Grammaticalization processes in the languages of South Asia

Coupe, Alexander R.

2018

Coupe, A. R. (2018). Grammaticalization processes in the languages of South Asia. In H. Narrog, & B. Heine (Eds.), *Grammaticalization from a Typological Perspective* (pp. 189-218). doi:10.1093/oso/9780198795841.003.0010

<https://hdl.handle.net/10356/146316>

<https://doi.org/10.1093/oso/9780198795841.003.0010>

© 2018 Alexander R. Coupe. First published 2018 by Oxford University Press. All rights reserved.

Downloaded on 26 Apr 2023 16:16:51 SGT

Grammaticalization processes in the languages of South Asia

ALEXANDER R. COUPE

10.1 INTRODUCTION

This chapter addresses some patterns of grammaticalization in a broad selection of languages of South Asia, a region of considerable cultural and linguistic diversity inhabited by approximately 1.7 billion people living in eight countries (Afghanistan, Bangladesh, Bhutan, India, Nepal, Maldives, Pakistan, and Sri Lanka) and speaking 655 known languages (Simons and Fennig 2017). The primary purpose of the chapter is to present representative examples of grammaticalization in the languages of the region—a task that also offers the opportunity to discuss correlations between the South Asian linguistic area and evidence suggestive of contact-induced grammaticalization. With this secondary objective in mind, the chapter intentionally focuses upon processes that either target semantically equivalent lexical roots and constructions or replicate syntactic structures across genetically unrelated languages.

The theoretical concept of ‘grammaticalization’ adopted here is consistent with descriptions of the phenomenon first proposed by Meillet (1912), and subsequently developed by e.g. Givón (1979a), Lehmann (1985), Traugott and Heine (1991) and papers therein, Bybee, Perkins, and Pagliuca (1994), Heine, Claudi, and Hünemeyer (1991a), and Heine and Kuteva (2002, 2005). In accordance with this preceding work, grammaticalization is viewed as a historical process in which lexical morphemes, or constructions involving lexical morphemes, are gradually bleached of their precise semantic specificity and develop more abstract grammatical meanings that permit the conventionalization of their use in a potentially widening range of functional domains. This process is usually complemented by some phonetic erosion of the grammaticalized morpheme(s), but a reduction in phonological bulk may not necessarily accompany the shift from a concrete lexical meaning towards a more abstract grammatical meaning; both the grammaticalized element and its lexical source(s) may coexist with an identical form for an extended period of time, thereby giving rise to ambiguous interpretations of meaning. For example, the light verbs of

Indo-Aryan languages are phonologically identical in form to their lexical roots, despite their grammaticalization processes being of considerable antiquity.

The chapter proceeds as follows. Section 10.2 introduces the reader to the languages of South Asia, discusses their present geographical distributions, and briefly outlines their typological profiles. Section 10.3 considers the factors that contribute to South Asia being recognized as a linguistic area. Also addressed here are some of the problems a researcher may face in deciding if a particular grammaticalization pattern is an independent phenomenon, or one induced by contact. Section 10.4 investigates the lexical sources of some body-part nouns, the trajectories by which they have developed as grammatical morphemes encoding case relations and other functional meanings, and how some have additionally developed clause-linking functions. Sections 10.5–10.7 examine instances of lexical verbs that have grammaticalized various valency-modifying, aspectual, and modality meanings from compounds, and section 10.8 examines the relative–correlative construction of South Asia and considers whether its wide distribution could be due to contact-induced grammaticalization. The chapter concludes in section 10.9 with a discussion of the findings and the implications for establishing grammaticalization patterns in linguistic areas.

10.2 THE LANGUAGES OF SOUTH ASIA

South Asia is home to representatives of at least six major linguistic stocks: Austroasiatic (Munda languages in eastern peninsular India, and Khasian languages in Northeast India), Dravidian (principally the south of peninsular India and northern Sri Lanka, plus one outlier in Baluchistan), the Indo-Iranian branch of Indo-European (namely the Indo-Aryan, Iranian, and Nuristani languages of the northern half of the subcontinent, plus Sinhala and Dhivehi, spoken in Sri Lanka and the Maldives respectively), the Tibeto-Burman languages of the Sino-Tibetan family (spoken predominantly in the Himalayas and the peripheral hill states of Northeast India), and the Tai branch of Tai-Kadai (eastern Assam and Arunachal Pradesh, Northeast India, not shown in the map in Fig. 10.1).

To this array we might add the Great Andamanese and Ongan (a.k.a. Angean) language families of the Andaman Islands. Although these languages are geographically located somewhat closer to mainland Southeast Asia, they are reported to have a head-final constituent order as well as a retroflex series of consonant phonemes that cannot be attributed to language contact (Abbi 2006). This suspiciously links them to many language families of the subcontinent, and distinguishes them substantially from the typologically very different languages of Southeast Asia. Masica (2016: 245) therefore ponders whether they could be the remnant of an ancient substratum formerly located on the Indian mainland. Lastly, there is a handful of language isolates, such as the Burushaski language of the Hunza and Gilgit districts in northern Pakistan, the Kusunda language of western Nepal, and (if it is still spoken) the Nihali language of central-west India. Collectively these languages offer a veritable smorgasbord of typological features and profiles, but also some interesting commonalities, particularly with respect to shared grammaticalization patterns and structural convergence.

FIG. 10.1. Distribution of South Asian language families. Nihali, Kusunda, and Tai-Kadai languages are not shown (adapted from *A Historical Atlas of South Asia*, Oxford University Press, 1992)

The majority of South Asian languages have a AOV/SV constituent order and demonstrate typological characteristics associated with head-final languages, as outlined by Greenberg (1963), such as genitive–noun and relative–noun order, postpositions, a dominant tendency for suffixal morphology, main verbs preceding auxiliary verbs, and standards of comparison preceding adjectives. The vast majority employ dependent marking at the level of the clause, and most also demonstrate the indexing of one or more arguments on matrix verbs. Word formation is typically agglutinative and synthetic, with the exception of Indo-Aryan languages, which demonstrate a moderate degree of fusion, a feature consistent with their Indo-European roots.

Narrative chaining via converb constructions (a.k.a. conjunctive participles or gerunds) characterizes clause linkage patterns, and was initially proposed as a key typological feature identifying South Asia as a linguistic area (e.g. Emeneau 1956), before subsequent work by Masica (1976) established the fact that similar converbal clause chaining patterns extend well into the 'Indo-Altaic' area of Central and Far East Asia, the Horn of Africa, and even into parts of Europe.

The only languages observed to deviate substantially from the general South Asian typological profile are the head-initial Khasian languages of Meghalaya state and adjacent regions of Bangladesh. These languages also demonstrate typological features that accord with Greenbergian universals, and thus have opposite orders to those outlined above for the head-final languages of South Asia. The relatively recently arrived Tai languages of eastern Assam and Arunachal Pradesh similarly conform to a head-initial typology, have isolating word-formation typology, and are tonal, in common with their Tai relatives in Southeast Asia. According to research by Morey (2006), Greenberg's characterization of Khamti (a Tai language of Assam) as being an exceptional AOV/SV language with prepositions was inaccurate. He concludes that the basic constituent order is AVO/SV, but that verb-final structures are possible under certain pragmatically defined circumstances, and he proposes that language contact with Assamese and other verb-final Tibeto-Burman languages of the region may have played a role in variant orders. Some Tai languages have also developed postpositional (anti-agentive) marking on O arguments (Morey 2006: 361–3), possibly due to the areal influence of neighbouring Indic and Tibeto-Burman languages.

Sources of data on Indo-Aryan languages are substantial, extending back in written form to the middle of the second millennium BCE, and Dravidian written sources in Tamil, Telugu, Kannada and Malayalam are extant from the early centuries of the Christian era (Southworth 2005: 39, 48). These textual sources are extremely valuable for investigating the diachrony of grammaticalization phenomena. With the exception of Tibetan (with the earliest written records dating from the eighth century CE) and the Ahom *buranjis*, which chronicled life and governance in the Ahom kingdom (1228–1826 CE) and were initially written in the Ahom script, the other languages of South Asia are unwritten.¹ As Heine (Chapter 2 this volume) notes, an investigation of grammaticalization phenomena in unwritten languages must therefore rely upon internal reconstruction, historical reconstructions, and typological considerations. Despite the limitations, it is still possible to reveal a good deal of evidence for grammaticalization using a combination of these techniques.

10.3 SOUTH ASIA AS A LINGUISTIC AREA

Heine and Kuteva (2005: 172) recognize three types of linguistic area: (1) those established by the presence of a shared set of linguistic features; (2) those in which the

¹ Minor exceptions are the Tibeto-Burman languages Newar, Lepcha, Limbu, and Meiteelon, for which writing systems were developed at different times within the past millennium.

languages share a high degree of mutual intertranslatability; and (3) those that share the same processes of grammaticalization (and thus form a grammaticalization area). As they note, these types are not mutually exclusive, and all three criteria arguably apply to the Sprachbund of South Asia.

Since Emeneau (1956) and later work by Masica (1976), the subcontinent of South Asia has been recognized as a linguistic area in which particular linguistic features have diffused across the genetic boundaries of unrelated languages as a consequence of longstanding stable multilingualism. Emeneau (1956: 16) defines a linguistic area as ‘an area which includes languages belonging to more than one family but showing traits in common which are found not to belong to the other members of (at least) one of the families’. One prominent trait is the presence of a phonemic contrast between dental and retroflex consonants in South Asian languages. Retroflex plosives can be reconstructed to Proto-Dravidian and have spread into Indo-Aryan (with the exception of Assamese),² but are extremely rare or nonexistent in other branches of Indo-European.³ Written records from Middle and New Indo-Aryan demonstrate an increasing occurrence of retroflex consonants over time, which Emeneau (1956: 7) holds to be a clear demonstration of the ‘Indianization’ of the Indo-Aryan branch of Indo-European. Retroflex consonants are also found in Munda, Burushaski, and Tibeto-Burman languages of South Asia in contact with Indic languages (but more rarely in related languages outside of the South Asian Sprachbund),⁴ so this distribution is plausibly attributed to convergence that began with the diffusion of Dravidian loanwords containing retroflexes into Vedic Sanskrit (Kuiper 1967: 89).

The languages of Kupwar village are a celebrated example of how six centuries of stable multilingualism have led to what is ostensibly a single grammatical template being employed for three distinct local varieties of Urdu, Marathi, and Kannada, as demonstrated by the data of (1). Gumperz and Wilson (1971: 155) observe that this has resulted in ‘a gradual adoption of grammatical differences to the point that only morphophonemic differences (differences of lexical shape) remain’.

² The retroflex and dental plosive series common to all other New Indo-Aryan languages appear to have settled on an articulatory compromise in Assamese, resulting in a single alveolar series of plosives. Because a dental-retroflex contrast is also represented in the Assamese orthography and in the phonological inventories of related languages, it is assumed that a phonological contrast must have been historically present at an earlier stage of the language (Mahanta 2012: 219). Bilingualism in Assamese varieties used as lingua francas in Northeast India may have contributed to a simplification of the Assamese phonological inventory. For example, it has been noted that when people converse in Nagamese, the Assamese-based lingua franca of Nagaland, they simply use their L1 phonology (Sridhar 1974; Burling 2007).

³ Retroflex consonants reported in North Germanic languages (e.g. Hamann 2003: 83–4) appear to be attributable to relatively recent mergers involving rhotics and stops, and thus have no bearing of the value of retroflexion as a defining feature of South Asia as a linguistic area.

⁴ Arsenault (2017: 12) finds that the retroflex consonants of Sino-Tibetan languages spoken in China are typologically divergent from those of South Asia, and Matisoff (2003: 21) proposes that retroflexes in Tibeto-Burman languages are secondarily derived from proto-clusters with medial liquids. This explanation accounts for the presence of a retroflex plosive phoneme with a rhotacized release in Sangtam, a Tibeto-Burman language of central Nagaland (Coupe, in prep.). Retroflex consonants appear to be exclusive to Austroasiatic languages spoken in South Asia (e.g. Jenny, Weber, and Weymuth 2015).

- (1) Kupwar Village (Maharashtra): (a) Urdu, (b) Marathi, and (c) Kannada
a.	pala	jəra	kaṭ-ke	le-ke	a	-Ø	-ya
b.	pala	jəra	kap-un	ghe-un	a	-l	-o
c.	tapla	jəra	khod-i	taḡond-i	bə	-Ø	-yn
- greens a.little cut-CNV take-CNV come (-1SG) -PFV
 'I cut some greens and brought them.' (Gumperz and Wilson 1971: 154)⁵

Anderson (2008: 4–5) remarks on the anomaly that, despite an extended period of coexistence of speakers of Munda and Indo-Aryan languages, Sanskrit and Middle Indic texts demonstrate no evidence of borrowing from Munda, even for plant or animal names. Similarly, Thomason and Kaufman (1988: 115) propose that Dravidian structural interference in Indic involved minimal lexical transfer. Since it is generally assumed that lexical borrowing normally precedes structural borrowing (e.g. as implied by Comrie's (1989: 210) suggested implicational universal for borrowability), this paradox can perhaps be explained with recourse to the following sociolinguistic considerations.

It is likely that autochthonous Munda speakers typically occupied a marginalized and lowly socioeconomic position in the Hinduized society of Vedic South Asia, just as the tribal (*ādivāsī*) people generally still do in modern India. In support of this assumption, Southworth (2005: 42), citing Thapar (1978), mentions the contempt expressed in Rigvedic hymns for non-Hindu indigenous people, their religious beliefs, and their languages. With the prevalence of such pejorative attitudes directed towards the indigenous cultures and languages of South Asia, it would be highly improbable for Munda lexical items to be borrowed by a superstrate language. Asymmetrical sociopolitical relationships between Indo-Aryan invaders and the conquered also possibly accounts for the paucity of old Dravidian loan words in Indic (Thomason and Kaufman 1988: 116). As lexicon is highly emblematic of caste or clan membership throughout South Asia, its important role in the identification of one's social affiliation potentially makes it quite resistant to borrowing anyway.⁶ Speakers instead tend to reduce the cognitive burden of needing to speak multiple languages in their daily lives by minimizing structural differences via convergence while maintaining more overt inter-group lexical differences.

⁵ The original glossing and interlinearization has been modified to more accurately represent the grammatical categories of these examples.

⁶ The Brahui language of Baluchistan clearly presents a counterexample to this statement. This Dravidian outlier was initially thought to be an Indo-Aryan language because of the preponderance of lexical items of non-Dravidian origin (Southworth 2005: 12). However, just as a speech community might eschew lexical borrowing to outwardly maintain a caste distinction, so too might it actively borrow vocabulary to hide one, particularly if there is social pressure to assimilate to the dominant culture or language of a region. Assimilation in Baluchistan was facilitated by intermarriage between tribes resulting in a high degree of bilingualism (Emeneau 1980: 318), and this must have facilitated lexical borrowing, much of which was probably unidirectional. See Bashir (2016: 273–5) for further discussion of Brahui–Baluchi convergence.

While it is clearly demonstrated that languages in sustained contact can converge in structure, as amply suggested by the data of (1), it has also been proposed that grammaticalization patterns may be sensitive to language contact (e.g. Heine and Kuteva 2005). However, proving that a particular grammaticalization pattern results from language contact is an endeavour potentially beset by a number of uncertainties that may complicate the picture. These are outlined below.

First, assuming that plausible criteria can be established for identifying a linguistic area, how does one decide if a particular grammaticalization shared by two or more languages in a multilingual contact zone is unambiguously a consequence of language contact? Obviously if a certain grammaticalization process is found only in languages located within the linguistic area, and the pattern is also unattested in related languages outside of that convergence zone, then this might be taken as convincing evidence that a borrowed conceptual schema has resulted in a shared grammaticalization.

The possibility that such a shared pattern could be contact-induced is further suggested by the observed grammaticalization and its outcome being cross-linguistically rare. To illustrate, one of the processes to be discussed in this chapter concerns the grammaticalization of a conative modality meaning from a compound involving a verb of perception (typically 'look' or 'see') in a number of unrelated South Asian languages—see section 10.7). Now, this could very well constitute a case of contact-induced grammaticalization across genetically unrelated languages within the contact zone, as it is a pattern not widely attested in the languages of the world. Conative meanings are reportedly more likely to be associated with imperfective aspect, or with different semantic classes of verbs and verbal constructions (e.g. 'try', 'obtain', 'taste', 'go' + INFINITIVE, VERB + light verb 'do'), and in some languages conativity may instead be expressed via partitive or dative case marking on an O argument (see Vincent 2013). It is also noteworthy that the development of conative markers from verbs of perception appears to cluster in regions with high linguistic diversity, further suggesting that a conceptual schema can diffuse through contact. For example, Foley (1986: 152) reports that this is an almost universal grammaticalization pattern in the Papuan languages of New Guinea:

- (2) Asmat (Asmat Family), Papua Province, eastern Indonesia (Drabbe 1959)
 yitim-**por**
 arise-**see**
 'try to awaken somebody'
- (3) Barai (Koiarian Family), Papua New Guinea (Olson 1975: 496)
 akoe **ga**
 throw **see**
 'try throwing it'⁷

⁷ Foley (1986) glosses Barai *ga* as 'see'. The source glosses *ga* as 'look' and translates the example as 'Try throwing it and see'. The meaning of *ko* as a main verb in Hua is translated by Haiman (1980: 147) as 'see, look'.

- (4) Hua (Yagaria, East Central Highlands Family), Papua New Guinea (Haiman 1980: 147)
 ke hu ko-mana
 talk do see:1-INCON
 'I tried to talk (but to no avail)'
- (5) Yimas (Lower Sepik Family) Papua New Guinea (Foley 1986: 152)
 na-mpɪ-kwalca-tay-ntut
 3SG.U-3DU.ACT-arise-see-REM.P
 'they both tried to wake him up'

The targeting of a verb of visual perception for the grammaticalization of conative modality in both New Guinea and South Asia may be due to a universal cognitive representation of human experience, but could additionally be the consequence of contact-induced grammaticalization resulting from such a diffusing cognitive schema. These two causal factors are not necessarily incompatible in a linguistic area (e.g. Heine and Kuteva 2005: 14). Certainly the concentration of this grammaticalization pattern in known contact zones with a high incidence of linguistic diversity, its relative rarity, and the lack of a random distribution of conativity based on a verb of perception across the languages of the world all lead to the conclusion that such a clustering is highly unlikely to be due to chance.⁸

Conversely, if a particular grammaticalization pattern is also found in languages more widely as well as in the contact zone, then that phenomenon might be justifiably attributed to parallel developments known as 'drift' (e.g. Sapir 1921; Robbeets and Cuyckens 2013). A globally attested example of this is the grammaticalization of postpositions from nouns in OV languages (e.g. Aristar 1991). Another is the development of indefinite articles from the numeral 'one' (Robbeets and Cuyckens 2013: 3). Even if most of the languages of South Asia also happen to demonstrate a propensity to grammaticalize postpositions from relational nouns and indefinite articles from the numeral 'one', the ubiquity of these patterns in hundreds of languages around the globe underscores the supposition that both processes represent universal grammaticalization pathways that could just as possibly arise independently in languages that happen to share a linguistic area.

With these caveats now stated, the chapter will proceed to describe some grammaticalization processes observed to occur in a selection of South Asian languages, and where the evidence is sufficiently convincing, some of these developments will be identified as plausibly resulting from language contact.

10.4 GRAMMATICALIZATION OF RELATIONAL MORPHOLOGY AND METAPHORICAL EXTENSIONS

Relational nouns denoting body-part terms and spatial locations are an especially rich source of case marking and converbal morphology in South Asian languages.

⁸ I am grateful to Heiko Narrog for comments that helped to clarify this discussion.

These typically develop out of compounded genitival [N_1 -GEN N_2] or appositional [N_1 N_2] constructions, in which the head noun N_2 is originally a body-part term or a noun denoting a spatial location. The compound's head loses its lexical status as it undergoes semantic bleaching, eventually permitting it to function as a grammatical morpheme encoding a purely relational meaning in this type of construction. Compounds involving a genitival suffix may initially retain the genitive in the ensuing grammaticalization of the postposition, thus revealing the diachronic origin of the grammaticalized collocation and giving rise to constructions such as the Hindi comitative *N-ke sāth* (from the noun *sāth* 'company, society'), and the postessive *N-ke pīche* (from the noun *pīchā* 'rear part, hindquarter'), e.g. *parivār-ke sāth* 'with (the) family', *ghar-ke pīche* 'behind (the) house'. Structurally similar examples of case compounding involving genitival morphemes in Bodic languages of the Himalayan region are discussed in Noonan (2008).

The following subsections describe some grammaticalization processes involving body-part nouns in Indo-Aryan and Tibeto-Burman languages. It will be demonstrated how, once a noun has grammaticalized as an oblique relational form, it can then be extended to even more abstract morphosyntactic functions in the grammar, such as non-finite and finite clause linkage.

10.4.1 'SIDE, FLANK' > POSTPOSITION

An oblique case-marking postposition has grammaticalized from a construction involving a relational noun with the meaning of 'armpit, side, flank' in a number of South Asian languages. According to Beames (2012[1875]: 257) and Chatterji (1970[1926]: 173), the lexical source of the dative marker found in New Indo-Aryan languages—e.g. Hindi *ko*, Bangla *ke* and Oriya *ku*—is the Middle Indo-Aryan locative declension of the Sanskrit noun *kakṣe* (armpit.LOC.SG.M) 'in the armpit'. Elaborating on the observations of these and other Indic scholars, Reinöhl (2016: 58–9) proposes that the starting point of the grammaticalization would most likely have been 'side (of the body), flank', as a metaphorical extension of the meaning of *kakṣa*.

The first attested uses of *ko* as a dative/accusative marker are from Old Urdu/Punjabi texts dating from the twelfth and thirteenth centuries (Butt and Ahmed 2011: 565, cited in Reinöhl 2016: 58–9). By 1800, *ko* was used in Hindi to mark both abstract and concrete goals.

(6) Hindi (New Indo-Aryan, circa 1800) (Butt and Ahmed 2011: 565–6)⁹

- a. is manzil **ko** kab poāco-ge
 this destination DAT/ACC when reach:FUT:2PL
 'When will (you) reach this destination?'
- b. apne haq **ko** poāc kar
 self right DAT/ACC reach having
 'having attained one's right'

The grammaticalization trajectory from body-part noun expressing a spatial location to a postposition encoding goals and recipients, and finally to marking a core

⁹ Glosses have been adjusted in these and the following examples for consistency.

argument in O function, follows the expected pathway to an increasingly more abstract function that characterizes the evolution of grammaticalized morphemes. The metaphorical shift to marking certain types of core arguments in O function in Indo-Aryan languages is of particular interest, because it is only obligatory when the referent of the O argument is human or highly referential.

LaPolla (1992, 1995b: 1175–6) suggests that the primary motivation for this type of case marking initially is the disambiguation of semantic roles. His observations are made in respect to pragmatically motivated relational marking in Tibeto-Burman, but they are just as relevant to Indo-Aryan: when there is no potential confusion as to which argument is the agent, there is consequently no requirement for overt marking, whereas the presence of two possible (especially human) agents in the sentence requires disambiguation via some morphosyntactic means. This is achieved via dative marking on the O argument in a grammaticalized extension of the older locative marking function.

Rather than extending the discriminatory marking to the patient, another option is for languages to instead mark the agent. This is a quite common pattern in Tibeto-Burman languages (see e.g. LaPolla 1995a; Noonan 2008). All but one member of the Ao group languages of central Nagaland have a syncretic postpositional clitic *nə~na* that is used to mark the agentive and instrumental cases. The Ao dialects of this sub-grouping are unique in additionally marking the allative case with the same form, and this constitutes a previously unattested agentive/instrumental/allative syncretism in the languages of the world (Coupe 2011). The same morpheme is also recognizable in the ablative forms of the Ao group languages, all of which have originated from old appositional *N₁ N₂* compounds that have been subjected to cycles of grammaticalization as relational morphemes over time.¹⁰ The examples of (7) respectively demonstrate the agentive, instrumental, allative and ablative functions of this isomorphic form in Mongsen Ao.

(7) Mongsen Ao (Tibeto-Burman), Nagaland (Coupe 2007a: 173–8)

- a. *mətʃatshəŋ nə pùŋì tʃu tsəŋ*
 mətʃatshəŋ **nə** pùŋì tʃu tsəŋ
 PN AGT wild.pig DIST spear.PST
 ‘Mechatseng speared the wild pig.’
- b. *təɪ məzəʔ nə ɬuŋukù tʃì*
 təɪ-əɪ məzəʔ **nə** ɬuŋ-ukù tʃì
 thus-SEQ fire INS burn-ANT PTCL
 ‘And, [he] cleared [the field] with fire.’
- c. ... *təpaʔ taŋ nə wai*,
 tə-paʔ taŋ **nə** wai-əɪ
 RL-father SIDE ALL go-SEQ
 ‘... after going to the father ...’

¹⁰ It is common in Tibeto-Burman languages for an ablative marker to have the form of a dimorphic agentive/instrumental-locative compound. For examples in the Ao group and beyond, see Coupe (2011) and references therein.

- d. *nuksənsaŋpaʔ áhlú phinə tʃhuwaɪ*
 nuksənsaŋ-pàʔ a-hlú **phinə** tʃhuwa-àɪ
 PN-M NRL-field **ABL** emerge-PRS
 'Noksensangba returns from the field.'

In common with the New Indo-Aryan languages discussed above, the most credible diachronic source for the agentive/instrumental/allative case marker is a lexical noun reconstructed to Proto-Tibeto-Burman as **ʔ-nam* 'side/rib' (Matisoff 2003: 100, 604). The root of a cognate Chungli Ao form *tena* is defined by Clark (1911: 626) as 'side at the waist where are no ribs', a meaning that seems wholly consistent with 'flank'. This would have initially grammaticalized as a locative marker **na* in Proto-Ao, probably in much the same way as the dative/accusative marker of New Indo-Aryan initially evolved from a body-part noun with a locative meaning.¹¹ The metaphorical extension to marking a core argument in A function is similarly motivated by the need to disambiguate semantic roles when this is pragmatically motivated, analogously to the way that dative marking becomes obligatory in Indic languages when an O argument is human or highly referential and thus could be mistaken for the A argument. Agentive marking in Tibeto-Burman languages is especially likely to appear when a non-canonical ordering of arguments places the A argument in non-initial position in the sentence.

What makes the grammaticalization of this syncretic form intriguing is its metaphorical extension from what must have originally been a locative marking function to marking an agent, as locations and agents seemingly have very little in common semantically. However, it is probable that at the earliest stages of grammaticalization, Proto-Ao **na* was a semantically underspecified oblique postposition that could be used for marking instruments and sources in addition to goals, and it was the instrumental meaning that was targeted and extended to marking a core argument. The semantic link shared by agents and instruments is that both are in some sense effectors of actions, so once the instrumental meaning of *nə* evolved to marking instruments, it would have been only a small metaphorical step to extend the marking to agents. In this respect Mongsen Ao is in harmonious accord with Narrog's (2014) revision of Heine et al. (1991a), which proposes that in instances of case polysemy, instrument marking precedes the progression to agent marking in the diachrony of grammaticalization chains. The path from an oblique case to a core syntactic case is also consistent with the cross-linguistically valid observation that grammaticalization chains evolve increasingly more abstract categories of grammar (Givón 1979a; Heine and Kuteva 2002, 2007; Bybee et al. 1994).

¹¹ Some Kuki-Chin languages of the northeastern region have an agentive/instrumental form *in* or *na*; Meiteelon (a.k.a. Meithei, Manipuri) and Tangkhul similarly have a form *nə* (LaPolla 1995a). The agentive/instrumental markers of Meiteelon, Tangkhul, and a number of Kuki-Chin languages are all suspiciously similar to the reconstructed Proto-Ao form **na*. Reflexes could have been genetically inherited from a common intermediate proto-language, although borrowing under an intense contact situation cannot be ruled out—consider the case of Chungli Ao, which has borrowed an agentive/instrumental marker from its Konyak neighbour Chang (Coupe 2007b: 356–7).

It is noteworthy that a lexical noun *tāŋi*, also meaning ‘side’, is currently undergoing a new cycle of grammaticalization in Mongsen Ao. An example is given in (7c) above. Its phonetically reduced grammaticalized form *tāŋ* (SIDE) is obligatorily used to mark NPs representing human goals of movement and speech. It appears to have grammaticalized relatively recently as a new postposition, as it is sometimes determined by a demonstrative, just as a lexical noun would be, and its noun phrase is also additionally case-marked by the older allative form *nā*.

- (8) Mongsen Ao (Tibeto-Burman), Nagaland (Coupe 2017a: 290)

tāŋjāi tŋū nā tǎpā? khā tǎjā nāt tāŋ tŋū nā wāi, ...

tāŋjāi tŋū nā tǎ-pā? khā tǎ-jā nāt tāŋ tŋū nā wā-āi
other DIST AGT RL-father CONJ RL-mother two SIDE DIST ALL go-SEQ
‘Others went to the mother and father, ...’ (<lit. ‘the mother and father’s side’)

This development is indicative of the cyclical nature of the grammaticalization of relational morphology, and provides a cogent explanation for how case compounding develops historically. New postpositions are continually emerging from body-part or spatial nouns and are put to new uses, but older forms of relational morphology are retained alongside the newly evolving forms. The position of a relational morpheme in the syntagmatic collocation and its proximity to the lexical head of the NP thus reflects the historical chronology of its emergence as a grammaticalized morpheme.

10.4.2 GRAMMATICALIZATION OF CONVERB MORPHOLOGY FROM POSTPOSITIONS

An additional grammaticalized function of postpositions in many languages of South Asia is clause linkage. Case-marking and nominalizing morphology is often recognizable in the forms of non-finite converb suffixes in Tibeto-Burman and Indo-Aryan languages of South Asia, and a recent investigation of their diachronic origins concludes that post-head and headless relative clauses plausibly provide a historical pathway for their reanalysis as clause-linking devices (Coupe 2017b). The reanalysis of a relativized noun phrase as a type of non-finite clause linker is demonstrated by the following Mongsen Ao examples. These capture collocations involving the agentive/instrumental case marker and the general nominalizer *-pā?* in respectively less and more advanced stages of evolving a predicative causal meaning out of an older referential meaning.

- (9) Mongsen Ao (Tibeto-Burman), Nagaland (Coupe 2017b: 226–7)

a. *tŋhā-pā? tŋū nā pā lūŋkī nā wā-Ø*
be.hurt-NMLZ DIST INS 3SG cave ALL go-PST

‘Because of the pain [caused by Mother Hen pouring boiling water on Rat-Pup], she (i.e. Rat-Pup) went into a cave’ (from a folktale explaining why rats kill poultry chicks).

- b. *tə-āɪ* *pā* *nā* ‘*ùʔ*’ *tə* ***āɪəm-pànā***
 thus-SEQ 3SG AGT ONOM thus **bear-CAU.CNV**
pā *tā-pūk* *tā-khūŋlān* *nā* *kāā-jūk-Ø-ùʔ*
 3SG RL-stomach RL-neck ALL come+ascend-PFV-PST-DEC
 And because ‘uh!’ she bore it (i.e. held her laughter in), her stomach rose up
 to her throat.’

The bolded non-finite constructions of (9a,b) both express causal presuppositions. First, the structure of (9a) is identical to a nominalization functioning as a headless relative clause that is determined by the distal demonstrative *tŭ* and case-marked by the instrumental marker *nā*. This is the order in which these constituents appear in a noun phrase modified by a post-head or headless relative clause, yet here it encodes a reason for the event expressed in the matrix clause. Secondly, the form *-pànā* in (9b) encodes an identical causal meaning, but demonstrates a reduction in phonological bulk that is the hallmark of grammaticalization (its source is the nominalizer *-pāʔ* and the instrumental *nā*, which have coalesced as a monomorphemic converb suffix in the reanalysis of function). That the two constructions are alternately used by the same speakers in narrative texts to express causal presuppositions is not surprising, as older structures are known to overlap synchronically with newer grammaticalizing structures in a process known as layering (Hopper 1991; Bybee et al. 1994).

Similar collocations expressing non-finite causal meanings that have grammaticalized from juxtapositions involving nominalized/participle forms and instrumental or ablative postpositions are found throughout Tibeto-Burman, Indo-Aryan, and Dravidian languages of South Asia (see Coupe 2017b: 225–9 for additional examples and discussion).

- (10) Old Tamil (South Dravidian) (Lehmann 1998: 95)
kolli *āram* ***ā-tal-in*** *am* *pukai* *tavaṇ-um...*
 firewood sandalwood **become-NMLZ-ABL** beauty smoke spread-NPS+3SG/PL
 ‘Because the firewood was sandalwood, a beautiful smoke spread.’
- (11) Marathi (Indo-Aryan), central western India (Pandharipande 1997: 114)
gāḍi *yeṇār* ***aslyāne / aṣlyāmuḷe*** *rastā* *banda*
 train come:FUT **be:NMLZ:INS / be:NMLZ:POSP** road close
dzhālā *āhe*
 happen:PST:3SG:M is
 ‘The road is closed because the train is going to come.’

If we may digress briefly, some New Indo-Aryan languages instead use a marker based on a converbal (conjunctive participial) form of the verb ‘say’ to encode a causal presupposition. This represents a grammaticalized usage that has developed via a different pathway, most likely out of a nominalized construction originally used for the expression of quoted speech.

- (12) Assamese (Indo-Aryan) Assam (author's fieldnotes)¹²
 ... manu zɔn upɔɪ-ɔt as-e bul-i
 person CLF:HUM above-LOC exist-3 SAY-CNV
 '(He is afraid) because the man is up there.'
- (13) Nepali (Indo-Aryan) Nepal (Clark 1963: 174)¹³
 hera uslāi sūdra bhan-era nikāl-na ðhik hoina
 EXCLM 3SG:DAT caste.name SAY-CNV take.out-INF correct be:3SG:NEG
 'Look here, it is not right to drive him away because he is a Sudra.'

The grammaticalized use of 'say' for encoding causal, conditional, evidential, and purposive meanings is widespread in Tibeto-Burman (Saxena 1988) as well as in many other language families of the world (see Heine and Kuteva 2002: 261–9 for additional examples). According to Masica (1991: 403), it has often been suggested that the source of the SAY construction in South Asia is Dravidian, although he notes that the pattern is also found in some Tibeto-Burman languages. If these do indeed represent Dravidian calques, then they have probably diffused into the grammars of Tibeto-Burman languages via contact with Indo-Aryan.

10.4.3 GRAMMATICALIZATION OF DISCOURSE CONNECTIVES

Yet another reanalysed function of the Mongsen Ao causal converb permits it to be used as a discourse connective for linking finite clauses. The causal discourse connective consists of a collocation involving the quotative particle *tə* and the conditional converb suffix *-pənə*. The grammaticalized construction never occurs with a demonstrative interpolated between the nominalizer and the postposition (cf. 9a), suggesting that this extension of function is at a more developed stage of evolution. Its purpose in narrative texts is to summarize preceding information and present it as the cause for the following statement that is asserted as the effect. We see this usage in the textual example of (14), which recapitulates the previously stated reasons for why a particular clan has very few members.

- (14) Mongsen Ao (Tibeto-Burman), Nagaland (Coupe 2017a: 228)
 tə-pənə thūkū táŋ-ù?
 thus-CAU.CNV nine only-DEC
 'That's why/because of that there are only nine.'

Other converb suffixes whose forms have grammaticalized from nominalized noun phrases are used in the same manner as discourse connectives. For example, the conditional converb suffix *-pālā*, formed by a fusion of the general nominalizer *-pəʔ* and the topic particle *lā*, similarly combines with the quotative particle *tə* to form a

¹² This example is taken from a recorded narration of the Pear Film (Chafe 1980) and immediately follows the sentence of (29) below.

¹³ Glosses have been added.

conditional discourse connective *tə-pālā*. Its function is to summarize preceding information terminating in a finite clause (representing the protasis) and to relate it to the following discourse (representing the apodosis). If it were used in place of the causal discourse connective of (14), the resulting meaning would be ‘If that be the case, there are only nine’. The functions of both of these discourse connectives are thus logical extensions of their grammaticalized uses as non-finite converb suffixes, and they express meanings consistent with their immediate sources in their respective grammaticalization chains. This correlates with Hopper’s (1991: 22) principle of ‘persistence’.

10.4.4 ‘HAND/ARM’ > TERMINATIVE POSTPOSITIONAL CLITIC, TERMINATIVE ASPECT

In Mongsen Ao, a possibly unique but entirely plausible grammaticalization of Proto-Tibeto-Burman **lak* ‘arm, hand’ has undergone metaphorical extension to a relational meaning initially expressing ‘terminal part in space’ in N_1 – N_2 compounds, and then extending to a meaning of ‘terminal part in time’ in verb stems. There is little doubt that these two grammaticalized meanings stem from the same source, as both express an identical meaning, the only difference being that one is situated in the dimension of space, the other in the dimension of time.

(15) Mongsen Ao (Tibeto-Burman), Nagaland (Coupe 2017b: 285)

a. *tə-mijün-lāk* RL-finger-TERM ‘fingertip’

b. *səntün-lāk* tree-TERM ‘apex of tree’

c. *tə-mī-lāk* RL-tail-TERM ‘tail tip’

d. *mī-lak* fire-TERM ‘flame’

e. *tə ākī t̪fū t̪hà... t̪hālākəɪ*,

tə ā-kī t̪fū t̪hà... t̪hà-lāk-əɪ

thus NRL-house DIST make make-TERM-SEQ

‘Thus, having finished (narrator stutters) building [his] house...’

10.5 ‘SEND’, ‘GIVE’ > MORPHOLOGICAL CAUSATIVE

The inherent lexical semantics of ‘give’ and ‘send’ makes them common targets for grammaticalization as causative morphemes. Masica (1976, 2016: 245) singles out morphologically marked causative verbs as one of the defining features of South Asia as a linguistic area, and therefore his observation invites a closer inspection of their lexical sources in unrelated South Asian languages to establish what they may have in common.

In some South Asian languages (especially Indo-Aryan, Munda, and Dravidian), causative morphemes are thought to develop out of ‘explicator’ compound constructions, in which a non-finite converbal or absolutive verb form is compounded with a clause-final main verb (e.g. Masica 1976; Hook 1974). But this is not necessarily the

case for all languages of South Asia, as the following data will demonstrate, and collocations of verb roots present another credible pathway for their grammaticalization as valency-modifying morphemes.

Lexical verbs expressing similar semantics are also frequently grammaticalized as causative morphemes in languages spoken in regions extending beyond South Asia, especially in Southeast Asia (e.g. Matisoff 1991). Jenny (2015: 156) proposes that the development of equivalent grammaticalized meanings from the same lexical source over such an expansive geographical region is suggestive of either language contact situations in the past, or possibly a chain of contact situations, or language-internal developments involving a shared cognitive schema. Each of these scenarios is equally plausible, but establishing precisely which is the most likely reason for a particular grammaticalization pattern occurring in a convergence zone is a potentially challenging task, as noted earlier. It is also possible that all these factors may contribute in some way to facilitating a particular lexeme's grammaticalization as a functional morpheme.

A verb with the meaning of 'send (on an errand, entrust with a commission)', 'make', or 'give' is extensively found to grammaticalize as a causative morpheme in Tibeto-Burman languages. LaPolla (1994: 70–1) views this development as an instance of 'drift' in a Sapirian sense (Sapir 1921: 150), as many of the lexical forms that grammaticalize as morphological causatives are demonstrated to be non-cognate in Tibeto-Burman languages. This is strongly suggestive of a causative cognitive schema that is associated with the semantics of such verbs, and which facilitates their metaphorical extension as grammaticalized morphemes.

In the examples of (16) presented below, the grammaticalized use of Mongsen Ao *zək* 'send' as a suffix *-zək* encoding causative-related meanings in (16a,b) is contrasted with the main verb usage of a cognate form in (16c). The causative morpheme is segmentally identical but carries a different tone, which is a common corollary to the grammaticalization of many functional morphemes in this language.

(16) Mongsen Ao (Tibeto-Burman), Nagaland (Coupe 2007a: 324, 309, 325)

a. *təɪ, tshəluŋla nə asijukzəkpa? sə ní? la phàtajù?*

tə-əɪ tshəluŋ-la nə asì?-juk-**zək**-pa? sə

thus-SEQ fox-F AGT deceive-PFV-SEND-NMLZ ANA.DEM

ní? la phàJa?-i-ù?

one.day TOP catch-IRR-DEC

'And, the fox that deceived [him] one day will be caught.' (lit. '... caused him to be deceived')

b. *tətʃhàku mitəm nə pi tshəmàzək*

tə-tʃhà-ku mitəm nə pi tshə-mà-**zək**-Ø

thus-do-LOC.CV pestle INST PROX pound-SPLIT-SEND-PST

'And then, this [cane] was split by the pestle.'

c. *kiphuɪ nə áwkla khə ajila nət áhlu nə zək*

kiphuɪ nə a-úk-la khə a-ji-la nət a-hlú nə **zək**-Ø.

owner AGT NRL-pig-F CONJ NRL-dog-F two NRL-field ALL send-PST

'[The] owner sent his pig and his dog to his field.'

In Kham, an obviously non-cognate grammaticalized form of the verb *pərĩ:-* ‘to send’ is used in a periphrastic causative construction when the causation is indirect, while an identical form continues to be used synchronically as a main verb. The periphrastic causative construction with *pərĩ:-* must be used when the causee retains volitional control over the caused event (e.g. see 17a,b). This usage contrasts with a morphological causative *-se*, which is obligatorily used when there is direct causation (as in 17c).

(17) Kham (Tibeto-Burman), W. Nepal (Watters 2004: 227–9)

- a. no-lai dō:h-wo ŋa-**pərĩ:-**ke
him-OBJ run-NMLZ 1SG-SEND-PFV
‘I made him run.’
- b. o-za:-lai syā:-wo **pərĩ:-**ke-o
3S-child-OBJ sleep-NMLZ SEND-PFV-3SG
‘She made her child go off to sleep.’
- c. baza-rə ya-sə-buhr-ke-o
bird-PL 3PL-CAU-fly-PFV-3SG
‘He flushed the birds. (lit. ‘made them fly’)¹⁴

An identical lexical source for a morphological causative is also reported in the Austroasiatic language Khasi, spoken in the West Khasi Hills of Meghalaya state, Northeast India. According to Temsen and Koshy (2011: 244–5), the causative function of *p^ha?* (one of two morphological causatives in the language) has grammaticalized from a lexical verb *p^ha?*, also expressing a core meaning of ‘send’. Furthermore, as in Kham, grammaticalized *p^ha?* is restricted to use in causativized clauses in which the causee retains control over the caused event, i.e. when there is indirect causation. The grammaticalized causative morpheme is contrasted with the main verb usage in (18a,b).

(18) Khasi (Austroasiatic), Meghalaya

- a. u-jən u-**p^ha?**-thya? ya-i-k^hillun
M-John 3M:SG-SEND-sleep ACC-n-child
‘John made the child sleep.’ (Temsen and Koshy 2011: 245)
- b. ša ka iyeng ki-n sa **p^ha?**
ALL F:SG house 3PL-IRR IRR send
‘To the house they will send.’ (Nagaraja 1985: 88)¹⁵

The form *p^ha?* ‘send’ also occurs frequently as a lexical verb in the closely related Khasian language Pnar, but a search of an extensive corpus of narrative texts reveals that it has not grammaticalized as a morphological causative synchronically (Hiram Ring, pers. comm.) Nevertheless, its participation as V₁ in V₁V₂ quasi-compounds such as *p^ha?* *sumar* in (19b) suspiciously makes these constructions syntactically identical to verb complexes involving the Khasi SEND causative, as demonstrated by (18a), and this structure is probably the prelude to the grammaticalization of a causative meaning.

¹⁴ Glosses have been added to the example.

¹⁵ Glosses have been added to this example.

(19) Pnar (Austroasiatic), Meghalaya (Ring 2015a: 270)

- a. man m̐ je ki jaitʃaʔ u-ð̌iʔ kat-kam
 become NEG able 3PL be.patient NFIN-drink as-like
 ka=k̐os wa da p^{ha}ʔ i
 F=course NMLZ REAL **send** 1PL

‘so they have no patience to drink (take medicine) according to the course that we (doctors) sent’

- b. t̐ ka wa sumar ŋa, ka wa paʔ̌
 NVIS 3SG:F:ACC NMLZ take.care 1SG:TOP 3SG:F:ACC NMLZ look
 ja ŋa, ka wa e ďʒa ja ŋa,
 BEN 1SG:TOP 3SG:F:ACC NMLZ give rice BEN 1SG:TOP
 p^{ha}ʔ sumar ki=sistar
send take.care PL=sister(nun)

‘so (it was) she that cared for me, she that looked after me, she that gave me food and (she) was sent to care (for me) (by) the sisters’ (Ring 2015b, Pnar_Language_Archive.FPAHM_007)

The causative SEND construction logically develops out of a compounded structure in which the head—initially a lexical verb meaning ‘send’—undergoes the process of grammaticalization and develops causative semantics as a metaphorical extension stemming from the lexical meaning of ‘send’. As the Khasian languages are head-initial and Tibeto-Burman languages are head-final, it stands to reason that the grammaticalized causative morpheme is a prefix in Khasi and a suffix in Mongsen Ao. The status of *p̐ri-* as the semi-grammaticalized head of its own predicate in Kham similarly reflects its historical source as a lexical verb that is in a nascent state of grammaticalizing as a periphrastic causative morpheme.

Such a lexical source for causatives is not reported in other Austroasiatic languages (e.g. Anderson 2007; Jenny and Sidwell 2015), or indeed further afield (e.g. no examples are discussed in Heine and Kuteva 2002 either), so the SEND causative of Khasi may well have developed via contact with Tibeto-Burman languages. It is furthermore highly probable that the lexical verb *p^{ha}ʔ* ‘send’ of the Khasian languages Khasi and Pnar is also a borrowing, as it reportedly has no parallels in any other Austroasiatic language (Mathias Jenny, pers. comm. 29 May 2017; Paul Sidwell, pers. comm. 30 May 2017). This makes it all the more likely that a contact-induced transfer is responsible for its causative meaning in Khasi; the causative grammaticalization trajectory as well as its lexical source may have resulted from a language-contact scenario.

Heine and Kuteva (2003: 350) discuss evidence that might be used in making the case for replication of a grammaticalization pattern, and propose that linguistic transfer can constitute any of the following:

- forms, that is, sounds or combinations of sounds;
- meanings (including grammatical meanings) or combinations of meanings;
- form–meaning units or combinations of form–meaning units;
- syntactic relations, i.e. the order of meaningful elements;
- any combination of (a)–(d).

The fact that neither a lexical verb with a cognate form meaning ‘send’ nor a SEND causative has been reported in any other Austroasiatic language strengthens the assumption that the pattern as well as the form has been borrowed from an unrelated neighbouring language, thus implying that (a–c) could all be involved, although at this stage it is not possible to identify the source that could have served as the model.

From a cognitive perspective, the inherently causative semantics associated with the meaning of ‘send (on an errand)’ seems to make it an ideal target for the grammaticalization of a causative meaning,¹⁶ but the most common lexical source for a morphological causative in both South Asian and mainland Southeast Asian languages turns out to be the verb ‘give’ (see Matisoff 1991). This verb is known to grammaticalize a wide range of meanings in addition to encoding causative semantics, including benefactive, permissive, and purposive senses (e.g. see Heine and Kuteva 2002: 149–55).

In Mongsen Ao varieties, a suffix *-(p)iʔ* serves as a morphological causative.¹⁷ This morpheme is cognate with the reconstructed Proto-Tibeto-Burman form **bɔɣ* ‘give’ (Matisoff 2003: 132, 200, 442, 480).¹⁸ A reflex no longer occurs synchronically in Mongsen Ao as a lexical verb, having since been replaced by a newer form *khiʔ* ‘give’. The older form only survives as a grammaticalized causative morpheme, as demonstrated in (20).

(20) Mongsen Ao (Tibeto-Burman), Nagaland (Coupe 2007a: 480)

maŋmətuŋ tʃakiɿ, túŋkhəla, kinijunəɿ hlaið kəm
 maŋmətuŋ tʃak-iʔ-əɿ túŋkhəla kinijunəɿ hlà-əɿ
 village.name leave-CAU-SEQ 3PL village.name descend+go-SEQ
 kəm
 become.PST
 ‘Having left [the corpse in] Mangmetong village, they went down and founded
 Kiniunger village.’

¹⁶ Although a causative meaning is not necessarily the only outcome for the grammaticalization of ‘send’ in every South Asian language. For example, Slade (2013: 558–9) notes that Nepali *paṭhāunu* ‘to send’ is used as a light verb to express regret:

mai-le ke gar-i paṭhā-em?
 I-AGT what do-ABS send:PP-1SG
 ‘Oh what have I done?’

This is evidence for assuming that there can also be completely unrelated metaphorical extensions of grammaticalized meaning for a given lexical morpheme in different languages, and that the unique conceptualizations of a speech community may result in substantial deviations from a putative universal cognitive schema based on shared human experience.

¹⁷ The form of the morphological causative of Mongsen Ao varies significantly from village to village in Nagaland. In the Waromung and Khar village varieties the initial consonant has been lost, and the vowel of this and other functional morphemes has centralized and rounded to /u/. In the Khensa village variety the initial consonant is retained (fortuitously revealing beyond doubt this grammatical morpheme’s lexical source), whereas in the Mangmetong village variety it has been lost (for further details see Coupe 2003: 41, 2017b: 293).

¹⁸ The ‘y’ in this reconstructed proto-form deviates from conventional IPA representation in Matisoff’s (2003) reconstructions, and actually represents the palatal approximant /j/, not the high front rounded vowel /y/.

In what appears to be the beginning of a renewed cycle of causative grammaticalization, *khi?* is in the process of developing a permissive meaning in periphrastic constructions, in addition to its basic meaning lexical meaning of ‘give’. This is demonstrated in (21), in which the meaning of *khi?* hovers ambiguously between the semantics of the original lexical sense of ‘give’ and a new permissive interpretation. The permissive meaning is abetted by the dative case marking, which encodes a volitionally acting causee argument in indirect causativized clauses formed with the morphological causative *-i?* for some semantic verb classes (see Coupe 2007a: 191–2). Note that both the structure and the semantics of the grammaticalized meaning of *khi?* align it with the periphrastic causative of Kham, as illustrated in (17a,b).

- (21) Mongsen Ao (Tibeto-Burman), Nagaland (Coupe 2007a: 198)

nì nà niŋ tʃāi li áhjá? phàjpà? khiw?

nì nà niŋ tʃāi li á-hjá? phà?-i-pà? khi?-ù?
1SG AGT 2PL.POSS SON DAT NRL-fish catch-IRR-NMLZ give.PST-DECL

(i) (Your son wanted to catch fish) ‘I let your son catch fish.’

(ii) ‘I gave fish to your son to catch.’

Anderson (2007: 237) reports that an auxiliary verb construction involving the root *beɽ-* ‘give’ in the Munda language Gutob encodes a kind of causative or resultative meaning denoting an effect on the causee argument. Examples (22a,b) respectively contrast the function of the grammaticalized causative morpheme with the main verb usage in (22c).

- (22) Gutob (Austroasiatic), Eastern India

a. uson gol-gol-te nom bobrig-o? beɽ-o?
today smoothly you make.enter-PST GIVE-PST
‘Today you put it in smoothly.’ (Hook 1991: 185)

b. sobu paiṭi niŋ ɖem-o?niŋ beɽbe?niŋ
all work I do-PST.II=1 RDP:GIVE=1
‘I do all the work.’ (Zide 1997: 323, cited in Anderson 2007: 237)

c. niŋ niŋ-nu onoo?n beɽ-o?=niŋ suŋ-tu
I I-GEN daughter give-PST.II=1 INT-FUT.II
‘I will give my daughter.’ (Zide 1997: 323, cited in Anderson 2007: 236)

Nagamese is best characterized as a creole-like language with an Assamese base that is widely spoken as a lingua franca in the northeastern state of Nagaland. In common with other languages of the subcontinent, a verb with the meaning of ‘give’ is used with causative semantics when occurring periphrastically in series with another verb. That the final verb is functioning as a grammaticalized element in (23) is proven by the fact that the construction predicates a single event: that of falling down. In contrast, buying a book is a separate event from the act of giving it in (24). It also seems to be the case, according to native-speaker consultants, that the matrix verb *dise* in (23) is essential for deriving the causative meaning, despite the presence of the causative suffix on the non-finite verb.

- (23) Nagamese (Indo-Aryan), Nagaland (author's fieldnotes—elicited data)
 didi laga kutta to ami-ke gira-a-i di-se
 older.sister GEN dog TOP 1SG-ACC/DAT fall-CAU-NFIN GIVE-PST
 'Older sister's dog made me fall down.'

- (24) Nagamese (Indo-Aryan), Nagaland (Bhattacharjya 2001: 214)
 Moy tai-ke ekta kitab kin-i-kena di-se
 1SG 3SG-ACC/DAT one-CLF book buy-NFIN-CNV give-PST
 'I bought her a book.'¹⁹

In Bhattacharjya's (2001) thesis, one can also find examples of causativized clauses in which the causative suffix *-a* is absent, so the entire functional load for expressing a causative meaning is then carried by grammaticalized 'give':

- (25) Nagamese (Indo-Aryan), Nagaland (Bhattacharjya 2001: 348, 377)
- a. bagan to ami-khan ini rakh-i di-le, ... ek-bar
 garden TOP 1-PL fallow keep-NFIN GIVE-COND, ... one-time
 lai.pata laga-bo are ini alchi pora rakh-i di-she
 greens plant-IRR and fallow idle ABL leave-NFIN GIVE-PST
 'If we leave the garden fallow ... if we plant greens once and then leave them uncultivated ...'
- b. Hey, sala-ke dhur-i-bi, no-char-i-bi;
 EXCLM, bastard-ACC/DAT catch-NFIN-IMP NEG-release-NFIN-IMP
 theng bhang-i di-bi;
 leg break-NFIN GIVE-IMP
 itu kotha ki band-i-kena yate rakh-i di-bi
 this talk what tie-NFIN-CNV here keep-NFIN GIVE-IMP
 'Hey, get the bastard; don't let him run off. Break his legs, tie him up and dump him right here.'

While the grammaticalization of the periphrastic GIVE causative may appear at first to be a straightforward case of layering in Nagamese, on deeper inspection it could be motivated by a functional need to express the indirect causation of intransitive verb bases. An intransitive verb stem taking the *-a* causative suffix seems to require its causee to be a patient, whereas this semantic entailment does not necessarily apply with the periphrastic GIVE causative in the absence of the causative suffix. This is captured by the elicited example in (26), in which a permissive meaning obtains from an infinitival verb stem + GIVE construction and the causee is interpreted to be acting volitionally as an agent (cf. the semantically and structurally equivalent Mongsen Ao example of (21) above).

- (26) Nagamese (Indo-Aryan), Nagaland (author's fieldnotes—elicited data)
 tai didi ke dza-bole di-se
 3SG older.sister DAT/ACC go-INF GIVE-pst
 'S/he let older sister go.'

¹⁹ Bhattacharjya's glossing and interlinearizations have been adjusted for consistency.

Masica (1991: 319) discusses the parallel case of Bengali, which has no means of deriving an indirect causative from an intransitive stem via the old Indo-Aryan vowel strengthening method (e.g. Hindi intransitive *uṭhnā* ‘to arise’ vs transitive *uṭhānā* ‘to raise, lift’) and consequently resorts to periphrastic expression. Watters (2004: 229) makes the relevant point that a periphrastic causative in Kham does not trigger a reassignment of semantic roles. The agent of a non-causativized clause remains the agent of the periphrastic causative, whereas a morphological causative may obligatorily recast the agent as a patient, potentially creating a semantic incompatibility. Speakers of languages that do not have a way of expressing indirect causation with the grammatical tools at their disposal may well take a periphrastic pathway to grammaticalizing a new causativization strategy—as suggested by the periphrastic causatives of Kham and Nagamese—but for the present this remains a topic in need of deeper exploration.

10.6 ‘EAT’ > PASSIVE/MIDDLE/RECIPROCAL/ REFLEXIVE MARKING

In Indo-Aryan and Munda languages, the verb ‘eat’ occurs in a range of idiomatic expressions, the majority of which are consistent with a general meaning of adverse experience.

- (27) Early New Indo-Aryan (Jaworski and Stroński 2017)²⁰
 paṃkhinha dekhi sabanhi ḍara khāvā
 bird:DEF:M:OBL:PL see:CNV all:OBL fear:INDF:M:NOM:SG EAT:M:PP:SG
 ‘... birds having seen all of that got scared / ... birds saw all of that and got scared.’
- (28) Sinhala (Indo-Aryan), Sri Lanka (Keenan 1985: 259)
 kikili lamajagen maerun kae:va
 chicken child(INS) death EAT
 ‘The chicken was killed by the child.’
- (29) Assamese (Indo-Aryan), Assam (author’s fieldnotes)
 naspati ε-khon lo-bō bisaI-is-e, lo-l-e,
 pear one-CLF take-PURP seek-PRS-3 take-PST-3
 kintu bhōi kha-i as-e
 but fear EAT-CNV exist-3
 ‘[He] is seeking to take a pear, [and] took one, but is afraid.’

²⁰ The example is from an epic poem composed c.1540 in Old Awadhi by Malika Mohammada Jāyāsī. See Mātāprasāda Gupta (ed.), *Padmāvatā* (Ilāhābāda: Bhārati Bhaṇḍāra, 1963), vol. 69.2, p. 61. I am grateful to Rafał Jaworski and Krzysztof Stroński for bringing it to my attention.

- (30) Nagamese (Indo-Aryan), Nagaland (author's fieldnotes)

bhoi **kha**-ise?

fear **EAT-PST**

[context: the interlocutors have just narrowly avoided a collision whilst driving] 'Did you get a fright?'

Anderson (2007: 41–2) speculates that a quasi-passive marker *-dʒom* in the Munda language Kharia might have grammaticalized from a verb 'eat', and a cognate form also occurs in the related language Juang. In addition, Kharia extends the use of the morpheme *-dʒom* to a kind of emphatic, reflexive, and 'indirect-middle' marking. Whereas the grammaticalized use of 'eat' with nouns such as 'fear' parallels the NV structure of light verbs in Indo-Aryan, in which the noun functions as the O complement of the verb, in Munda and Tibeto-Burman the grammaticalized morpheme occurs in the V_2 slot of what was almost certainly a compound verb construction at an earlier stage.

- (31) Kharia (Austroasiatic), eastern India

io-**dʒom**-ta

see-EAT-PRS

'it is seen' (Grierson 1906: 195, cited in Anderson 2007:42)

- (32) Juang (Austroasiatic), eastern India

aiŋ ma'd-**dʒim**-sɛkɛ

I beat-EAT-PFV:PRS

'I am beaten' (Pinnow 1960: 115, cited in Anderson 2007: 42)

The Kiranti language Yakkha of eastern Nepal has similarly grammaticalized a lexical root meaning 'eat' as a type of reflexive/reciprocal marker. Grammaticalized 'eat' is used in the V_2 position in compound verbs, where it can additionally express autobenefactive meanings.

- (33) Yakkha (Tibeto-Burman), eastern Nepal (Schackow 2015: 14)

nda (aphai) moŋ-**ca**-me-ka=na

2SG (self) beat- V_2 .EAT-NPS-2=NMLZ:SG

'You beat yourself.'

Schackow (2015: 303–5) notes that *ca* in V_2 position has a number of polysemous meanings; e.g. in *kon-ca* 'walk-EAT', she interprets grammaticalized *ca* as contributing a nuance of 'consuming' the enjoyment of taking a walk. As in other languages of South Asia, EAT is used in some contexts to express an adversative passive-like meaning.

- (34) Yakkha (Tibeto-Burman), Eastern Nepal (Schackow 2015: 305)

moŋ-**ca**-khuba babu

beat- V_2 .EAT-S/A:NMLZ boy

'the boy who gets beaten up (regularly)'

Looking beyond the subcontinent, the metaphorical extension of a Turkish verb meaning 'eat' to a broad range of adversative idiomatic expressions is reported by Friedman (2003: 13–14), who observes that a similar usage has been calqued in

Macedonian from the verb *jade* ‘eat’ (cited in Heine and Kuteva 2005: 53). This results in adversative expressions such as *jade k’otek* ‘get a beating’ (lit. ‘eat a blow’) and is based on the equivalent Turkish expression *kötek yemek*.

Lastly, Hook (1991: 190) presents some examples of grammaticalized ‘eat’ in the Munda languages Mundari and Ho that do not express adversative meanings:

- (35) Mundari (Austroasiatic), eastern India (Hoffman 1950: vol. 7, 2098)

en horoko lel **jom-me**
 those people see **EAT-IMP**
 ‘Take a look at those people.’

- (36) Ho (Austroasiatic), eastern India (Burrows 1915: 89)

umbul-re dub **jom-pe**
 shade-in sit **EAT-IMP**
 ‘Sit (at ease) in the shade.’

Given the diverse and rather arbitrary range of meanings attributable to grammaticalized ‘eat’, it seems to be the case that there is no uniform shared grammaticalization trajectory that this verb could have taken. The polysemous meanings that have grammaticalized furthermore suggest that these are unlikely to be the outcome of a contact-induced grammaticalization in the languages of South Asia.

10.7 ‘SEE/LOOK’ > CONATIVE MODALITY ‘TRY, TEST OUT’

Representatives of Munda, Dravidian, Indo-Aryan, and Tibeto-Burman languages of South Asia all have a verb with the meaning of ‘see’ or ‘look’ that appears to be on the pathway to grammaticalizing, or to have already grammaticalized, a conative modality expressing a meaning of ‘to try, test out’.

- (37) Gorum (Austroasiatic), eastern India (Rau, in prep.: 393)

pans din zom-ej-ju? sun ga?-t-ej *giʔ*-t-ej
 five day gather-LV-MID:PST say eat-ACT:NPS-3PL **see-ACT:NPS-3PL**
 ‘When it has gathered for five days, they try it (by drinking).’

- (38) Tamil (Dravidian), Tamil Nadu and Sri Lanka (Lehmann 1993: 226–7)

- a. kumaar caṭṭai-y-ai-p poot-t-p **paar-tt-aan**
 Kumar shirt-ACC put-CNV **see-PST-3SG:M**
 ‘Kumar put on the shirt (e.g. to see if it fits).’
 b. kumaar inta naaval-ai-p paṭi-ttu-p **paar-tt-aan**
 Kumar this novel-ACC study-CNV **see-PST-3SG:M**
 ‘Kumar tried reading the novel (to see how it was).’

- (39) Nagamese (Indo-Aryan), Nagaland (author’s fieldnotes)

kha-i **sa-bi** **na**
 eat-NFIN **look-IMP** PTC:EXHORT
 [in the context of utterance:] ‘Try tasting it.’

A grammaticalized form of the lexical verb *atsə* ‘look’ occurs in Mongsen Ao verb stems, where it expresses a conative modality meaning consistent with ‘test, try out’.

(40) Mongsen Ao (Tibeto-Burman), Nagaland (Coupe 2007a: 332–3)

a. *təʃhàku, tənì nə ‘màkəʔatsəʔu? pi.’*

tə-ʃhà-ku tənì nə mə-kəʔa-tə-əʔ-ù? pi
 thus-do-LOC.CNV RL-wife AGT NEG-ascend+come-CON-PRS-DEC PROX
 ‘And so, the wife says [of her husband] “[He] isn’t attempting to come up
 [from the Assam plain], this one”.’

b. *təʔ lɪŋəʔ, ‘àhlù nə watsəŋ, tə ʃhàw.ʔə.*

tə-əʔ lɪŋ-əʔ a-hlù nə wə-tə-əŋ
 thus-SEQ plant-SEQ NRL-field ALL go-CON-IMP
 tə ʃhà-ù? təʔ
 thus do.PST-DEC RPR
 ‘And then, after [he] had done the planting, [she] said “Go and have a look
 at the field”.’

The grammaticalization of a conative modality meaning from verbs of visual perception is observed to cluster in areas of high linguistic diversity, as already demonstrated by the languages of New Guinea and discussed in section 10.3. A similar clustering is presented by a number of languages belonging to four language families of South Asia. The ubiquity of an identical grammaticalization pattern targeting semantically equivalent verbs in unrelated languages must surely be attributable to the contact-induced transfer of a conceptual schema, as in both the New Guinea and South Asian regions it is too concentrated to be merely due to chance.

10.8 RELATIVE–CORRELATIVE CONSTRUCTIONS

Up to this point we have considered grammaticalization processes applying mostly to individual lexical items that gradually evolve as functional morphemes in syntagmatic collocations. We now turn to a consideration of South Asian developments applying to larger syntactic structures—specifically, relative–correlative constructions—which demonstrate a grammaticalized function for interrogative pronouns in some languages, as well as evidence of an areal diffusion of the pattern.

While it can be potentially challenging to identify grammaticalization outcomes that result unambiguously from language contact, the case for relative–correlative constructions arising out of contact scenarios in South Asia seems beyond doubt. According to Nadkarni (1975), the relative–correlative construction is native to Indo-Aryan languages of South Asia, and has spread into Dravidian as a result of contact-induced convergence. Supporters of this conjectured direction of diffusion have argued that the relative–correlative construction with its two finite verbs violates a constraint on Dravidian syntax that only permits a single finite verb per sentence (with the exception of quotations), thus suggesting that the structural pattern must have been borrowed from Indo-Aryan. However, an opposing view proposes that relative–correlative constructions are native to Dravidian, and that the

spread has gone in the opposite direction (see Kolichala 2016: 88 for discussion and references, also Hock 1989). Regardless of what may be the correct interpretation for the direction of spread, it is indisputable that the relative–correlative structure has been replicated in Tibeto-Burman languages as a result of contact; and it will be shown that these languages are very similar to Dravidian in co-opting interrogative pronouns in an innovated role for marking the relative clause constituent.

The structural and morphological characteristics of the New Indo-Aryan relative–correlative construction are as follows. Two clauses are adjoined, both containing a finite verb. The dependent ‘relative clause’ is preceded by a so-called ‘j-class’ relative pronoun (REL) that marks the relativized argument and indicates the subordinate status of its clause, and the relativized NP argument is coreferential with a potentially optional noun or a pronoun in the matrix clause that functions as the correlative NP (CORREL). The Hindi example of (41) illustrates this structure.

- (41) Hindi (Indo-Aryan), North India (Kachru 2006: 4)

jo kitaab mez per he
 REL book.F.SG table.F.SG.OBL on be.PRS.SG
 vāh merī he
 CORREL 1.POSS.F.SG be.PRS.SG
 ‘The book which is on the table is mine.’

The relative–correlative construction is of considerable antiquity, and is attested as early as Vedic Sanskrit. As in all the daughter languages of Indo-Aryan, the position of the matrix clause vis-à-vis the dependent clause is pragmatically determined according to whether the modifying information is restrictive or non-restrictive, and this may have been a factor influencing its diffusion into the grammars of other South Asian languages. Examples (41) and (42) respectively contrast restrictive and non-restrictive interpretations of meaning.

- (42) Vedic Sanskrit (Indo-Aryan) (Hock 2016: 20)

tvaṁ taṁ ... bādhasva ...
 you:NOM:SG that:ACC:SG:M bind:IMP:2SG
 ... yo no jighāmatī
 who:NOM:SG:M we:OBL.CLIT slay:DES:PRS:3SG
 ‘You ... tie down that (evil-doer) who ... tries to slay us.’ (Rig Veda 6.6.2)

Turning now to the Dravidian relative–correlative, we find that it similarly involves two finite verbs; but because Dravidian languages lack a form class of relative pronouns, speakers must press into service an interrogative pronoun for marking the dependent relative clause of the construction. This is demonstrated in the Malayalam example of (43).

- (43) Malayalam (Dravidian), Kerala (Asher and Kumari 1997: 53)

aarə manassə aṭakkunnuvo-o avannə samaadhaanam kiṭṭunnu
 who mind control:PRS-IP he:DAT peace obtain:PRS
 ‘He who controls the mind obtains peace.’

In some Dravidian languages, a particle (typically an interrogative with the form *-o*) marks the end of the relative clause constituent. An identical pattern of using an interrogative pronoun in lieu of a relative pronoun is replicated in Tibeto-Burman languages employing the relative–correlative construction. Mongsen Ao is also somewhat similar to Dravidian languages in using a topic particle *la* at the end of the relative clause to indicate its dependent status.

- (44) Mongsen Ao (Tibeto-Burman), Nagaland (author's fieldnotes)
 sápa? nā ì tǝlāj ā-tshā phānā tsāŋ-i?-.iū lā
 who AGT 1PL.INCL daughter NRL-mithun five attach-CAU-IMM TOP
 āji tǝlāj pā tsā-i-ù? tē sā-Ø
 1PL.POSS daughter 3SG take-IRR-DEC thus say-PST
 ‘“Whoever ties five mithuns (*Bos frontalis*) [as a bride price for] our daughter, he can take our daughter,” [he] said.’

Tibeto-Burman languages that make limited to extensive use of the relative–correlative construction are typically spoken in locations buffering communities speaking Indo-Aryan languages, and where bilingualism in an Indo-Aryan language is common. Elsewhere in the Tibeto-Burman domain the relative–correlative pattern is not attested, and speakers exclusively use the native Sino-Tibetan nominalized participle type of relative clause construction, as demonstrated by the Khamniungan example of (45). See (16a) to compare an internally headed example in Mongsen Ao.

- (45) Khamniungan (Tibeto-Burman), Nagaland (author's fieldnotes)
 ...fawʔ¹¹ nā³¹, ko³³-khej³³-le³³ nōj¹¹-tʃən³³ nā³¹,...
 rat DEM earth-INSIDE-LOC stay-NMLZ DEM
 ‘... the rat, which lives inside the earth, ...’

A possible functional motivation for South Asian languages replicating the relative–correlative pattern is that, whereas access to relativization using the participle type of relative may be limited by language-specific constraints, there appears to be no such restrictions on access to relativization using the relative–correlative strategy. This logically motivates the replication of the relative–correlative construction in Tamil, a language that prohibits relativization on an instrument and other oblique positions using the standard participle relativation strategy native to Dravidian, but permits it using the relative–correlative pattern.

- (46) Tamil (Dravidian), Tamil Nadu and Sri Lanka (Keenan and Comrie 1977: 73)
 Enna(k) katti(y)-āl korī(y)-ai anta maṇitaṇ kolaippi-tt-āṇ
 which knife-with chicken-ACC that man kill-PST-3SG:M
 anta katti(y)-ai jāṇ kaṇ-ṭ-āṇ
 that knife-ACC John see-PST-3SG:M
 ‘John saw the knife with which the man killed the chicken’ (lit. ‘with which knife the man killed the chicken, John saw that knife’)

Genetti (2007: 488) additionally notes that the relativized referents in relative–correlative constructions of Dolakha Newar are typically indefinite, unknown, and non-specific, a pattern that is also common to the majority of relativized referents of relative–correlative clauses in Mongsen Ao. The relative–correlative construction may therefore fill a functional gap in the structural inventories of some languages by facilitating relativization on indefinite and non-specific referents, or by making it possible to relativize on arguments that are not accessible using a native participle relativization strategy.

- (47) Dolakha Newar (Tibeto-Burman), Nepal (Genetti 2007: 487)
gunān bāmā=e khā ṇen-ai āmun sukha sir-ai
 who:ERG parent=GEN talk listen-3SG:PR 3SG:ERG happiness know-3SG:PRS
 ‘Whoever listens to his/her parent’s advice, s/he knows happiness.’

As demonstrated above for Indo-Aryan, the position of the relative clause constituent in the relative–correlative construction can be before the matrix correlative clause, where it encodes restrictive reference, or alternatively after the matrix correlative clause, where it encodes non-restrictive reference. A number of Tibeto-Burman languages also permit pre- and post-head relative clause placement using their participle relativization pattern similarly to encode a restrictive~non-restrictive contrast (e.g. see Coupe 2017b), but perhaps not all do. If there are rigid constraints on the position of the head that precludes encoding this contrast by means of constituent order, then this could provide another motivation for languages of South Asia replicating the relative–correlative structure and using it alongside the participle type of relative clause.²¹

The relative–correlative construction is also found in Munda languages. Kharia, for example, has borrowed the set of j-class relative pronouns from Indo-Aryan, one of which is used to mark the beginning of the relative clause in (48a). Intriguingly, Kharia has an additional set of relativizing forms that are homophonous with the proclitic interrogative markers (e.g. 48b), paralleling the grammaticalized use of interrogative morphemes in Dravidian and Tibeto-Burman languages for marking relative clauses (see Peterson 2011: 408 for the full paradigm).

- (48) Kharia (Austroasiatic), eastern India (Peterson 2011: 408–9)
 a. ...*je khajar tar=sikh=o?*=may *ho=kəɽ=a?* *komaŋ=ko*
 REL deer kill=PFV=ACT.PST=3PL that=SG.HUM=GEN meat=CNTR
 nalage, ...
 NEG.QUAL.PRS
 ‘...it isn’t the meat of the deer they had killed ...’ (...which deer they had killed, **his** meat it is not ...)

²¹ See Heine and Kuteva (2005: 127ff.) for similar examples of replicating languages filling functional gaps in their grammars in other linguistic areas of the world.

- b. ... *a=boʔ=te* *pujapaʔh* *karay=na* *aw= ki* *ho* *boʔ=te*
 Q=place=OBL sacrifice do=INF QUAL-MID.PST that place=OBL
 $\dot{d}am=ke$ $ho=ki$ ho $\dot{d}oli=te$ $ma\dot{r}ay=oʔ=may$
 arrive=SEQ that=PL that palanquin=OBL put.down=ACT.PST=3PL
 ‘... having arrived at the place where the sacrifice was to be done, they put
 the palanquin down.’

10.9 CONCLUDING COMMENTS

This chapter has considered grammaticalization phenomena from an areal perspective, and has found correlating evidence for the contact-induced transfer and replication of patterns involving unrelated languages of South Asia. Perhaps the most convincing of these is the SEND causative of Khasi. This satisfies all of the criteria for having spread through contact, since neither the pattern nor the lexical source can be linked to related languages outside South Asia. If the Khasi SEND causative prefix has indeed resulted from a contact-induced grammaticalization, then it may be added to the list of counterexamples demonstrating that there is no requirement for languages to share structural compatibility for contact-induced grammaticalization to occur (see e.g. Harris and Campbell 1995: 124–6).

Secondly, the development of conative modality from verbs of visual perception in four language families of South Asia presents yet another convincing case of contact-induced grammaticalization of a conceptual category, due to the fact that the verbs of these unrelated languages all follow the same grammaticalization trajectory precisely leading to a conative modality outcome, and the pattern is furthermore observed to be cross-linguistically rare.

Lastly, the relative–correlative construction similarly presents robust evidence for contact-induced grammaticalization, as relative–correlative constructions in Tibeto-Burman languages conform to the criterion of only being found in languages within the linguistic area of South Asia. The fact that Dravidian, Munda, and Tibeto-Burman languages all press their interrogative pronouns into service as relative pronoun equivalents (or even borrow the *j*-class relative pronouns for this function, as in the case of Kharia) is very convincingly the replication of a syntactic pattern. Functional motivations for copying this construction can be identified, as noted in the case of Tamil, which can use the relative–correlative construction to relativize on arguments that are otherwise inaccessible to relativization using the standard Dravidian participle construction.

The data presented in this chapter collectively demonstrate the transfer of seemingly identical conceptual schemas across the genetic boundaries of languages in contact; these target morphemes or constructions with identical meanings in unrelated languages, and they produce the same grammaticalization outcomes. Such replicated patterns must cater to a multilingual community’s communicative needs, while at the same time reducing the cognitive burden imposed by multilingualism in a linguistic area.

ACKNOWLEDGEMENTS

I thank Sander Adelaar, Nikolaus Himmelman, Uta Reinöhl, and the editors for their comments and suggestions on an earlier draft. I alone bear responsibility for the conclusions reached and any misinterpretations of analysis. The chapter was written while I was affiliated to the Institute for Linguistics at the University of Cologne, and the background research was facilitated by an Alexander von Humboldt Research Fellowship for Experienced Researchers. I am grateful to both of these institutions for their generous support.