clause which has made a recent appearance in Hong Kong Cantonese. The two main relative clause structures can be thus described: colloquial Cantonese employs classifiers as relative markers as in (1) while formal Cantonese employs a structure using the possessive

ge^{3} which mirrors the use of Mandarin *de* as a relativizer. Compare the following two examples:

(1) Colloquial Cantonese: Relative Clause + DET + CL + HEAD NP

佢	唱	個	首	歌	好	好	听
Koei^{5}	*coeng*^{3}	*go*^{2}	*sau*^{2}	*go*^{1}	*hou*^{2}	*hou*^{2}	*teng*^{1}
3sg	sing	that	CL	song	very	good	listen

'the song she sings is very nice'

(2) Formal Cantonese: Relative Clause + GEN + HEAD NP

佢	唱	嘅	歌
Koei^{5}	*coeng*^{3}	*ge*^{3}	*go*^{1}
3sg	sing	PRT	song

'the song(s) she sings'

Typologically, the relational, including possessive, use of the classifier in colloquial Cantonese given in (1) is characteristic of other southern Chinese dialect groups such as Southern Min but also of Tai and Hmong-Mien languages, showing further evidence of the affinity among the Sinospheric languages (see Bisang 1992). The construction in (2) is an example of metatypy based on the prestige language, Mandarin. A third and innovative construction represents a hybridization of these two, where both the classifier and *ge*^{3} are present with the form [DET + CL + GEN (=g*e*^{3}) + N]:

(3) Hybridization: Relative Clause + DET + CL + GEN + HEAD NP

佢	唱	個	首	嘅	歌
Koei^{5}	*coeng*^{3}	*go*^{2}	*sau*^{2}	*ge*^{3}	*go*^{1}
3sg	sing	that	CL	PRT	song

'the song she sings'

At this point, a reasonable surmise might be that such examples of Cantonese show a lack of mastery over the newer Mandarinized form of the relative clause structure. It is interesting to learn, however, that the hybrid relative clause

construction tends to be used in more formal and public registers such as broadcasting and sermons, and is therefore classified as pseudo-High in register by Matthews and Yip. Possibly it serves a double purpose: on the one hand it has an emblematic status for Cantonese speakers - it can be used to show linguistic solidarity and Cantonese identity by retaining the classifier as a marker of the relative clause - yet on the other hand speakers retain the use of 'posh' Cantonese by means of the counterpart of the Mandarin relative clause, which uses the genitive marker *ge*[3] (see Aikhenvald, this volume, on the topic of emblematicity). An explanation involving syntactic hypercorrection does not appear to be relevant in this case.

A similar phenomenon can be observed in both Taiwanese Southern Min and Hakka for the comparative construction where the native strategy using an adverb 'more' is combined with the cognate for Mandarin *bǐ* 比'compare' (see Ansaldo 1999 on this kind of double-marking). Zhu (1990) also examines a hybrid structure for neutral Yes/No questions where an adverbial interrogative marker is used together with a VP-NEG-VP form. This is found in some Lower Yangtze Mandarin dialects, in the Suzhou dialect (Wu) and in the Shantou dialect (Southern Min) (see also §3.2.3). Similarly, Chappell (1992b and in press [c]) notes hybridization for the evidential (or experiential aspect) marker in Taiwanese Southern Min, where the native strategy of a preverbal marker *bat* 别 from the verb 'know' is combined with the verb enclitic *koè*, calqued on Mandarin *guò* 过'cross, pass through'.

3.5 *Sùliào pǔtōnghuà* 塑料普通话 'the plastic common language'

Wu (1992) describes a variety of Changsha Mandarin called *sùliào* or 'plastic' *pǔtōnghuà*' in which convergence is taking place between the local Xiang dialect and the official language, *pǔtōnghuà*. *Pǔtōnghuà* literally 'the common language', is based on the pronunciation of educated speakers of the prestige dialect of Beijing Mandarin in combination with the vocabulary and grammar of model works of vernacular literature written in Northern Chinese dialects. This definition was promulgated for the official language of China in 1955 (Chen 1999: 25). Speakers prefer to use Changsha Xiang but in official and formal situations they are encouraged to use *pǔtōnghuà*. Although the convergence is unidirectional - in the direction of Mandarin - it is far from complete.

When speakers accommodate to *pǔtōnghuà*, a language over which they may not have full command, a special tone correspondence is set up which neither belongs to the Changsha Xiang dialect nor to *pǔtōnghuà*, yet symbolizes that speakers have adopted an official speech level which is as close as they can possibly come to *pǔtōnghuà*. Even when non-standard lexical items are used, specific to the Xiang dialect, or speakers are unable to distinguish velar from alveolar nasal endings, let alone retroflexes from dental sibilants (as they should in standard Mandarin), the mere fact that they are using this special tone correspondence suffices for their speech to be considered 'official', that is, as plastic *pǔtōnghuà*.

By way of contrast, if speakers use the right lexicon and grammar for *pǔtōnghuà* but retain their own Changsha Xiang tone pattern, their speech remains irredeemably Changsha Xiang. The reason is as follows: first, it needs to be noted that Changsha Xiang has seven tones, whereas both plastic *pǔtōnghuà* and 'real' *pǔtōnghuà* have only four. Wu (1992: 137-138) explains how the correspondences between the Middle Chinese sources for the modern tones in standard Mandarin and colloquial Changsha Xiang differ. Changsha speakers base their rules for conversion of Xiang tones into plastic *pǔtōnghuà* on the historical relationships for their *own* dialect with Middle Chinese. It is this local interpretation which has created the special tone correspondences that act as a marker of plastic *pǔtōnghuà*.

In the final section, I examine the outcomes of language change: are pathways of grammaticalization triggered by a certain set of typological preconditions in the given language; is it due to areal diffusion of a morphosyntactic feature or, more broadly, merely attributable to common language universals of grammatical change?

4. Shared grammaticalization pathways in Sinitic, areal diffusion and language universals

In this section, I examine five sets of data in Sinitic: the source of the diminutive suffix, the feature of negative existential verbs 'there is not/there are not', the development of complementizers from verbs of saying, adversative passives and some constructions which express inalienable possession. Some of these phenomena unify Sinitic as a family while others bear witness to the grouping of languages in the Southeast and East Asia zone as a *Sprachbund* or linguistic area. In this section, the attempt is made to distinguish which features represent a pathway of grammaticalization that is crosslinguistically unremarkable, which are

the result of areal diffusion, and which could be seen as special typological features of Sinitic languages.

4.1. *Early Southern Min dialect grammar and evidence for grammaticalization: the diminutive*

Early 17th century texts on Southern Min dialects provide an invaluable source for the diachronic study of the grammar of their modern counterparts in that they are largely written in the special dialect characters for vernacular Hokkien. Below, I compare the diminutive of modern Southern Min dialects such as Taiwanese and Amoy (Xiamen) with those found in the *Arte de la lengua Chiõ Chiu* (1620), a grammar on the same type of dialect written in Spanish.[iv]

In Sinitic languages, the diminutive has its source in various morphemes for 'son' which may have 'child' as a secondary meaning. A morpheme for 'child' is the common source crosslinguistically for diminutives (see Heine et al 1991: 79-88; 1993: 38). For example, Mandarin uses the suffix ㄦ< ér 兒 'son' while Cantonese employs tone sandhi, changing the citation tone to high rising tone, the cheshirization of an earlier segmental morpheme meaning 'son'. Cheshirization refers to the attrition of segmental phonemes, which leave a mere trace of their former phonetic substance, such as the tone.[v] In Taiwanese Southern Min, the diminutive is formed with the suffix *-á.* It can be related to the lexeme for 'son', 子 *kǐa*^{n}, used in the *Arte* (1620: 2b, 11a,12b) and to *kiá*^{n} 'son' in contemporary Taiwanese and Amoy, for which the character 囝 is used as well.[vi] Note that the stem of the word used for 'child' in the *Arte* —简仔 *kǐn nǐa* (1620: 15) or 囝仔 *gín-á ~ gín-ná* in contemporary Taiwanese — cannot be the source for this diminutive on phonological grounds (see Lien 1998).

In the early 17th century grammar of Southern Min, the following description is given for the diminutive (1620:10):

(4) *Arte de la lengua Chiõ Chiu* (1620)

"The diminutive is formed with the final particle *ia* or *nia* or *guia*:

kéiguìa	圭仔	'little chicken' [*pollito*]
bôguìa	帽仔	'little hat' [*sonbrerillo*]
tóguìa	刀仔	'little knife'[*guedillito*]."

In contemporary Taiwanese, the three corresponding words are *ke-á* 'chicken, little chicken'; *bō-á* 'hat' and *to-á* 'knife, small knife' respectively, indicating partial bleaching of the diminutive feature.[vii]

I suggest that in this early grammar of Southern Min, the *Arte*, an incipient stage of development for the diminutive can be viewed, where its form can still be clearly related to the morpheme for 'son', unlike contemporary Southern Min where the form has atrophied to *-á* and can be used not only as a diminutive but also as a marker for the noun category:

(5) Taiwanese Southern Min:

一	张	桌仔	合	两	张	椅仔
chit	*tè*	*toh-á*	*kap*	*nīng*	*tè*	*í-á*
one	CL	table	and	two	CL	chair

'a table and two chairs' (not: 'a small table and two small chairs')

It is interesting to find that the lexeme *kiá*^{n} can nonethelesss still be used as a kind of suffix to mark the young of animal species, postposed after the reduced diminutive form:

(6)

牛仔囝	狗仔囝
gû-á-kiá^{n}	*káu-á-kiá*^{n}
ox-DIMN-offspring	dog-DIMN-offspring
'calf'	'puppy'

Further support for the proposed grammaticalization pathway of 'son' > DIMINUTIVE comes from Yang (1991: 166) who points out that the diminutive suffix in the Chaozhou dialect of Southern Min retains the full form of *kiá*^{n}.

(7) 鼎仔 Chaozhou:*tia*^{n} *kiá*^{n} contrasting with Xiamen, Zhangzhou, Taiwanese: *tia*^{n}*-á*.
'a small cooking pot'

Yang also quotes the Tang poet 顾况 Gu Kuang who annotates the character 囝 , pronounced with an alveopalatal initial / *tɕiǎn/* in modern Mandarin, as having the meaning 'son' in colloquial Min in §13 of his poem 上古之什 *Shànggǔzhīshé*.

(8) 囝 音 蹇 闽 俗 呼 子 为 囝.
Jiǎn yīn jiǎn mǐn sù hū zǐ wéi jiǎn
(word) sound jian Min custom call son as jian
"The sound of this character 囝 is *jiǎn*, the Min usually call 'son' *jiǎn*."

The more general case of semantic change from 'child' to diminutive morpheme is well-attested in other languages of the world, for example, in Jurafsky (1996) and Heine et al (1993: 38) while the use of diminutives with probable source morphemes in sex-specific 'son' is characteristic of Sinitic (for more data, see Huang 1996). The *Arte* provides the hard evidence for this semantic change into a diminutive suffix, affecting the morpheme 'son' in Southern Min (see also Chappell in press [b]). Given the widespread occurrence of the first type of conceptual shift cross linguistically, I conclude that while this more semantically specific case may be a shared development in Sinitic languages, it only partially characterizes it typologically.

4.2. *Negative existential constructions: 'there is no/there are no'*

Southern Sinitic languages display a large number of negative morphemes which can be used to negate propositions at clause-level. Furthermore, the semantic space for negation is carved up by subtle modal and aspectual nuances. In particular, Southern Min languages show a highly differentiated set of negative adverbs, most being fused forms combining one of the first two negatives listed in Table 3 with various modal verbs and showing different degrees of bondedness:

Table 3: Taiwanese Southern Min negative markers:

bô+ V	無	Negation of perfective contexts, attributive predicates
m̄+ V	唔	Negative marker for property verbs, imperfective contexts & unwillingness to V
(iá) bē+ V	(猶)未	Negation of expectation: 'have not (yet) V-ed'
boē+ V	勿会	Negation of ability/possibility to V: 'unable to V'
boài+ V	勿爱	Negation of perfective desiderative: 'didn't want to V'
m̄mài+V	唔爱	Negation of imperfective desiderative: 'don't want to V'
mài+ V	莫	Negative imperative: 'Don't V!'
m̄mó+ V	唔爱	Negative hortative: 'You shouldn't V!'
m̄bién + V	唔免	Negation of necessity: 'You don't need to V'

In Sinitic, it is typically the marker used to negate perfective clauses which also has a fully verbal use meaning *there is no Y/there are no Y* with one nominal argument. This set of verbs in Southern Sinitic can also occur in a transitive syntactic frame as the negative possessive verb: *X has no Y*. I describe the semantic and syntactic features of negative existential verbs in more detail in Chappell (1994) and observe that their prior lexical meaning is often 'lose' as exemplified by (9) for Cantonese 冇 *mou*^{5} where the meaning is ambiguous between the two uses:

(9) Cantonese:

已经	冇	哩	個	嘅	权势	啊
yi^{5}*ging*^{1}	***mou***^{5}	*lei*^{5}	*goh*^{3}	*ge*^{3},	*kuen*^{4}*sai*^{3}	*a*^{1}
already	**NEGV**	this	CL	PRT	power	PRT

'(This prime minister) had already lost his power.' OR:
'The prime minister no longer had any power.'

Standard Mandarin does not possess such a negative existential or negative possessive verb. It must use the negative perfective marker *méi* preposed before the verb *yǒu* 'there is', shown in (10).

(10) Mandarin:

没	(有)	人	了
méi	*(yǒu)*	*rén*	*le*
NEG	(there:be)	person	CRS

'There's nobody here.'

Omission of *yǒu* 'there is' is possible but should not be confused with an analysis of *méi* as a monomorphemic negative existential verb (which it is not), since *yǒu* can always be added back in. It appears that the same situation applies in many Tibeto-Burman languages where a negative adverb or prefix beginning with *m-* is used (see Matisoff 1991: 388, 393-394), and also in Thai. In other words, these languages similarly do not have a special negative existential verb. Hence, this is a Southern Sinitic feature, not attested in either northern Chinese or evidently in the other half of the Sino-Tibetan language family. It is neither a Sinospheric typological feature nor a pan-Sinitic one. Nor is it well-documented cross-linguistically, given that Payne (1985) discusses this type of negation for only a few Austronesian languages but does not include it as a negation type.

4.3 *Complementizers*

In Taiwanese Southern Min, a complementizer similar in function to English *that* has grammaticalized out of the verb 'to say' *kóng* 講. Matisoff (1991: 398-400) describes this path of grammaticalization as an example of the general category of verbs developing into verb particles in Southeast Asian languages, represented by Thai, Khmer and Lahu. Like these three languages, the Southern Min verb 'say' is also used at the end of a non-final clause and before the intonation break to introduce the complement clause. It is not fully grammaticalized since it may be omitted. Moreover, it forms a kind of verb complex with the preceding matrix verb which must belong to one of the following verb classes: speech act, cognition or perception, and it directly introduces the embedded clause, as in (11):

(11) Taiwanese Southern Min

遐	個	敌对	的	武将	共	笑	講
Hia	*ê*	*<MC: díduì >*	*ê*	*búchiòng*	*kā*	***chhiò***	***kóng,***
that	CL	opposing	L	general	PRETR	**laugh**	**SAY$_{\text{that}}$**

这	是	号作	<J:猴面	冠者>.
che	*sì*	*hō-tsò*	*<J: Sarumen*	*Kanja >.*
this	be	name:as	monkey:face	youngster

'Those generals who opposed him mocked him (General Toyotomi) as the one who should be called "monkey-face boy".' [*Japanese tales* 629-630]
(Note: MC = Mandarin Chinese insert; J = Japanese insert)

In this first stage of grammaticalization, when *say* verbs are used as quotative markers, the lexical meaning is not completely bleached. Examples such as *chhiò kóng* could still be rendered as 'laughed (at him) saying' while in the second stage where *kóng* is used with cognitive verbs such as *siū*^{n} 'think', its literal meaning is less plausible: 'think saying'. The putative path of development is outlined in Chappell (in press [e]) in addition to other grammaticalized or partially grammaticalized uses of *kóng* as a metalinguistic marker of explanation; an evidential marker of hearsay; a component of a compound conditional marker; a topic introducer and as a clause-final marker of assertions and warnings. It has not yet developed a purposive function, which may indicate that certain of its several grammaticalization pathways are relatively 'young' (Bernd Heine pers. comm.).

There has been only very little study of this phenomenon in typological work on Sinitic languages to date. In Chappell (in press [e]), I show that this development has proceeded as far as the quotative stage in some Yue and Wu dialects and to a lesser extent in standard Mandarin. For the Yue dialect of Cantonese, ample evidence can be found of the use of *wa*^{6} 'to speak' in conversational and narrative texts where it functions as such a quotative marker with speech act verbs. Note, however, that *wa*^{6} does not form a verb complex with the preceding speech act verb: this is clear in the fact that it can be separated from the verb by a noun denoting the direct object:

(12) Cantonese

赞	哩	個	男仔	话...
jaan^{3}	*lei*^{5}	*goh*^{3}	*laam*^{4}*jai*^{2}	*wa*^{6}...
praise	this	CL	young:man	say…

"(She) praised this young man saying..."

Although a verb complex with 'say' as V_2 is not a possible strategy for introducing complement clauses in standard Beijing Mandarin, or *pǔtōnghuà* (as opposed to such a use for quotations), it is in the regional variety known as Taiwanese Mandarin. It is striking that Taiwanese Mandarin does not choose the cognate verb for *kóng*, which is *jiǎng* in Mandarin, to create the new syntactic calque but instead makes use of its functional equivalent, the high frequency verb *shuō* 说, in the configuration NOUN - SUBJECT - $VERB_1$ - *shuō* + CLAUSE:

(13) Taiwanese Mandarin

那	我	希望	说		这	個	愿望
nà	*wǒ*	*xīwàng*	*shuō*		*zhèi*	*ge*	*yuànwàng*
CONJ	1SG	hope	SAY_{comp}	this	CL	wish	

很	快	就	到	了
hěn	*kuài*	*jiù*	*dào*	*le*
very	quickly	then	arrive	PFV

'So I hope that this wish will be realised very soon.'

(14) Beijing Mandarin

*我	希望	说
* *wǒ*	*xīwàng*	*shuō*
1SG	hope	say

However, this does not provide supporting evidence just for the North-South divide for Sinitic languages: it appears that Sinitic is encircled by language families and

language isolates (such as Japanese and Korean) that all possess complementizers which have developed from verbs of saying. This feature has been described in the relevant literature for individual languages belonging to Tibeto-Burman, Tai-Kadai, Hmong-Mien, Indic, Dravidian and Altaic (see Matisoff 1991, Saxena 1988).

Since this semantic change is also crosslinguistically well-attested (it occurs widely in various language families of Africa - see Frajzyngier 1996 for Chadic, Amberber 1995 for Amharic, Heine et al 1991: 216-7; 246-7; Heine et al 1993: 190-8 for a larger sample of languages), it seems that the grammaticalization of *kóng* into a complementizer in Taiwanese Southern Min is most likely a language internal development. It has simply drawn on its own resources (Dixon 1997) to recreate a syntactic device which was in fact available in Classical and Middle Chinese, as attested in the written register.

Indeed, earlier periods of written Chinese made use of verbs of saying such as *yuē* 曰(Classical Chinese) and *dào* 道(Medieval Chinese) as quotative markers, although not as fully-fledged complementizers (described in Chappell in press [e]). This means that not only does Sinitic have its own inherited language-internal devices upon which to analogize but it also has access to patterns and processes which can be imitated from surrounding unrelated language families.

It seems that this has taken place in recent times for sister languages within Sinitic, the case in point being the calquing of the Taiwanese Southern Min complementizer into Taiwanese Mandarin. This is an unusual development in terms of the direction of metatypy from a less prestigious to a more prestigious language and note that there are many other examples of Taiwanese Southern Min constructions which have been borrowed into the Taiwanese variety of Mandarin (see Kubler 1985). This probably reflects linguistic creativity in transferring favoured syntactic forms and devices into Mandarin where gaps exist, rather than a negative description in terms of interference from L_2.

Further research on dialect materials would be in order to show irrefutable evidence for the view that the development of a complementizer in Taiwanese Southern Min is a purely independent innovation, triggered however by a combination of factors: a conducive environment in terms of areal typological features and the existence of appropriate language internal characteristics.

Unlike the case for negative existential verbs, the existence of a complementizer in Southern Min and some Wu and Yue dialects tallies well with Matisoff's inclusion of Southern Sinitic in the Southeast Asian linguistic area. The theoretical problem remains however of distinguishing areal diffusion from a putative language universal for the development of complementizers from verbs of saying, given the right typological preconditions.

4.4. *Adversative Passives*

Matisoff (1991) points out that verbs of giving typically develop into causatives and benefactives in Southeast Asian languages. In Southern Sinitic languages, verbs of giving are also used to form the passive construction. For example, most Hakka dialects use the high frequency verb *pun* ^{44}'to give' as both the passive and the benefactive marker, while Cantonese does the same with *bei*^{2} < 'give'.

A further characteristic feature of passives which unites Sinitic is that the colloquial forms are both adversative and agentful. This appears to be an unusual development for 'give' (compare this with data in Heine et al 1993: 97-103). Such a description applies to standard Mandarin as well where only the *bèi* passive has an agentless form although it has lost its adversative feature in some contexts. Note that the *bèi* passive belongs to more formal discourse, in contrast with the agentive colloquial passives formed by *jiào* 'make' and *ràng* 'let' (see Chappell 1986).

Norman (1982: 245) observes that these two Northern Chinese passives formed with the causative verbs *jiào* 'make' and *ràng* 'let' are unique amongst Sinitic languages, as opposed to the use of verbs of giving. He argues that this is not an independent development in Mandarin but rather is due to Manchu superstrate influence on Chinese. In Manchu and other Altaic languages, the same structure can be used for both passive and causative meanings. In support of this view, an earlier study by Hashimoto (1987: 46) contrasts standard Mandarin with Mandarin dialects on the periphery of the Northern Chinese zone which continue to use verbs of giving as passive markers. This suggests that give verbs as passive markers are an older feature.

The adversative feature appears to be an areal feature as not only do Southeast Asian languages such as Thai and Vietnamese show this restriction, but also Japanese (see Shibatani 1994). Hence, there are different allegiances for each of these features: some evince the North-South divide in Sinitic (verbs of giving versus causative verbs

used as passive exponents), some are relevant to the Southeast and East Asian area (the adversative feature), while this particular development for 'give' is possibly specific to Southern Sinitic within the Asian zone, and is quite rare crosslinguistically (Bernd Heine pers. comm.).

4.5 *Possession*

4.5.1 *Pronominal systems and inalienable possession*

In general there are no separate morphological classes for alienable and inalienable possession in Sinitic languages; nonetheless, there is a weaker reflection of this distinction in the fact that genitive marking is facultative for kin relationships as well as other important social relationships, body parts and spatial orientation, particularly when the possessor is pronominal (see Chappell & Thompson 1992 on Mandarin genitives):

PRONOUN POSSESSOR (genitive marker) NOUN POSSESSED

(15) Mandarin

你	(的)	母亲	先生	(的)	耳朵	里
nǐ	*(de)*	*mǔqin*	*xiānsheng*	*(de)*	*ěrduo*	*li*
2sg	(GEN)	mother	teacher	(GEN)	ear	in
'your mother'			'in the teacher's ears'			

Hakka is unusual within Sinitic in having a special portmanteau genitive form for pronominal possessors which can be considered as a kind of case marker:

Table 4: Meixian Hakka pronouns

	Nom/Acc	**Genitive**
1SG	ŋai^{11} 崖	ŋa^{44}
2SG	ŋ^{11} 你	ŋia^{44}
3SG	ki^{11} 佢	kia^{44}

These special genitive forms are not generally used, however, with inanimate nouns such as ‘fountain pen’ in (17) but, again typically, with kin as in (16):

(16) 我 老弟
ŋa^{44} *lau*^{31}*-t‘ai*^{44}
1sg younger:brother
‘my younger brother’

With inanimate nouns, as in example (17), the genitive marker *ke* is used with the Nom/Acc form of the pronominal possessor:

(17) 我 嘅 钢笔
ŋai^{11} *ke*^{53} *kongbit*^{11}
1sg GEN pen
‘my fountain pen’ (**ŋa*^{44} *Ø kong*^{53}*bit*^{11})

This semi-covert distinction is reflected more clearly in syntax in the form of the double patient construction, discussed next.

4.5.2. *Double patient constructions*

The double patient construction is shared by all Sinitic languages. It is syntactically unusual in that the intransitive process verb appears to take two arguments, one more than the verb valency should allow, recalling the ‘one-too-many-argument’ problem described in Shibatani (1994). The two arguments of the intransitive verb designate possessor and possessum. Furthermore, the nouns in this possessive relationship occur non-contiguously and belong to different constituents. Specifically, the possessor appears in the canonical position for grammatical subject (S) clause-initially, while the possessum appears postverbally in the canonical object position (O). The verb must be a so-called ‘unaccusative’ non-volitional one such as ‘go red’, ‘go white’, ‘limp’, ‘increase’ (literally: ‘become more’), ‘fall out’ or ‘rot’, which takes a semantic undergoer as its subject. An example of this construction from Cantonese is given with its structural formula:

Double patient construction:

NOUN$_{\text{POSSESSOR}}$ VERB$_{\text{INTRANSITIVE}}$ NOUN$_{\text{PART/KIN TERM}}$

(18) Cantonese Yue

棵	树	落	咗	好	多	叶
Poh^{1}	*sue*^{6}	*lok*^{6}	*joh*^{2}	*ho*^{2}	*doh*^{1}	*yip*^{6}
CL$_{\text{REF}}$	tree	fall	PFV	very	many	leaf

'That tree has lost many leaves [more literally: The tree fell very many leaves].'

In Chappell (1999), I argue that the relationship of inalienable possession licenses the use of two arguments with an intransitive verb. It can only be used for part-whole relations and, in a more restricted fashion, for kin. While this construction is a shared feature of Sinitic, as with the study of complementizers, it has not been extensively researched. The same situation applies for Southeast Asian languages: it is not possible in Lahu (Matisoff pers. comm.) but a similar construction appears to exist in Lao (Nicholas Enfield pers. comm.). At this stage, it is difficult to determine if such a construction is typologically defining for Sinitic.

5. Conclusion

The family tree model appears to work reasonably well for Sinitic as far as phonology and some aspects of morphology are concerned; nonetheless, this only accounts for a small part of a much more complex linguistic picture: the family tree model is unable to capture the effect of successive waves of Mandarinization of Southern Sinitic languages, stratifying lexical and syntactic components as shown in §3 for nominal affixes in Southern Min and interrogative constructions in Southern Sinitic languages. Nor can it handle the cases where convergence is well under way with the Mandarinization of Changsha Xiang, though be it by means of an intermediate language known as *sùliào* or 'plastic' *pǔtōnghuà*. The initial stages of this process of convergence includes widespread occurrence of metatypy and hybridization of syntactic forms in Sinitic, as illustrated by the example of Hong Kong Cantonese relative clause constructions. Hence, a more delicate and subtle treatment of the question of genetic affiliation is needed.

Note that the processes of metatypy and convergence may not always be in the direction of the official language of prestige: in Taiwan, massive calquing and metatypy from Southern Min into Taiwanese Mandarin is taking place, as briefly described for the use of complementizers. It can be conjectured that this is because Southern Min, and not Mandarin, is emblematic of current loyalties and serves as a 'badge' of being Taiwanese. Such developments involving language contact cannot be easily captured in terms of genetic affiliation while they would skew the data in any study using the comparative method.

§4 investigated the problems of determining whether certain syntactic and morphological features could be the outcome of shared developments in a language family, while others are simply the result of areal diffusion or are common crosslinguistically, requiring no particular typological preconditions. Five areas of morphosyntax were thus examined: similarities and differences with crosslinguistically attested pathways of language change were described for the five areas of diminutives, negatives, complementizers, passives and inalienable possession with additional language-specific features being noted in some of these cases: first, diminutive suffixes in Sinitic were shown to have their source not in a morpheme for 'child' but in the more sex-specific 'son' (which nonetheless may have the secondary meaning of 'child' or 'offspring' in some, but not all, of these languages).

Second, the large inventory of negative markers in Sinitic languages was also briefly described. The fact that these grammaticalize out of a fusion of basic negative markers and modal verbs appears to be typologically unusual in the light of crosslinguistic studies such as Payne (1985). Third, it was observed that complementizers with a source in a verb of saying are common crosslinguistically although the Southern Min development is relatively young, while that for Cantonese Yue is only in an incipient stage. Fourth, passive exponents in Southern Sinitic languages were described as typically having their source in verbs of giving, yet it is unusual crosslinguistically for this type of passive to also express adversity and to require an agent. Fifth, for the expression of inalienable possession at the level of nominal syntax, the Meixian Hakka dialect presents an interesting and typologically uncharacteristic case for Sinitic since it uses a portmanteau morpheme in precisely this function. This distinction is typically covert in most Sinitic languages, and can at best be only detected for syntactic constructions such as the one described above with intransitive verbs and two patient nouns. Yet different pronouns and nominal

constructions to code alienable versus inalienable possession are very common crosslinguistically (see Chappell & McGregor 1995).

To adequately reconstruct the history of a language family, a model is needed which is significantly more sophisticated than the family tree based on the use of the comparative method. It needs to incorporate the diffusion and layering process as well as other language contact phenomena such as convergence, metatypy and hybridization. The desideratum is a synthesis of all the processes that affect language formation and development.

References

Amberber, M. 1995. 'Functional verbs in predicate formation. Event type hierarchy and grammaticization', *Proceedings of the Twenty-First Annual Meeting of the Berkeley Linguistics Society*, 1-11.

Ansaldo, U. 1999. *Comparative constructions in Sinitic: areal typology and grammaticalization in Chinese languages*. PhD thesis. University of Stockholm, Sweden.

Arte de la lengua Chiõ Chiu. 1620. Ms. from Biblioteca Universitaria Provincial Barcelona entitled *Gramatica China*. Anonymous author(s). Manila.

Baxter, W. H. III. (1992). *A Handbook of Old Chinese Phonology.* Berlin: Mouton de Gruyter.

Beijing Daxue Zhongguo Yuyan Wenxue Xi Yuyanxue Jiaoyanshi [Linguistics section of Beijing University Chinese Languages and Literatures Department] (eds.). 1964. *Hanyu fangyan cihui* [A lexical list for Chinese dialects]. Beijing: Beijing University.

Bisang, W. 1992. *Das Verb im Chinesischen, Hmong, Vietnamesischen, Thai und Khmer.* Tübingen: Gunter Narr.

Bisang, W. 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.

Chappell, H. In press [a]. Editor of *Sinitic grammar: synchronic and diachronic perspectives.* Oxford: Oxford University Press.

Chappell, H. 1986. 'Formal and colloquial adversity passives in standard Chinese', *Linguistics* 24.6: 1025-1052.

Chappell, H. 1992a. 'Causativity and the *ba* construction in Chinese', pp. 509-530 of *Partizipation: Das sprachliche Erfassen von Sachverhalten*, edited by Hansjakob Seiler & Walfried Premper. Tübingen: Gunter Narr,.

Chappell, H. 1992b. 'Towards a typology of aspect in Sinitic languages', *Zhongguo Jingnei Yuyan ji Yuyanxue*: *Hanyu Fangyan* [Chinese Languages and Linguistics: Chinese dialects] 1.67-106.

Chappell, H. 1994. 'A typology of negative verbs of existence and possession in Sinitic languages', pp. 303-311 of *Current Issues in Sino-Tibetan Linguistics*, edited by Hajime Kitamura, Tatsuo Nishida & Yasuhiko Nagano. Osaka: National Museum of Ethnology.

Chappell, H. 1999. 'The double unaccusative in Sinitic languages', pp. 197-232 of *External Possession and related noun incorporation constructions* (Typological Studies in Language Series 39), edited by Doris L. Payne & Immanuel Barshi. Amsterdam: John Benjamins,.

Chappell, H. In press [b]. 'Dialect grammar in two early modern Southern Min texts: A comparative study of dative *kít,* comitative *câng* and diminutive *–guià*', *Journal of Chinese Linguistics* 28.2.

Chappell, H. In press [c]. 'A semantic typology of evidential markers in Sinitic languages', *Sinitic grammar: synchronic and diachronic perspectives*, edited by H.Chappell. Oxford: Oxford University Press.

Chappell, H. In press [d]. 'A brief history of Chinese dialects', *Sinitic grammar: synchronic and diachronic perspectives*, edited by H.Chappell. Oxford: Oxford University Press.

Chappell, H. In press [e]. 'Grammaticalisation of the verb *kóng* 'say' in Taiwanese Southern Min', *Studies in Chinese Historical Morphology and Syntax* (Studies in Language Companion Series), edited by Charles N. Li & Alain Peyraube. Amsterdam: John Benjamins.

Chappell, H. & W. McGregor. 1995. Editors of *The grammar of inalienability. A typological perspective on body part terms and the part-whole relation.* (Empirical approaches to language typology 14). Berlin: Mouton de Gruyter.

Chappell, H. & S. A. Thompson. 1992. 'The semantics and pragmatics of associative *de* in Mandarin discourse', *Cahiers de Linguistique Asie Orientale* 21: 199-229.

Chen, P. 1999. *Modern Chinese: History and sociolinguistics.* Cambridge: Cambridge University Press.

Cheung, S. H-n. 1992. 'The pretransitive in Cantonese', *Zhongguo Jingnei Yuyan ji Yuyanxue. Hanyu Fangyan* [Chinese languages and linguistics: Chinese dialects] 1.241-303.

Dixon, R.M.W. 1997. *The rise and fall of languages*. Melbourne: Oxford University Press.

Frayzyngier, Z. 1996. *Grammaticalization of the complex sentence: A case study in Chadic.* Amsterdam: John Benjamins.

Grootaers, W. 1943. 'La géographie linguistique en Chine - necessité d'une nouvelle méthode pour l'étude linguistique du chinois', *Monumenta Serica* 8. 103-166.

Hashimoto, M. 1973. *The Hakka Dialect: A linguistic study of its phonology, syntax and lexicon.* Cambridge: Cambridge UP.

Hashimoto, M. 1974. 'Aspect and tense in Asian languages', *Computational Analyses of Asian and African Languages* 1.15-25.

Hashimoto, M. 1976a. 'Language diffusion on the Asian continent: Problems of typological diversity in Sino-Tibetan', *Computational Analyses of Asian and African Languages* 3.49-65.

Hashimoto, M. 1976b. 'The double object construction in Chinese', *Computational Analyses of Asian and African Languages* 6.33-42.

Hashimoto, M. 1986. 'The Altaicization of Northern Chinese', pp. 76-97 of *Contribution to Sino-Tibetan Studies*, edited by J. McCoy and T. Light. Leiden: E.J. Brill.

Hashimoto, M. 1987. 'Hanyu beidongshi de lishi, quyu fazhan [The history and areal development of the Chinese passive]', *Zhongguo Yuwen* 1.36-49.

Hashimoto, M. 1992. 'Hakka in Wellentheorie perspective', *Journal of Chinese Linguistics* 20.1: 1-49.

Heine, B., U. Claudi & F. Hünnemeyer. 1991. *Grammaticalization: a conceptual framework.* Chicago: University of Chicago Press.

Heine, B., T. Güldemann, C. Kilian-Hatz, D. Lessau, H. Roberg, M. Schladt & T. Stolz. 1993. *Conceptual shift: a lexicon of grammaticalization processes in African languages.* [Afrikanistische Arbeitspapiere 34/35]. Cologne: Institut fur Afrikanistik.

Huang B. 1996. Editor of *Hànyǔfāngyán yufǎlèibiǎn* [Concordance of Chinese dialect grammar]. Qingdao: Qingdao Publishers.

Jurafsky, D. 1996. 'Universal tendencies in the semantics of the diminutive', *Language* 72.3: 533-578.

Karlgren, B. 1915-1926. *Etudes sur la phonologie chinoise. Archives d'études orientales,* v.15 (in 4 parts). Leiden: E.J. Brill.

Karlgren, B. 1940. 'Grammata Serica, script and phonetics in Chinese and Sino-Japanese', *Bulletin of the Museum of Far Eastern Antiquities* 12.1-471.

Karlgren, B. 1957. 'Grammata serica recensa', *Bulletin of the Museum of Far Eastern Antiquities* 32.1-332.

Kubler, C. 1985. *The development of Mandarin in Taiwan: A case study of language contact.* Taipei: Student Book Company.

Lien, C. 1998. 'Taiwan Minnanyu cizhui 'a' de yanjiu [A study of the affix 'a' in Taiwanese Southern Min]', *Di'er jieTaiwan Yuyan Guoji Yantaohui Lunwen Xuanji*. Taipei: Crane Publishing Co., 465-483.

Lien, C. In press. 'Competing morphological changes in Taiwan Southern Min', *Sinitic grammar: synchronic and diachronic perspectives*, edited by H.Chappell. Oxford: Oxford University Press.

Matisoff, J. 1991. 'Grammatization in Lahu', pp. 383-453 of *Approaches to Grammaticalization*, Vol 2, edited by Elizabeth Closs Traugott & Bernd Heine. Amsterdam: John Benjamins.

Matthews, S.& V.Yip. In press. 'Aspects of contemporary Cantonese grammar: The structure and stratification of relative clauses', *Sinitic grammar: synchronic and diachronic perspectives*, edited by H.Chappell. Oxford: Oxford University Press.

Norman, J. 1982. 'Four notes on Chinese-Altaic linguistic contacts', *Tsinghua Journal of Chinese Studies* 14.1-2: 243-247.

Norman, J. 1988. *Chinese.* Cambridge: Cambridge UP.

Payne, J. 1985. 'Negation', pp. 197-242 of *Language typology and syntactic description*, Vol 1, *Clause structure*, edited by Timothy Shopen. Cambridge: Cambridge University Press.

Ross, M. 1996. 'Contact-induced change and the comparative method: Cases from Papua New Guinea', pp. 180-217 of *The comparative method reviewed:*

irregularity and regularity in language change, edited by M. Durie & M. Ross New York: Oxford UP.

Sagart, L. 1988. 'On Gan-Hakka', *Tsing Hua Journal of Chinese Studies New Series.* 18.1: 141-160.

Sagart, L. 1997. 'On distinguishing Hakka and non-Hakka dialects', *Journal of Chinese Linguistics* 26.2: 281-302.

Sagart, L. 1999. *The Roots of Old Chinese.* Amsterdam: John Benjamins.

Sagart, L. In press. 'Vestiges of Old Chinese derivational affixes in modern Chinese dialects', *Sinitic grammar: synchronic and diachronic perspectives*, edited by H.Chappell. Oxford: Oxford University Press.

Saxena, A. 1988. 'The case of the verb "say" in Tibeto-Burman,' *Proceedings of the Berkeley Linguistics Society* 14.375-388.

Serruys, P. L.-M. 1943. 'Philologie et linguistique dans les etudes sinologues', *Monumenta Serica* 8.167-221.

Shibatani, M. 1994. 'An Integrational Approach to Possessor Raising, Ethical Datives, and Adversative Passives', *Proceedings of the 20th Annual Meeting of the Berkeley Linguistics Society,* 461-486.

Ting, P-H. 1983. 'Derivation time of colloquial Min from Archaic Chinese', *Bulletin of the Institute of History and Philology* 54.4: 1-14.

Van der Loon, P. 1966. 'The Manila incunabula and early Hokkien studies', Part I. *Asia Major. New Series* 12: 1-43.

Van der Loon, P. 1967. 'The Manila incunabula and early Hokkien studies', Part II. Asia Major. New Series 13.95-186.

Wu, Y. 1992. *Studies in Chinese Lexicology: Investigations into the Xiang dialect.* PhD thesis. Melbourne: The University of Melbourne.

Yang, H-f. 1991. *Taiwan Minnanyu Yufagao* [A grammar of the Minnan language of Taiwan]. Taipei: Ta An Press.

Yuan, J. 1989 [1960]. Compiler of *Hanyu fangyan gaiyao* [An outline of Chinese dialects]. Beijing: Wenzi Gaige.

Yue-Hashimoto, A. 1991. 'Stratification in comparative dialectal grammar: A case in Southern Min', *Journal of Chinese Linguistics* 19.2: 139-171.

You, R. 1992. *Hanyu fangyanxue daolun* [Chinese dialectology]. Shanghai: Shanghai Jiaoyu Press.

Zhu, D. 1990. 'Dialectal distribution of V-neg-VO and VO-neg-V interrogative sentence patterns', *Journal of Chinese Linguistics* 18.2: 209-230.

ENDNOTES

[i] Thanks to the following colleagues for comments and critique: Sasha Aikhenvald, Tim Curnow, Bob Dixon, Nick Enfield, Geoffrey Haig, Bernd Heine,Tania Kuteva, Randy LaPolla, Jim Matisoff, Alain Peyraube, Laurent Sagart and participants at the workshop on "The connection between areal diffusion and the genetic model of language relationship" held at the Research Centre for Linguistic Typology at ANU in August 1998.

This research forms part of an Australian Research Council Large Grant project "A semantic typology of complex syntactic constructions in Sinitic languages" (1997-1999).

[ii] I am indebted to Laurent Sagart for this clarification of Karlgren's approach.

[iii] For more discussion of the reconstructions for either Old or Middle Chinese, see Norman (1988), also Baxter (1992) and Sagart (1999) and in press.

[iv] This work was most likely a collaborative effort of Spanish Dominican missionaries and Chinese interpreters living in a Chinese Sangley community near Manila in the late 16th and early 17th centuries. On phonological grounds, Van der Loon identifies the dialect used in these manuscripts as the vernacular of Hai-cheng as spoken around the turn of the seventeenth century (1967:132). He shows conclusively that it differed in certain phonological features from the dialect of Zhangzhou city, to which prefecture this harbour town belonged. It appears that the Sangleys or Chinese traders had migrated from this port in southern Fujian province during the late 16th century, with many eventually settling in and around Manila.

[v] See Matisoff (1991) for more on 'cheshirization', to whom we owe the coining of this evocative term.

[vi] This morpheme *kiã*^{n} 'child, son' is in fact used to exemplify the tone category which is accompanied by nasalization, according to the missionaries' classification. Note that in the Spanish romanization *k-* is used interchangeably with *gu-*and *qu-* for the unaspirated voiceless velar plosive initial /k/, as seen in the diminutive forms given in (4). Furthermore, nasalization has not been marked for these diminutive forms, suggesting that it had already been lost at this stage, in contrast to its lexical use as 'child'.

[vii] Note that only one of the variants listed by the Arte is illustrated by the examples in (4). This is discussed further in Chappell in press [b].