

Parallel innovation and ‘coincidence’ in linguistic areas: on a bi-clausal extent/result construction of mainland Southeast Asia¹

N. J. ENFIELD

Max Planck Institute for Psycholinguistics, Nijmegen

The notion of a ‘linguistic area’—‘a geographical region containing a group of three or more languages that share some structural features as a result of contact rather than as a result of accident or inheritance from a common ancestor’ (Thomason 2000:311)—presupposes that neighbouring languages can share structural features for three reasons: (1) due to ‘genetic inheritance’ from a common ‘ancestor’ language; (2) due to ‘diffusion’ of features from one language into the other (or indeed, from a third language into both) via language contact; (3) due to accident or coincidence. There is another way, however, in which neighbouring languages may share structure, namely due to a certain kind of parallel yet independent innovation, one which is not due *directly* to language contact, but which also is not accidental. Certain grammatical structures which languages share due to previous historical contact may cause certain innovations to be very likely for language-internal reasons, and these innovations may arise spontaneously and independently in separate languages.

The issue arises when we consider the status of grammatical constructions—i.e. morphosyntactic structures with specifiable meanings—in linguistic areas. This paper discusses a bi-clausal construction with ‘extent/result’ semantics found in a number of languages of the mainland Southeast Asia area. Mainland Southeast Asia is the home of hundreds of languages from at least six distinct language families—Austronesian, Hmong-Mien, Mon-Khmer, Sinitic, Tai, and Tibeto-Burman—which show a high degree of parallelism in a number of structural domains (Enfield in press, Chapter 2; cf. Capell 1979, Clark 1989, Bisang 1991, Matisoff 1991, 2001, Enfield 2001). While the structure and meaning of a construction can appear closely analogous across languages, the lexical ‘marker’ employed for that construction is in many cases not cognate (even when the two

¹ This paper is dedicated to Jim Matisoff, in humble appreciation of a true pioneer and a very nice fellow.

languages share cognates of a given marker). When two languages share a construction marked by non-cognate lexical material, one would normally assume they did not inherit the construction from a common ancestor. In such a case, the construction may have been calqued (i.e. where a language has borrowed an *idea* of structuring meaning in a certain way, and has not borrowed any linguistic *form* with it), or it may have been separately and independently innovated in the languages. The existence of non-cognate structural parallelism raises interesting questions regarding language contact and change. The construction and its 'marker' (i.e. the stable lexical material which identifies it) are logically separate with respect to the possible avenues of explanation of structural parallelism—shared genetic inheritance, contact-related diffusion, or 'coincidence'.

1. Extent/result complement constructions

Many languages of mainland Southeast Asia share a construction with the following properties:

structure: CLAUSE-1 "MARKER" CLAUSE-2
 meaning: 'Clause-1 was the case to such an extent that
 Clause-2 became the case as a result'
 (or: 'It was so Clause-1 that Clause-2').

The semantic relationship between CLAUSE-1 and CLAUSE-2 involves both cause and temporal subsequence (namely, 'CLAUSE-1; because of this, after this, CLAUSE-2', with the more specific idiomatic English translation *so V1 that V2*).

We first consider Lao (Southwestern Tai), Kmhmu Cwang (Eastern Mon-Khmer), Hmong (Hmong-Mien) and Vietnamese (Eastern Mon-Khmer). In these languages, the markers of the extent/result construction are derived from lexical items which are non-cognate, but which each express at an abstract level an idea of 'coming to be in the same place', common to concepts such as 'reach', 'arrive', 'connect', 'up to', and 'touch'.

Speakers of Lao mark the construction with *con*³ 'up to, until':

- (1) *khaw*³*ten*⁴ *con*³ *muaj*² *laaj*³
 Lao 3PL jump until tired very
 'They jumped until they were very tired; They jumped so much they were very tired.'
- (2) *khaw*³*haj*⁵ *con*³ *phaa*⁵-*set*¹-*muu*² *piak*⁵ *met*²
 Lao 3PL cry until cloth-wipe-hand wet all
 'They cried until their handkerchiefs were all wet; They cried so much their handkerchiefs got all wet.'
- (3) *dang*³ *con*¹ *khon*² *nəən*² *bə*⁰ *lap*²
 Lao loud until person lie NEG sleep
 '(They) made so much noise people couldn't sleep.'

Bi-clausal extent/result construction of MSEA

The following example shows the marker *con*³ with its purely temporal meaning ‘until’:

- (4) *man*² *juu*¹ *huan*² *con*³ *haa*⁵ *moong*²
Lao 3SG be.at houseuntil five o'clock
‘He was at home until five o’clock.’

In Kmhmu Cwang, the construction is marked with *cam* ‘connect to, link’:

- (5) *No* *tèng nan cam gôn yat da' gang leh*
Km 3PL make loud connect person LOC be.at housebeside
ni' sih am bwan.
this sleep NEG can
‘They made such a noise that the people in the next house were unable to sleep.’
- (6) *Ge 'mook nangsw cam uat tnoh.*
Km 3MSG tell writing connect tired mouth
‘He taught until his mouth was tired.’

Here are two examples showing *cam* with the meaning ‘up to’, ‘touching’, ‘connecting’:

- (7) *Ô' yoh cam Viangcan.*
Km 1 go reach V.
‘I went (all the way) to Vientiane.’
- (8) *cam yo'*
Km touch RCP
‘touching (of two things)’

In Hmong, the extent/result complement construction is marked with *txogqhov* ‘until, up to’:

- (9) *Nws ua ntshoo txogqhov kuv pw tsis tsawgzog ib hmos.*
Hm 3 do/make noise until I lie NEG sleep one night
‘They made (such a) noise that I couldn’t sleep all night.’
- (10) *Nws sau ntawv txogqhov tsaug leeg tes tas.*
Hm 3 write writing until tired sinew hand all
‘He wrote so much his hands were completely tired.’

The following example shows Hmong *txog*, the first element in *txogqhov*, meaning ‘reach’:

- (11) *Kuv mus tsis txog phoosxabvam.*
 Hm 1 go NEG reach P.S.
 'I didn't get to Phone Savanh.'

In Vietnamese, the extent/result complement construction is marked by *đến-nối* 'until' (where *đến* appears elsewhere as a main verb meaning 'reach'):

- (12) *Anh ấy làm tiếng động đến-nối nhà bên cạnh*
 Vn man that make sound loud until housedirection side
không ngủ được.
 NEG sleep can
 'He made loud noise such that the (people in the) house next door couldn't sleep.'
- (13) *Anh ấy dạy đến-nối mệt.*
 Vn man that teach until tired
 'He taught until he was tired.'

The structure and meaning of these constructions are essentially the same across the languages (although of course one would expect to find minor grammatical distinctions). One explanation for this would be that the constructions have been calqued across two or more of these languages (i.e. the *idea* of the construction was borrowed across languages, with each individual language providing its own constructional marker). Another explanation would be that the similarity of these constructions is merely a coincidence. But how could such a 'coincidence' come about? A semantically and structurally specific construction does not just happen to exist in a language, but develops historically. And historical development of particular grammatical constructions is contingent upon existing grammatical and semantic structure, which constitute input to well-known processes of inference and extension which, in turn, drive semantic/grammatical change (Hopper and Traugott 1993, Harris and Campbell 1995, Traugott and Dasher 2002, *inter alia*). It is often difficult if not impossible to tell whether a construction common to two neighbouring languages has been calqued or independently innovated, and perhaps this is because the distinction is in reality not especially clear (Enfield *in press*: 368). If neighbouring languages independently innovate a semantically and structurally similar construction, chances are that the grammatical and semantic structures which provided the environment for the innovation to develop were also shared, likely due to contact at an earlier stage.²

² This point may be more forcefully made with reference to constructions of greater semantic and structural specificity. See Enfield (*in press*) for detailed discussion.

We turn now to Sinitic languages. The marker of the extent/result complement construction in Modern Standard Chinese (MSC) is *de*, a morpheme historically meaning ‘acquire, obtain’. There is a whole family of constructions in which *de* links a predicate with a complement of some kind, and these have been intensively studied in MSC (cf. Enfield in press: 268). We are here interested only in the construction expressing extent/result, as in the following examples:

- (14) *tā men tiào de hěn lèi*
 MSC 3 PL jump DEvery tired
 ‘They jumped till (they) got very tired.’ (Huang 1988:274)
- (15) *tā men kū de shǒupà dōu shī le*
 MSC 3 PL cry DEhandkerchief all wet PFV
 ‘They cried so much that even the handkerchief got wet.’ (Huang 1988:274)
- (16) *chǎo de rén.jiā shuì bù zháo*
 MSC noisy DEother.people sleep NEG can
 ‘(They) made so much noise that other people couldn’t sleep.’ (Chao 1968:355)

Lamarre (2001) surveys constructions of this kind across Sinitic languages, and shows that despite these languages indisputably belonging to the same language family, they can each display a single construction with very similar semantics and very similar structure, yet use a wide range of different non-cognate lexical items to mark the construction. While the constructional marker in MSC is a verb meaning ‘obtain, acquire’, many neighbouring languages (Sinitic and otherwise) mark the same construction with other verbs, including ‘go’, ‘reach’, ‘come’, ‘attach’, and ‘arise’ (Lamarre 2001). There is a certain abstract commonality in the semantics of these markers, all of them associated with ‘reaching’ or ‘coming together’ (compare the other Southeast Asian languages discussed above).

The following examples show structures analogous to the MSC structures in (14-16), from other Sinitic languages Xinyi (Yuè), Beijing Mandarin (Mandarin), and Cantonese (Yuè), respectively, in which the constructional marker is not cognate with MSC *dé* ‘obtain’, but rather with *dào* ‘reach’ (examples from Lamarre 2001:99-101):

- (17) *k'æi¹³ fun⁵³ hei³⁵ tou³⁵ tai¹¹ tai¹¹ seŋ⁵³ kom³⁵ ham³³*
 Xinyi3SG happy ‘reach’ loud-RDP voice PCL shout
 ‘He was so happy he shouted out.’ (Tang 1986:101)
- (18) *wǒ hèn tā hèn dào tīng-jiàn tā de míng-zì jù shēng-qì*
 BjM 1 hate 3 hate ‘reach’ hear 3 POSS name so get.angry
 ‘I hate her to the point that [or: ‘so much so that’] I get angry simply by hearing her name.’ (Chao 1926:876)

(19) *kui⁵ bei² ngoh³ man⁶ do³ hau² nga²nga²*
 Cant 3 by I ask 'reach' mouth mute

'He was questioned by me so closely he could not answer anything.' (Peng 1993:91; transliteration and translation as given in Lamarre 2001)

While the three languages possess morphemes cognate with MSC *de* (**tak* in Proto-Chinese), it happens that they do not use those morphemes to mark this construction, despite its close parallelism with the same construction in their sister language MSC (with which they are each now in contact). There is, of course, no possibility that when these languages inherited their cognate of MSC *de*, they would have inherited this modern construction with it. Proto-Chinese **tak* itself had little of the present range of functions it shows in MSC today. With extensive historical written records available, historical linguists have attempted to reconstruct the development of **tak* in Sinitic, and they are unanimous that its original meaning was 'obtain, acquire' (Sun 1996:108, Lamarre 2001), and, furthermore, that its modern 'auxiliary' functions are relatively recent developments. In many Sinitic languages the modern descendent of **tak* does not perform the same range of functions as it does in MSC. The extent/result construction described here has either been more recently borrowed across Sinitic languages or it has been independently innovated in the languages, and this distinction has no correlation with whether or not the constructional marker used is cognate. The 'typological poise' of the languages is very similar, increasing the likelihood of independent innovation toward the same outcome (Enfield 2001: 284-287, in press: 358-361).

3. Discussion and conclusion

In the absence of a single genetic origin common to two neighbouring languages, contact-related historical diffusion is often presumed to be the cause of structural parallelism. A possible assumption is that if neighbouring languages *are* known to be 'genetically related', then structural parallelism will *not* be due to diffusion but to their common genetic background. This assumption is not necessarily correct, as the examples from Sinitic languages above have shown. Rightly, the definition of linguistic area given at the beginning of this paper does not require that languages in a linguistic area be genetically *unrelated*. To the contrary, shared 'genetic' background of two languages with a common ancestor may be no more likely an explanation for modern parallels in their grammatical and semantic structure than processes of contact-related diffusion. Once a proto-language has split into two or more separate languages, speakers of those related languages *may remain in contact (or later come back into contact)*, and resultant processes of structural diffusion will be essentially the same as those which pertain between unrelated languages. Languages may have specific constructions in common, and these constructions may or may not be marked by cognate material. But if certain analogous constructional markers in two separate languages *are* shown to be cog-

nate, one cannot conclude from this that the *constructions* they mark were inherited from a common ancestor.

Thus, we may ask: When a language calques a construction from a genetically and/or areally related language, what is the principle for selection of an appropriate constructional 'marker' to mirror the marker used in the source language? The form of the marker in the source language may of course be borrowed along with the construction, but this does not necessarily happen. The borrowing language may select from its own resources a morpheme unrelated to the constructional marker used in the source language. One expects, however, that the marker selected should be semantically appropriate, given the meaning of the construction.

To conclude, we have seen that structurally and semantically parallel constructions can be shared by areally contiguous languages, where the constructional pattern and the morpholexical marker of the construction are logically independent with respect to possible explanations for their sharedness (i.e. as due to borrowing or independent innovation). Further, the use of cognate morpholexical means to mark the same construction in two (even closely) related languages is no guarantee that the construction was not calqued through contact. It is necessary to recognise a type of explanation for shared semantic/grammatical structure in neighbouring languages ('genetically related' or not) aside from the standard options 'shared due to borrowing through contact' and 'shared due to coincidence', as follows. Two languages independently innovate the same new structure, but far from being coincidental, the process is licensed in the respective languages by common semantic and grammatical features which themselves came about as a result of common inheritance or direct diffusion in earlier contact between the languages. Shared semantic and grammatical structure in separate languages provides similar input structures for the processes of inference and extension which drive structural change. The result is parallel yet independent innovation due neither to direct borrowing nor to mere 'coincidence'.

References

- Bisang, Walter. 1991. Verb serialization, grammaticalization and attractor positions in Chinese, Hmong, Vietnamese, Thai and Khmer. In *Participation: das Sprachliche Erfassen von Sachverhalten*, edited by Hansjakob Seiler and Waldfried Premper, 509-562. Tübingen: Gunter Narr Verlag.
- Capell, Arthur. 1979. Further typological studies in Southeast Asian languages. In *South-east Asian linguistic studies, Volume 3*, edited by Nguyen Dang Liem, 1-42. Canberra: PL.
- Chao, Yuen Ren. 1926. Research on Beijing, Suzhou, and Changzhou auxiliaries. *Qinghua Xuebao*, 3.2, 865-917.
- Chao, Yuen Ren. 1968. *A grammar of spoken Chinese*. Berkeley and Los Angeles: University of California Press.

- Chappell, Hilary. (ed.) 2001. *Sinitic Grammar: synchronic and diachronic perspectives*. Oxford: Oxford University Press.
- Clark, Marybeth. 1989. Hmong and areal Southeast Asia. In *Southeast Asian syntax*, edited by David Bradley, 175-230. (A-77) Canberra: PL.
- Enfield, N. J. 2001. On genetic and areal linguistics in mainland Southeast Asia: parallel polyfunctionality of 'acquire'. In *Areal diffusion and genetic inheritance: problems in comparative linguistics*, edited by Alexandra Y. Aikhenvald and R. M. W. Dixon, 255-290.
- Enfield, N. J. In press. *Linguistic epidemiology: semantics and grammar of language contact in mainland Southeast Asia*. London: Routledge.
- Harris, Alice C. and Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.
- Hopper, Paul J. and Elizabeth Closs Traugott. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.
- Huang, C.-T. James. 1988. *Wǒ pào de kuài* and Chinese phrase structure. *Language*, 64.2, 274-311 .
- Lamarre, Christine. 2001. Verb complement constructions in Chinese dialects: types and markers. In *Sinitic grammar: synchronic and diachronic perspectives*, edited by Hilary Chappell. Oxford: OUP.
- Matisoff, James A. 1991. Areal and universal dimensions of grammatization in Lahu. In Traugott, Elizabeth Closs and Bernd Heine, eds. 1991. *Approaches to grammaticalization*. Amsterdam: Benjamins, 383-453.
- Matisoff, James A. 2001. Genetic versus contact relationship: prosodic diffusability in Southeast Asian languages. In *Areal diffusion and genetic inheritance: problems in comparative linguistics*, edited by Alexandra Y. Aikhenvald and R. M. W. Dixon, 291-327.
- Peng, Xiao Chuan. 1993. The structural auxiliary 'dao' in Cantonese. In *Research and pedagogy of Cantonese*, edited by Zheng Ding Ou, 89-100. Guangzhou: Zhongshan Daxue Chubanshe.
- Sun, Chao Fen. 1996. *Word-order change and grammaticalization in the history of Chinese*. Stanford: Stanford University Press.
- Tang, Zhi Dong. 1986. Demonstratives in Xinyi dialect. *Yuyan Yanjiu*, 2, 98-108.
- Thomason, Sarah G. 2000. Linguistic areas and language history. In *Languages in contact*, edited by Dicky Gilbers, John Nerbonne, and Jos Schaeken. Amsterdam: Rodopi, 311-327.
- Traugott, Elizabeth Closs and Richard B. Dasher. 2002. *Regularity in semantic change*. Cambridge: Cambridge University Press.

Language and Cognition Group
Max Planck Institute for Psycholinguistics, Nijmegen
PB 310, 6500 AH, The Netherlands

Nick.Enfield@mpi.nl

PROCEEDINGS OF THE TWENTY-EIGHTH ANNUAL MEETING
OF THE
BERKELEY LINGUISTICS SOCIETY

February 15-18, 2002

SPECIAL SESSION
ON
TIBETO-BURMAN AND SOUTHEAST ASIAN LINGUISTICS

in honor of Prof. James A. Matisoff

edited by:
Patrick Chew
BERKELEY LINGUISTICS SOCIETY