"Split grammaticalization" of 'be' in New Indo-Aryan
Paolo Milizia (Università di Cassino e del Lazio meridionale)

This presentation focuses on some aspects of the grammaticalization of 'be' as an auxiliary within the formation of New Indo-Aryan participle-based verb forms. A typologically non-obvious feature of this diachronic developement lies in the fact that two among the major functions characteristically carried out by auxiliaries, i.e. 1) to act as a TAM marker and 2) to act as a Person-Agreement marker, become autonomous from each other. As we will try to show, a non-trivial property of some New Indo-Aryan periphrastic verb forms, i.e. the fact of exhibiting two instances of Person-Agreement (i.e. one on the lexical verb and another one on the auxiliary) is strictly related to this split.

The participle-based periphrastic constructions relevant to our subject are both those employing the old present participle (i.e. the imperfective participle) and those employing the old verbal adjective in *-ta* (i.e. the perfective participle) – on the typology of grammaticalization paths involving 'be' plus participle, see Heine (1993: 36), Krug (2011).

Crucial facts are: 1) the peculiar kind of grammaticalization undergone by 'be' in a series of varieties, i.e. not simply a grammaticalization of 'be', but the grammaticalization of the opposition between presence and absence of 'be' as a formal device for distinguishing different TAM categories; 2) the fact that the constructions in which the original participles figure without being accompanied by a form of 'be', i.e. the participial predicates with purely nominal syntax, are diachronically reanalysed as "normal" finite verb constructions – by which we do not allude, of course, to the morphological properties typically but not necessarily associated with finite verb forms (since these forms may continue not to inflect for person), but to the purely syntactic/distributional property of being able to stand in a simple declarative sentence.

From a cross-linguistic point of view, we can posit three major patterns according to which a participle-based construction is integrated into the inflectional paradigm of a verbal lexeme: Pattern A: participle with 'be' for all person values; Pattern B (auxiliary-less): participle without 'be' for all person values; Pattern C (mixed): participle with 'be' for 1st and 2nd persons, and participle without 'be' for 3rd persons. Exemplifications are provided by the past tense of the Slavonic languages: Serbo-Croatian has A, Russian has B, Polish and Czech have C (Vaillant 1966: 87-89; cf. also Andersen 1987; Hopper, Traugott 2003: 145-147).

What is remarkable in the development of a series of New Indo-Aryan languages – we will focus in particular on Hindi, Marathi and Gujarati – is that we find more than one pattern in a single verbal system. This is self-evident in Hindi, where we find both the A and the B pattern (cf. Montaut 2004, 2006): Pattern A (full auxiliary): with the present participle → present progressive (subsequently developed into a general imperfective present); with the perfective participle → present perfect; Pattern B (auxiliary-less): with present participle → conditional (counterfactual) and "past habitual"; with the perfective participle → preterite (perfective past).

Significantly, the functions associated with these four types continue uses already documented in Middle Indo-Aryan (including Late Sanskrit) and Early New Indo-Aryan (as for the present participle, cf. Sen [1953] 1965, §§ 150-153; Singh 1980: 150-151; Bubeník 1998: 108-109; Strnad 2013: 415-416; as for the perfective participle, see Bloch 1906: 66-67). In fact, the crucial Modern New Indo-Aryan innovation was to made the presence of the auxiliary either categorically required or categorically excluded according to the different functions (or function bundles). Thus in the Hindi system (see table 1 below) the present indicative ~ conditional opposition and the present perfect ~ preterite opposition are expounded by the presence vs. absence of the auxiliary (and by the related selection of the nasalized feminine plural ending in auxiliary-less contructions). Though the auxiliary contained in these constructions does inflect for person, person-inflection is not a real requirement for Hindi finite verbs, as is shown by preterite and conditional, which are, synchronically, finite verb forms.

An interesting development is found in Marathi, a language with a dispreference for finite verbs non-inflecting for person. Marathi shows a univerbized pattern C (i.e. a resynthetized paradigm where non-third persons have person agreement endings which are the outcomes of the cliticization and univerbation of inflected forms of the auxiliary 'be' – cf. Bloch 1920: 246) exactly for the same functions for which Hindi has pattern B: conditional and preterite. These univerbized occurrences of 'be', which serve as person markers, do not prevent the simultaneous presence of autonomous non-univerbized forms of 'be', which serve as TAM markers, in the same periphrastic construction. This development constitutes a "grammaticalization split" in the sense that the inflected forms of 'be' underwent a split whereby they were grammaticalized as Person Agreement markers along one path and as TAM markers along another. Thus in Marathi, like in Hindi, the present perfect is identical to the simple preterite except for the additional presence of the auxiliary, but unlike in Hindi, it shows two occurrences of person agreement: one in the form of the lexical verbs (which is a participle with an univerbized auxiliary) and another one in the unbound auxiliary form. In fact this constructions represents a sort of C + A double pattern (see, e.g, the 2nd plural forms in tables 2 and 3 below; cf. Navalkar 1880: 158; Katenina 1983: 223, 235-238).

As for the other periphrases, the Marathi situation is complicated by the fact that this language manages to form three different verbal categories based on the gender-distinguishing imperfective participle. Indeed, in the first place, it employs the opposition between two different sets of participial endings (the first, which is the regular outcome of the *-aka-/ikā*-enlarged participle, with masc. sing. *-ā-* ~ fem. sing. *-ī-*; the second, which is probably taken from the 3rd person pronoun *to*, with masc. sing. *-o-* ~ fem. sing. *-e-* – cf. Bloch 1920: 247-248) to distinguish the conditional (in *-ā-* ~ *-ī-*) from the general imperfective present (in *-e-* ~ *-o-*); in the second place, it employs the presence of the auxiliary in order to formally characterize the "emphatic" present progressive, which, like the present perfect, exhibits both univerbized and unbound forms of "be" and is, therefore, a double finite verb (tables 4, 5, 6; cf. Navalkar 1880: 157, 160; Katenina 1963: 222-223, 229-230).

As we will try to show, a hypothesis can be proposed about the origin of the type of double finite present shown by Gujarati (but also by other varieties of a wide western South-Asian area, including the Khari Boli dialect; cf. LSI 9/1: 65, 73, 81, 255; LSI 9/2: 26, 41, 48, 57; LSI 9/3: 13; Tessitori 1915: 74; Varma 1935: 119). Such a construction consists of the combination of the old present with the 'be' auxiliary (table 7). Indeed it can be thought that this kind of double finite verb owes its origin to the fact that in these varieties the old present also got sucked into the verbal subsystem marked by formal oppositions based on the presence vs. absence of 'be'. Sanskrit uses of *asti* with meanings like "once upon a time" or "it is the case that" have been cautiously and tentatively invoked by Hock (2014) as possible precursors of the formation at issue. However, if our hypothesis holds, such an attempt at explanation will turn out to be unnecessary.

Tables:

1)	Hindi: inflection of *calnā* 'go, walk'				
	present		conditional		
	m.	f.	m.	f.	
1sg.	*caltā hũ̃*	*caltī hũ̃*	*caltā*	*caltī*	
2sg.	*caltā hai*	*caltī hai*	*caltā*	*caltī*	
3sg.	*caltā hai*	*caltī hai*	*caltā*	*caltī*	
1pl.	*calte hãĩ*	*caltī hãĩ*	*calte*	*caltĩ̃*	
2pl.	*calte ho*	*caltī ho*	*calte*	*caltĩ̃*	
3pl.	*calte hãĩ*	*caltī hãĩ*	*calte*	*caltĩ̃*	
	present perfect		preterite		
1sg.	*calā hũ̃*	*calī hũ̃*	*calā*	*calī*	
…	…	…	…	…	…

2)	Marathi: perfect of 'walk'		
	m.	f.	n.
1sg	*cālălõ āhẽ*	*cālălẽ āhẽ*	*cālălẽ āhẽ*
2sg	*cālălā āhes*	*cālălī āhes*	*cālălẽ āhes*
3sg	*cālălā āhe*	*cālălī āhe*	*cālălẽ āhe*
1pl	*cālălõ āhõ*		
2pl	*cālălã̃ āhã̃*		
3pl	*cālăle āhet*	*cālălyā āhet*	*cālălĩ̃ āhet*

3)	Marathi: preterite of 'walk'		
	m.	f.	n.
1sg	*cālălõ*	*cālălẽ*	*cālălẽ*
2sg	*cālălās*	*cālălīs*	*cālălẽs*
3sg	*cālălā*	*cālălī*	*cālălẽ*
1pl	*cālălõ*		
2pl	*cālălã̃*		
3pl	*cālăle*	*cālălyā*	*cālălĩ̃*

4)	Marathi: conditional of 'walk'		
	m.	f.	n.
1sg	*cāltõ*	*cāltẽ*	*cāltẽ*
2sg	*cāltās*	*cāltīs*	*cāltẽs*
3sg	*cāltā*	*cāltī*	*cāltẽ*
1pl	*cāltõ*		
2pl	*cāltã̃*		
3pl	*cālte*	*cāltyā*	*cāltĩ̃*

5)	Marathi: present of 'walk'		
	m.	f.	n.
1sg	*cāltõ*	*cāltẽ*	*cāltẽ*
2sg	*cāltos*	*cāltes*	*cāltẽs*
3sg	*cālto*	*cālte*	*cāltẽ*
1pl	*cāltõ*		
2pl	*cāltã̃*		
3pl	*cāltāt*		

6)	Marathi: emphatic present progressive of 'walk'		
	m.	f.	n.
1sg	*cāltõ āhẽ*	*cāltẽ āhẽ*	*cāltẽ āhẽ*
2sg	*cāltos āhes*	*cāltes āhes*	*cāltẽs āhes*
3sg	*cālto āhe*	*cālte āhe*	*cāltẽ āhe*
1pl	*cāltõ āhõ*		
2pl	*cāltã̃ āhã̃*		
3pl	*cālte āhet*	*cāltyā āhet*	*cāltĩ̃ āhet*

7)	Gujarati: present of 'go'	Braj (Bulandshahar): present of 'go'
1sg	*avũ chũ*	*calũ hũ*
2sg	*ave che*	*calae hae*
3sg	*ave che*	*calae hae*
1pl	*avie chie*	*calaẽ haẽ*
2pl	*avo cho*	*calao hao*
3pl	*ave che*	*calaẽ haẽ*

References:

Andersen, H. 1987. From auxiliary to desinence. In M. B. Harris, P. Ramat, *Historical Development of Auxiliaries*. Berlin: Mouton de Gruyter: 21-51.

Bloch, J. 1906. *La phrase nominale en sanskrit*. Paris: Champion.

Bloch, J. 1920. *La formation de la langue marathe*. Paris: Champion.

Bubeník, V. 1998. *A Historical Syntax of Late Middle Indo-Aryan (Apabhraṃśa)*. Amsterdam: Benjamins.

Heine, B. 1993. *Auxiliaries. Cognitive Forces and Grammaticalization*. Oxford: OUP

Hock, H. H. 2014. A morphosyntactic chain shift in the Hindi-Panjabi area: Explications and implications. *Journal of South Asian Languages and Linguistics* 1(1): 5-29.

Hopper, P., Traugott, E. C. 2003. *Grammaticalization*. 2nd ed. Cambridge: CUP.

Katenina, T. E. 1963. *Očerk grammatiki jazyka maratxi*. Moskva: Izd. Literatury na inostrannyx jazykax.

Krug, M. 2011. Auxiliaries and Grammaticalization. In B. Heine, H. Narrog (eds.), *The Oxford Handbook of Grammaticalization*. Oxford: OUP: 457-558.

LSI = Grierson, G. A. (ed.) 1903-1928. *Linguistic Survey of India* (11 vols.) Calcutta: Government of India.

Montaut, A. 2004. *A Grammar of Hindi*. Muenchen: Lincom.

Montaut, A. 2006, The evolution of the tense-aspect system of Hindi/Urdu: The status of the ergative alignment. In M. Butt, T. Holloway King (eds.), *Proceedings of the LFG06 Conference*. Stanford: CSLI: 365-385.

Navalkar, G. 1880. *The Student's Marathi Grammar*. Bombay: Education Society.

Sen, S. [1953] 1965. Historical Syntax of Middle Indo-Aryan, *Indian Linguistics (Repr. Edition)* 3: 355-473 (1st ed. *Indian Linguistics* 13: 1-140).

Singh, R. A. 1980. *Syntax of Apabhramsa*. Calcutta: Simant.

Strnad, J. 2013. *Morphology and Syntax of Old Hindī*. Leiden: Brill.

Tessitori, L. P., 1914, 1915, 1916. *Notes on the grammar of the old western Rajasthani*. Indian Antiquary 43, 44, 45.

Vaillant, A. 1966. *Grammaire comparée des langues slaves*. Tome III: *Le verbe*. Paris: Klincksieck.

Varma, Dh. 1935. *La langue braj (Dialecte de Mathurā)*. Paris: Adrien-Maisonneuve.