Peter Arkadiev

The Berber “state” distinction: Dependent marking after all? A commentary on Mettouchi & Frajzyngier (2013)

Abstract: Critically reviewing the recent analysis of the distinction between the so-called “absolute” and “annexed” “states” in Kabyle (Berber) by Mettouchi & Frajzyngier (2013) (M&F), it is shown that it does not fare better than the existing treatments in its account of the empirical facts, mainly due to the assumption of monosemy. It is further argued that M&F’s rejection of “case” as a valid notion for the description of the Kabyle data rests on simplistic and ill-informed views on the nature of case, and that adoption of a case analysis allows one to compare the Berber data to a wide range of languages with similar peculiarities in the distribution of dependent marking. Finally, M&F’s claim that their analysis of Kabyle has pointed out a “previously unrecognized typological category” is refuted, showing that it stems from an unwarranted mixing of language-particular descriptive categories and crosslinguistic comparative concepts.

Keywords: annexation, Berber, case, grammatical relations, head-/dependent-marking, inflection, Kabyle, noun phrase, state, syntax

1 Introduction

In an interesting and thought-provoking article in this journal, Mettouchi & Frajzyngier (2013), its authors (henceforth M&F) presented a novel and ingenious analysis of the so-called “state” distinction in Kabyle Berber. The “state” in Berber languages is a morphological inflectional category of nominals with two values, traditionally called “absolute” and “annexed”. The distribution of “states” is far from trivial, not being reducible to any single dimension such as grammatical relations (subject vs. object), information structure (topic vs. focus),
word order, or syntactic embedding. M&F analyze the distribution of the two “states” in Kabyle, and propose that the function of the “annexed state” is to “provide the specific value for a grammaticalized meaning encoded earlier in the sentence” (Mettouchi & Frajzyngier 2013: 1). M&F also argue against the previous attempts to account for the distribution of “states” in terms of the notion of case, and claim that “state” is a “previously unrecognized typological category” (this is, in fact, the title of their article).

The aim of this contribution is to critically review the arguments put forward by M&F against the treatment of the “state” category of Kabyle as a particular instantiation of case, and to ultimately cast doubt on the validity of the category “provid[ing] the specific value for a grammaticalized meaning encoded earlier in the sentence” for linguistic typology. In particular, after briefly outlining the essence of M&F’s proposal in Section 2, I will try to show, first, that their analysis of Kabyle has important shortcomings (Section 3), then, second, that most of the arguments against the case analysis of “state” adduced by M&F stem from a simplistic and typologically not very well informed preconception of case, and that treating the Berber “states” as cases is legitimate from a typological perspective (Section 4), and, third, that even if the functions M&F attribute to the Kabyle “states” are indeed best suitable to describe their distribution, these functions cannot be reasonably generalized crosslinguistically (Section 5). Therefore, I conclude that M&F’s article represents a common fallacy whereby linguists fail to keep apart descriptive categories used in the analyses of particular languages and comparative concepts useful for crosslinguistic generalizations (Haspelmath 2010).

Before proceeding I must stress that presenting an alternative and improved analysis of the Kabyle facts is not among the goals of this contribution. Being not a specialist on Berber languages, I am not in a position to propose such an analysis. Moreover, I do not consider such a task really relevant in the context of this discussion, since the aims of LT, as I understand them, primarily concern advancement of crosslinguistic studies and scrutiny of typologically valid notions rather than language-specific analyses. However, in Section 4 I will nevertheless try to show that an analysis of Kabyle in terms of case is not only possible but also typologically adequate.¹

¹ I would like to thank one of the anonymous reviewers and Frans Plank for convincing me that such a point should indeed be elaborated upon in this contribution.
2 M&F’s analysis in a nutshell

The starting point of M&F’s analysis of the “state” distinction in Kabyle is the observation that the distribution of the “annexed” state is constrained by the occurrence of certain grammatical elements to the left of the relevant nominal in the same clause. For instance, the “annexed” state is used to mark the subject of the verb cross-referenced on the latter by a special set of obligatory subject markers, as well as the complement of certain prepositions and relational elements. By contrast, the “absolute” state is free from such restrictions; thus, it marks the grammatical object when it is not cross-referenced on the verb, all preverbal (topicalized) nominals, including the subject, nominal attribute, and predicate nominal. Crucially, in cases when the grammatical object is cross-referenced by the special set of object clitics and functions as a kind of “anti-topic”, separated from the preceding clause by a pause (see, e.g., Mettouchi 2005, 2006, 2008), it is marked by the “annexed state” (it must be noted that the connection between “state”, bound pronominal cross-referencing, and information structure in Berber has been observed already in Basset (1950), and Galand (1964, 1979)). Illustrative examples of all the mentioned constructions can be found in Mettouchi & Frajzyngier (2013: 5–7).

Based on the above-mentioned facts, M&F propose that, first, the “absolute state” in Kabyle “is the default form of the noun and does not carry any specific function” (Mettouchi & Frajzyngier 2013: 12) and, second, that the “annexed state provides the value (in the logical sense) for the variable of the function grammaticalized in a preceding constituent” (Mettouchi & Frajzyngier 2013: 13). M&F represent their hypothesis about the function of the “annexed state” as \( f(x) = N_{\text{ann}} \), where \( f \) is the function, \( x \) is the variable, and \( N_{\text{ann}} \) is the noun in the annexed state” (Mettouchi & Frajzyngier 2013: 13). The “grammaticalized function” may be expressed by closed-class (grammatical) elements such as bound pronominals (expressing the functions of “subject” and “object”) and prepositions. When the nominal cannot be construed as “providing the value” for such a grammaticalized function, either because it precedes rather than follows the constituent bearing the relevant cross-referencing marker (as in the case of topicalized preverbal noun phrases), or because such a marker does not exist at all (as in the case of non-cross-referenced objects, adverbials, and nominal attributes), the nominal appears in the “absolute state”.

The analysis which I have succinctly and crudely presented here indeed provides a uniform and straightforward account of most of the uses of the two “states” in Kabyle and strikes one as both ingenious and appealing. However, the devil is in the details, and there are some aspects to M&F’s analysis which in fact may cast doubt on its validity for the description of the Kabyle facts.
3 Problems with M&F’s analysis

One of the puzzling aspects of the distribution of the “absolute” and “annexed” states in Kabyle is the fact that with prepositions, both “states” can occur. (For the early systematic and theoretically informed study of this phenomenon see Guerssel (1987, 1992), revised in Achab (2003), see also Bendjaballah & Haiden (2005) for an attempt at a phonological account of this distribution, which M&F seem to ignore.) Thus, such prepositions as g ‘at, in’, d ‘with’, am ‘like’, and some others require their complement to appear in the “annexed state”, while such prepositions as siwa ‘except’, ar ‘except, until’, and bla ‘without’, by contrast, take the complement in the “absolute state” (see Mettouchi & Frajzyngier 2013: 22–23). Finally, at least one preposition, s, takes the complement in the “annexed state” when used in the instrumental meaning, and in the “absolute state” when used in the directional meaning (Mettouchi & Frajzyngier 2013: 24). M&F account for the distribution of “states” with prepositions in Kabyle in the following way.

First, prepositions, being a closed class of elements, themselves serve as expressions of grammaticalized functions, so that their nominal complements provide the values for these functions, thereby, as predicted by M&F’s theory, appearing in the “annexed state”. Why then do some of the prepositions require or allow their complement to appear in the “absolute state”? M&F answer this question as follows. The grammaticalized meaning of the prepositions siwa ‘except’, ar ‘except, until’, and bla ‘without’ is “negative”, and “[t]he variables [...] cannot be given a value” (Mettouchi & Frajzyngier 2013: 22), and therefore their complement appears in the “absolute state”. M&F represent the contrast between the two kinds of preposition as follows: d ‘with’ f(associative) vs. bla ‘without’ f(∼associative).

Such an account strikes me as, first, stipulative, and, second, unconvincing. Nowhere in the article do M&F actually explain why prepositions with a “negative” meaning “cannot be given a value”, or, to put it differently, why their complement, which is arguably required for the sentence to be grammatically felicitous and semantically well-formed, cannot nevertheless provide the necessary value. In my view of semantics, both in with the dog and in without the dog the noun phrase the dog provides the value for the grammatical function, or, to use a different terminology, serves as an argument of the relator, expressed by the preposition. Both with and without select for an expression denoting the entity accompanying and, respectively, not accompanying the situation, and I see no reason (and M&F fail to provide one) for treating these prepositions as inherently different in their semantic or functional requirements. The same
applies to the other two Kabyle prepositions taking the complement in the “absolute state”. This is puzzling also since in fact not all Kabyle elements with a putatively “negative” meaning are treated by M&F in this way. For instance, the negative existential *ulaf*, after which the noun appears in the “annexed state” (Mettouchi & Frajzyngier 2013: 16–17), is analyzed as $f$(absent argument), not as $f$(~existing argument). It is not clear what the relevant semantic difference between “except that woman” and “the child is not here” can be which would necessitate two different “function-argument” analyses of these Kabyle elements, and so M&F’s account appears circular.2

Even more problematic is the case of the preposition *s*, which can co-occur with both “states”. M&F’s reasoning concerning this preposition deserves to be quoted in full (Mettouchi & Frajzyngier 2013: 24):

> The preposition *s* has an instrumental meaning. The complement in the annexed state provides the value for the grammaticalized function $f$(instrumental) of the preposition […]. This preposition can also be followed by a noun in the absolute state. The complement noun does not provide the value for the instrumental function coded in the preposition. Hence the fact that it is not in the annexed state, but in the only other possible form for the noun, the absolute. The absolute state of the noun after this preposition directs the hearer to interpret the complement as a directional adverb of manner, not as an instrumental complement.

I consider this reasoning non-explanatory and descriptively inadequate. Indeed, M&F not only fail to explain why the “absolute state” of the complement with the preposition *s* is interpreted as directional (*s + house.ABS means ‘home’*) and not in any other conceivable way, especially in some “negative” fashion, which would be expected given the nature of the other prepositions taking the “absolute state”; M&F also fail to explain why at all the “absolute state” of the noun can co-occur with this preposition and yield a grammatically correct construction, and why this happens only with this particular preposition, and not just in Kabyle, but in other Berber languages as well, e.g., in Tamashek (Heath 2005: 276–277). It is important to underline that M&F’s reasoning does not by any means preclude other prepositions to combine with nouns in the “absolute state” yielding any kind of meaning different from the one combining with the “annexed state”. Nothing in M&F’s theory of the Kabyle “states” and their relationships with prepositions predicts that the following “pseudo-Kabyle” preposition would be impossible: *n ‘of’ + “annexed state”, n ‘across’ + “absolute state”.

---

2 I thank an anonymous reviewer for pointing out this issue.
Moreover, the very idea that nouns in the “annexed state” can appear after (some) prepositions because the meaning of the latter “contains a variable that can be given a value” is not unproblematic, either. The main reason for such an account adduced by M&F is that “prepositions belong to a closed set” and “[t]he function of each preposition is determined by its contrast with other members of the set” (Mettouchi & Frajzyngier 2013: 21). This sounds fine in M&F’s framework, but one wonders why such an analysis cannot be applied to some other elements, for instance, to the copula d. The copula in Kabyle requires the predicate nominal to appear in the “absolute state” (Mettouchi & Frajzyngier 2013: 19–20), and M&F motivate this saying that “the copula itself does not bear any of the semantic functions grammaticalized in Kabyle, for example, it has no gender or number marking”. However, the same applies to prepositions, which bear no gender or number marking, either, and the copula can be considered entering into and defined by such grammatical contrasts as “copular vs. existential vs. verbal predication”. What, besides the need to account for the “absolute state” of the predicate nominal, precludes analyzing the copula as containing a variable, e.g., f(property)?

I conclude that M&F have failed to adequately account for the distribution of “states” with prepositions (and, admittedly, in some other contexts) in Kabyle, and that this failure stems from the very theoretical foundations of their analysis, i.e., their definition of “grammaticalized function” and its “value”, and especially from the – at least apparent – lack of principled constraints on the application of such notions in linguistic descriptions. It must be stressed in this respect that functional explanations of the distribution of grammatical categories, such as the one proposed by M&F, are subject to stricter adequacy requirements than purely descriptive statements in such terms as, e.g., case, which M&F dismiss as inadequate. (I will turn to this issue in the next section.) A functional explanation should rest on unequivocal and explicit assumptions about the relation between the explanans (functional principles) and explanandum (linguistic distribution) and should be applied consistently, notably not containing loopholes allowing the researcher to “adjust” the functional principles in such a way that they match the observed linguistic facts. Moreover, a functional explanation need not necessarily and in fact should preferably not be a “monosemy” explanation, positing a one-to-one relation between linguistic form and linguistic meaning or function, and, as has been convincingly shown by, e.g., Dahl (1987) and Croft (2001: 112–119), monosemy analyses of linguistic categories, language-specific and especially typological, actually often fail to provide satisfactory and non-ad hoc accounts of their distribution, almost

---

3 I thank an anonymous reviewer for raising this issue.
inevitably falling into some sort of circular reasoning of the kind “the clearly identifiable function F in context C1 is expressed by form A, and therefore form A in a different context C2 also corresponds to function F”. A widely accepted alternative to monosemy accounts is provided by polysemy analyses, whereby different uses of a linguistic element are not just listed but organized in a network reflecting possible links between them (see, e.g., Haspelmath 2003, Evans & Green 2006: 331–355). Such semantic networks, especially when based on crosslinguistic and historical data, allow for both principled synchronic accounts and construction of diachronic scenarios. By contrast, it is hardly conceivable how a monosemy analysis proposed by M&F could be useful for an understanding of the historical development of the Berber “state” category, even if it were free of the descriptive weaknesses I have pointed out above.

4 Why (not) case?

The “state” distinction of Berber has been analysed as dependent marking or nominal case ever since Berber data started to be treated in theoretical and typological literature; see, e.g., Bader & Kenstowicz (1987), Guerssel (1992) for influential generative analyses and Sasse (1984), Aikhenvald (1990, 1995), König (2006, 2008: 180–187), Creissels (2009), and Arkadiev (2009b) for functional-typological approaches. M&F specifically target this kind of analysis of the category of “state”, claiming that treating it as case is not adequate. The main arguments M&F provide against the case analysis of “state” in Berber all hinge on the (implicit) assumption that, first, the function of case is to mark grammatical relations and, second, case should mark these grammatical relations in a straightforward way. For instance, M&F say that “the coding of the grammatical relations subject and object is not the function of the two states” (Mettouchi & Frajzyngier 2013: 11), and later on that “some subjects and some objects bear the annexed state and other subjects and objects bear the absolute state” (Mettouchi & Frajzyngier 2013: 12), and therefore that “both the claim of marked nominative and of dependent marking fail to provide an explanation for why nouns are marked by the absolute state in some constructions and by the annexed state in others” (Mettouchi & Frajzyngier 2013: 12). More particularly, on the basis of examples like their (21), where both the subject and the (antitopic) object are

4 I thank Frans Plank and an anonymous reviewer for suggesting to raise this issue.
marked by the “annexed state”, M&F claim that it is illegitimate to characterize the “states” as cases (Mettouchi & Frajzyngier 2013: 15):

The two examples of nouns in the annexed state, where one is interpreted as the subject and the other is interpreted as the object, contradict Aikhenvald’s (1995) and König’s (2006, 2008) claim that the annexed state marks the nominative case and the absolute state marks the accusative case.

Though one has to agree with M&F that the term “case” by itself cannot explain the use and distribution of the two “states” in Kabyle, I do not find their argumentation against treating this distinction as case convincing and well-motivated. First of all, one has to distinguish between labeling a linguistic phenomenon and explaining its distribution. To take a simple example from a different domain, it is well known that present and past tenses in many languages are often used in ways which are not straightforwardly connected to the idea of “event occurring simultaneously to the time of speech” and “event occurring before the time of speech”, such as, e.g., the historical use of the present tense to refer to past events or irrealis use of the past tense (as in If I were you, I wouldn’t say that). Would M&F argue that what linguists working on English have been calling “tense” is in fact a “previously unrecognized typological category”? Labels such as “tense” or “case” are not explanatory by themselves, they only reflect the linguists’ intuition concerning the nature of the phenomena described and the expectations linguists (implicitly or explicitly) have about the behaviour of linguistic categories – intuitions based on theoretical assumptions and crosslinguistic knowledge. Thus, in calling a language-particular category “tense”, a responsible linguist assumes that a substantial portion of its distribution can be systematically accounted for by the notion of temporal reference, and that the rest of this distribution can be linked to this notion in some principled way (which, of course, often remains a desideratum, especially in the case of underdescribed languages). However, calling a language-particular phenomenon “tense” does not imply that all its uses would be straightforwardly and unequivocally related to temporal reference in some evident and trivial way – and linguistic descriptions trying to achieve this goal of using monosemous labels as explanatory devices in order to account for clearly polysemous categories usually end up in some or other kind of descriptive and explanatory reductionism (see again Croft 2001: 112–119).

The same reasoning applies to case. M&F argue against calling the Kabyle “states” “nominative” (“annexed state”) and “accusative” (“absolute state”) as if these labels themselves would serve as explanations (i.e., predictors) of the distribution of the relevant grammatical forms. The reason why, for instance, König (2008: 180–182) calls the “annexed state” nominative and the “absolute
case” accusative is that there are (arguably morphosyntactically basic) constructions in Kabyle where the “annexed state” (“nominative”) marks the subject and the “absolute state” (“accusative”) marks the direct object, cf. examples (16a) and (16b) in Mettouchi & Frajzyngier (2013: 12). However, such labeling does not imply that this is all what can or should be said about the distribution of the relevant morphological categories in Kabyle, neither does it aim at explaining all the ranges of use of these categories. Nevertheless, classifying Kabyle and other Berber languages as possessing case systems allows one to make both typological insights – see König (2006, 2008, 2009) on the typology of marked-nominative systems and the place of Berber among them,5 and Arkadiev (2005, 2009b) on the place of Berber in the typology of two-term case systems – and descriptive observations regarding the polysemy of the relevant categories and their possible historical evolution, whereas M&F’s monosemous proposal, as I have argued in Section 3, is not descriptively adequate and predictive on purely language-internal grounds, and, as will be shown below and especially in Section 5, is of no much use for typological research, either.

Furthermore, I would like to argue that M&F’s reasoning exemplified above is misguided from a crosslinguistic perspective on case as well. In modern typological work, case is understood as morphological means of marking dependent nouns for their relationship to their heads (see Blake 2001: 1), or, more generally, to the constructions in which they occur (cf. Bickel & Nichols 2007: 197). Though “prototypically” (or, to take a different perspective, “canonically”, see Corbett 2008), one of the major crosslinguistic functions of case is marking grammatical relations such as subject or object (or semantic roles such as agent, patient, etc.), it is well known that the mapping between morphosyntactic case values or case markers, on the one hand, and syntactic or semantic roles, on the other hand, is far from straightforward and trivial, being subject to language-particular syntactic, semantic, or pragmatic conditions (though both this mapping and these conditions lend themselves to reasonable crosslinguistic generalizations). Linguists working on case languages are pretty well used to situations when the same grammatical relation or semantic role is marked by different cases depending on certain properties of the construction, and to situations when one and the same case marks different grammatical relations or semantic roles – either paradigmatically or syntagmatically.

Consider, e.g., Lithuanian (Baltic, Indo-European; for more details, see Arkadiev 2013), where the grammatical relation “subject” can be marked by the nominative case in ordinary finite clauses (1a), by the accusative case in

5 Unfortunately, Berber languages are not discussed in the otherwise quite comprehensive typological study of marked nominative languages by Handschuh (2014).
non-finite complement clauses with verbs of perception or speech (1b), and by
the dative case in non-finite adverbial clauses (1c). Note that in (1b) and (1c),
there are two different grammatical relations – subject and direct object in (1b),
subject and indirect object in (1c) – marked by the same case.

(1)  a. Violet-a paraš-ė t-q skund-ą.
    Violeta-NOM.SG write-PST.3 that-ACC.SG complaint-ACC.SG
    ‘Violeta wrote that complaint.’ (constructed)

    b. [...] direktor-ė įtar-ė [Violet-q paraš-ius
director(F)-NOM.SG suspect-PST.3 Violeta-ACC.SG write-PST.PA
    t-q skund-ą].
    that-ACC.SG complaint-ACC.SG
    ‘the [woman] director suspected that Violeta had written that complaint’

(2)  a. Smėl-yje buv-o [maty-ti pėdsak-ai].
    sand-LOC.SG be-PST.3 see-INF footprint-NOM.PL
    ‘Footprints could be seen on the sand.’

    b. [...] j-ie ne-skait-o lietuvi-ų spaud-os.
    3-NOM.PL.M NEG-read-PRS.3 Lithuanian-GEN.PL press-GEN.SG
    ‘they don’t read the Lithuanian press.’

    c. darbuotoj-ai tur-i daugią laik-o [aptarnau-ti
    worker-NOM.PL have-PRS.3 more time-GEN.SG attend-INF
    klient-ams]
    client-DAT.PL
    ‘... workers have more time to attend to clients.’
Though the exact conditions on the case marking of grammatical relations in Lithuanian are very different from those in Kabyle (just as the inflectional systems themselves are different), the phenomena we are dealing with in the two languages belong essentially to one and the same type, i.e., constructionally conditioned differential dependent marking of arguments (see, e.g., de Hoop & de Swart (eds.) 2009, de Hoop & Malchukov 2008, Malchukov & de Swart 2009, Malchukov & de Hoop 2011).

Situations when morphological cases fail to distinguish between grammatical relations such as subject or object are especially well-attested in languages with two-term case systems (see Arkadiev 2009a, b), and in this respect Kabyle is not at all special. Consider, for instance, the following example from Western Balochi (Iranian, Indo-Iranian, Indo-European).

(3) \textit{bačak-\textit{k-a watı dant-\textit{ān-ā prōšt.}}}
\textit{boy-OBL.SG his tooth-PL-OBL.PL broke}
‘The boy broke his (own) teeth.’ (Korn 2009: 53)

Again, conditions triggering the so-called “double-oblique” construction in various Iranian and Indo-Aryan languages (see, e.g., Arkadiev 2009a, Haig 2008, Stilo 2009; generally speaking, these involve past tense or perfective aspect of the verb and animacy and/or definiteness of the direct object) are very different from the conditions under which both subject and object get the “annexed state” in Berber (non-topicalized subject and an anti-topic object cross-referenced by object pronominials on the verb), but both kinds of phenomena naturally fall under the general notion of differential dependent marking.

In this connection it must be noted that the fact that the “state” distinction in Kabyle applies only to postverbal nominals, all preverbal nominals invariably appearing in the “absolute state” regardless of their semantic role or grammatical relation, is an instance of the “no case before the verb” constraint, which is well attested in various African languages (see König 2008: 240–273), including those whose dependent marking systems are very different from that of the Berber languages. Consider, for instance, Coptic (Egyptian, Afro-Asiatic; Grossman 2015), where both the subject and the object are overtly case-marked by prefixes in postverbal position only (4a), losing this case marking when moved before the verb (4b,c).

(4) (a) \textit{a-f-t\textit{h}amio nće-p\textit{h}[nou]t\textit{č} m-pi-tačro}
\textit{PST-3SG,M-create NOM-G[0]d ACC-DEF,M-firmament}
‘God created the firmament.’ (Early Bohairic, Genesis 1:7; Grossman 2015: 208)
b. **Adam**=de  a-f-t⁴-ran  e-ni-tebnôou
   Adam=CONN  PST-3SG.M-give-name  ALL-DEF.PL-animals
   ‘And Adam named the animals.’ (Early Bohairic, Genesis 2:20; Grossman 2015: 209)

c. **pi-kʰaki**=de  a-f-t⁴-ren-f  če-pi-ečôrh
   DEF.M-darkness=CONN  PST-3SG.M-give-name-3SG.M  QUOT-DEF.M-evening
   ‘The darkness, he named it “evening”.’ (Early Bohairic, Genesis 1:5; Grossman 2015: 209)

A similar constraint against overt case marking in preverbal position is also attested in the Western Malayo-Polynesian language Nias (Brown 2001: 361) and in the Oceanic language Vaeakau-Taumako (Næss & Hovdhaugen 2011: 312–316), as well as in the Mon-Khmer language Semelai (Kruspe 2004: 251–266). Like Coptic, Semelai (Aslian, Mon-Khmer, Austro-Asiatic; Malaysia) has both overt marking of transitive agents and patients, see (5a), which is suspended when any of these constituents is fronted, see (5b, c), and the same even applies to locative-marked recipients, see (6a, b).

(5) a. **ki**=bukə?  **la**=knlək  **hn**=pintu?
   3.A=open  ERG=husband  ACC=door
   ‘The husband opened the door.’ (Kruspe 2004: 255)

b. **tilam**  **ki**=goŋ  **he?**  **raboŋ**
   mattress  3.A=bring above roof.ridge
   ‘The mattress, he brought up onto the ridge of the roof.’ (Kruspe 2004: 251)

c. **kah**,  **ki**=ʔye  **creh**
   he  3.A=see fish
   ‘Him, he saw the fish.’ (Kruspe 2004: 257)

(6) a. **ki**=ʔur  **ʔen**  **kmpan**
   3.A=instruct  LOC  wife
   ‘He instructed his wife.’ (Kruspe 2004: 257)

b. **kmpan**,  **ki**=ʔur
   wife  3.A=instruct
   ‘His wife, he instructed her.’ (Kruspe 2004: 257)

Thus, in Coptic and in Semelai, in distinction to Berber and many of the marked-nominative languages of Africa, the unmarked form occurring in preverbal fronting does not coincide with any of the forms regularly used for marking arguments in the postverbal position (though Kruspe (2004: 260–262) notes that both the agent and the object case markers are often omitted even in postverbal
position, as in (5c)), but the general mechanism of suspension of overt case marking in topicalization is evidently the same in all these languages.

Moreover, it is by no means obvious that it is legitimate to speak about the (even formally covert) distinction between different grammatical relations in the preverbal position in Kabyle and other languages with the “no case before the verb” rule at all. For example, Mettouchi (2005) presents convincing arguments that the preverbal position in Kabyle is that of the topic. In some syntactic frameworks (e.g., Lexical-Functional Grammar, see Bresnan & Mchombo 1987) the latter is treated as a grammatical function in its own right, different from subject or object and thus admitting its own (often though by no means always zero) morphological marking – in case of Kabyle, the “absolute state” (cf. “hanging topics” in the unmarked nominative case in languages such as German, see, e.g., Nolda 2004).

The issue of the choice of “annexed” or “absolute” “state” with prepositions in Kabyle (see Section 3 above) is also not a problem for the analysis in terms of case. It is well known (see, e.g., Lestrade 2006 for a typological survey) that in many languages different adpositions require different cases on their nominal complements, and that in some languages the same adposition may combine with complements in different cases depending on the particular meaning of the adposition. Consider once again Lithuanian, where the directional adposition į ‘into’ requires the accusative case (7a), and the ablative preposition iš ‘from, out of’ requires the genitive case (7b):

(7) a. ei-ti į kambar-į
go-INF in room-ACC.SG
‘to go into the room’

b. ei-ti iš kambari-o
go-INF from room-GEN.SG
‘to go out of the room’

Next consider Russian (Slavic, Indo-European), where the preposition s(o) with a complement in the instrumental case denotes accompaniment (8a), and with a complement in the genitive case denotes a source of motion (superrelative) (8b).

(8) a. id-ti so stul-om
go-INF with chair-INS.SG
‘to walk with a chair’ (e.g., holding it in one’s arms)

b. vsta-t’ so stul-a
stand.up-INF from chair-GEN.SG
‘to stand up from the chair’
One might dismiss the examples from Lithuanian and Russian on the basis of these being languages with large case systems where several morphologically marked cases can appear with prepositions. However, phenomena of this kind are attested in languages with just two cases one of which is formally unmarked as well, cf. the following examples from Nias (Western Malayo-Polynesian, Austronesian; Sumatra), where in (9a) the locative preposition \( \textit{ba} \) is followed by a noun in the so-called “mutated form” which can be termed “oblique case”, whereas in (9b) the instrumental preposition \( \textit{faoma} \) requires its complement to occur in the “unmutated” form (“absolutive case”).

(9) a. so \( \textit{ndro} \) \( \textit{ba} \) \( \textit{mbafa-nia} \).
   exist \( \text{blood.OBL} \) \on face.OBL-3SG.POSS
   \'He has blood on his face.' (Brown 2001: 351)

b. \( \textit{u-tab} \) \( \textit{na} \) \( \textit{gole} \) \( \textit{faoma} \) \( \textit{balatu} \).
   \( 1SG.RLS\)-cut.up \( \text{meat.OBL} \) \with \( \text{knife(ABS)} \)
   \'I cut the meat with a knife.' (Brown 2001: 361)

In my view, the distribution of “states” with prepositions in Kabyle can be straightforwardly described in terms of case government, i.e., certain prepositions assigning particular cases to their complements. The synchronic pattern of use of the “annexed” and “absolute” “states” with prepositions in Kabyle, including the case variation with the preposition \( \textit{s} \) ‘with/towards’, should be taken at face value, i.e., as unmotivated by any coherent and straightforward semantic feature.\(^6\) Any valid and non-circular explanatory hypothesis about this issue must inevitably refer to the diachronic dimension; see, e.g., Mettouchi (2008: 31–34) for one plausible hypothesis about the origins of the contemporary distribution. By contrast, attempts to account for this distribution in synchronic functional terms such as “providing a specific value for a grammaticalized meaning” lead to counterintuitive and non-explanatory generalizations of the type criticized in Section 3. Here I would like to reiterate that I am not against functional explanations of linguistic phenomena by themselves and I admit that

---

\(^6\) Again, see Bendjaballah \& Haiden (2005) for an attempt of a phonological account of this distribution; it must be noted, however, that these authors do not consider the issue of variable “state assignment” of the preposition \( \textit{s} \). It should also be noted that in closely related Tamashek (Heath 2005: 291–292) those preposition-like elements which require their complement to appear in the “absolute” state are morphosyntactically distinct from “true prepositions” (Heath 2005: 272–275) requiring the “annexed state” in that, among other things, they combine with independent, and not suffixed, pronouns. The elements in question should then be analyzed as a separate lexical category distinct from prepositions, and probably entering into a different syntactic configuration with their arguments.
the functional account of the Kabyle “annexed state” proposed by M&F is able to coherently motivate most of its various uses. It is the attempt to reduce all the variety of these uses, even those which do not naturally lend themselves to a characterization in terms of “providing a specific value for a grammaticalized meaning encoded in the preceding constituent”, to a single all-encompassing function, which, in my view, seriously undermines the conceptual and empirical merits of M&F’s account. Methodological reductionism, aiming at an account of a phenomenon as complex and diverse as the one found in Kabyle in terms of a single and not fully explicitly formulated function, should be avoided.

I would like to conclude this section by briefly providing a characterization of the Kabyle “state” system in terms of the typology of marked-S languages set up in Handschuh (2014), where, as I have already mentioned above, Berber languages were not investigated. In this work the comparison of marked-S systems is based on the selection of contexts where either the morphologically marked form used for the S relation or the formally unmarked nominal occurring in the P function can appear in different languages. These contexts include a variety of nominal predication constructions, existential and locational predication, preverbal topicalization of subjects (“emphatic subjects” in terms of Handschuh 2014: 99), subjects of dependent clauses, subject of predicates marked with valency-decreasing operations such as passive or antipassive, as well as adnominal modifiers and citation forms of nouns. Since M&F provide a detailed and well-illustrated account of the various uses of the Kabyle “states”, assigning Kabyle the values of the above parameters is mostly unproblematic. The results are presented in Table 1, which reproduces Table 8.1 from Handschuh (2014: 204) with the column for Kabyle added; the order of rows reflects the decreasing frequency of the occurrence of the form used for S (“S-case”) in a particular context, based on a crosslinguistic sample. For the sake of comparability, I identify the “annexed state” of Kabyle with the “S-case” and the “absolute state” with the “zero-case”. The numbers in the Kabyle column refer to the examples in Mettouchi & Frajzyngier (2013); where the data in this paper was insufficient or lacking, I consulted other sources, as indicated in the notes; in some cases I had to revert to the closely related Tamashek as described by Heath (2005), assuming there being no significant difference between it and Kabyle; such cases are indicated by asterisk in the table.

Table 1 clearly shows that Berber is in fact not very “exotic” as a representative of the marked-S type: all functions which are more often than not covered by the S-case in marked-S languages are expressed by the “annexed state” in Kabyle as well. The only really divergent feature of Kabyle is the use of the same form both for S and for the adnominal possessor; the only two other such
languages reported by Handschuh (2014: 194) are Dinka (Nilotic) and the already mentioned Nias. It must be noted, however, that the co-expression of the adnominal possessor and the transitive agent in ergative alignment is quite common crosslinguistically, see, e.g., Palancar (2009: 568) and Lander (2009: 590), and although none of the plausible historical scenarios advanced for the emergence of the Berber “state” distinction assumes original ergativity, I nevertheless find it non-coincidental that “marked subjects” can be crosslinguistically expressed in the same way as adnominal possessors. Moreover, if we take into consideration what Handschuh (2014: 131) calls “subjects of non-basic clauses”, we will find subject-possessor syncretism to be actually quite widespread. Consider again Lithuanian, which has several impersonal passive constructions in which both transitive and intransitive verbs may appear (see, e.g., Timberlake

<table>
<thead>
<tr>
<th>Role</th>
<th>Kabyle</th>
<th>Ø-case</th>
<th>S-case</th>
<th>Other</th>
<th>Total languages</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>%</td>
<td>Number</td>
<td>%</td>
<td>Number</td>
</tr>
<tr>
<td>S argument</td>
<td>S (9)</td>
<td>0</td>
<td>23</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>Locational S</td>
<td>S (35)</td>
<td>0.5</td>
<td>21.5</td>
<td>98</td>
<td>0</td>
</tr>
<tr>
<td>S VDC</td>
<td>S (35)</td>
<td>3.5</td>
<td>11.5</td>
<td>77</td>
<td>0</td>
</tr>
<tr>
<td>Positive</td>
<td>S (42)</td>
<td>4.5</td>
<td>14.5</td>
<td>73</td>
<td>1</td>
</tr>
<tr>
<td>existentials</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adverbial clauses</td>
<td>S (5)</td>
<td>2</td>
<td>22</td>
<td>6.5</td>
<td>72</td>
</tr>
<tr>
<td>Nominal predication</td>
<td>S (15)</td>
<td>7</td>
<td>33</td>
<td>14</td>
<td>67</td>
</tr>
<tr>
<td>Negative existentials</td>
<td>Ø-S (40), (41)</td>
<td>4</td>
<td>40</td>
<td>6</td>
<td>60</td>
</tr>
<tr>
<td>Relative clauses</td>
<td>S (12)</td>
<td>4</td>
<td>41</td>
<td>8</td>
<td>47</td>
</tr>
<tr>
<td>Emphatic S</td>
<td>Ø (4)</td>
<td>7</td>
<td>41</td>
<td>8</td>
<td>47</td>
</tr>
<tr>
<td>Complement clause</td>
<td>S (10)</td>
<td>3</td>
<td>60</td>
<td>2</td>
<td>40</td>
</tr>
<tr>
<td>Predicate nominal</td>
<td>Ø (1)</td>
<td>17</td>
<td>77</td>
<td>4</td>
<td>18</td>
</tr>
<tr>
<td>Term of address</td>
<td>Ø (8)</td>
<td>8.5</td>
<td>65</td>
<td>0.5</td>
<td>4</td>
</tr>
<tr>
<td>Attributive possessor</td>
<td>S (12)</td>
<td>10</td>
<td>45.5</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Citation form</td>
<td>Ø (7)</td>
<td>22.5</td>
<td>98</td>
<td>0.5</td>
<td>2</td>
</tr>
</tbody>
</table>

a. Only indirect evidence.
b. Valency-decreasing constructions.
1982, Wiemer 2006, Lavine 2006), and in both cases the argument corresponding to the subject of the regular active clause appears in the genitive, see (10a–c).

(10) a. Paskutini-u met-u vaik-o mieg-a-m-a blogai.
    last-INS,SG,M year-INS,SG child-GEN,SG sleep-PRS-PP-DF badly
    ‘The child has been sleeping badly recently.’ (Wiemer 2006: 277)

b. Vaik-o sudaužy-t-a puodel-is.
    child-GEN,SG break-PST,PP-DF cup-NOM,SG
    ‘The child apparently broke the cup.’ (Lavine 1999: 45)

c. vaik-o puodel-is
    child-GEN,SG cup-NOM,SG
    ‘(the/a) child’s cup’

On the other hand, in the Uto-Aztecan languages it is normal for the subjects of certain kinds of embedded clauses headed by nominalized verbs to be marked in the same way as adnominal possessors (these subjects can probably even be analyzed as adnominal possessors synchronically), see (11a–c) from Yaqui (Southern branch of Uto-Aztecan, Mexico).

(11) a. mache’eta-m ne tea-k [em am su’utoja-ka-po]
    machete-PL 1SG,NOM find-PFV 2SG,GEN 3PL,ACC release-PFV-SBD
    ‘I found the machetes where you left them.’ (Guerrero 2004: 42)

b. emp-o au wawate-k [beas ketgo em yi’i-ne-’u].
    2SG,NOM that remember-PFV this morning 2SG,GEN dance-EXPE-SBD
    ‘You remembered that you would dance this morning.’ (Guerrero 2004: 323)

c. em achai
    2SG,GEN father
    ‘your father’ (Guerrero 2004: 180)

However remote these typological parallels to the Berber subject-possessor syncretism may be, they are potentially instructive, especially for possible historical accounts; for instance, the development from a possessive or genitive marker to a (non-topical) subject marker via marking the subject of embedded clauses is documented, e.g., in the history of Japanese, see Syromjatnikov (1972: 98–100, 1983: 55) and Bentley (2001: 90–92). Drawing such parallels, however, would not have been possible if the notion of “case” had not been applied to the Berber phenomena.

The place of Kabyle among the other marked-S languages can be clearly seen on the SplitsTree diagram produced by the NeighborNet algorithm (Huson
The diagram shows that Kabyle patterns squarely within the typological space of marked-S languages, displaying more or less “regular” behavior, in distinction to such clear “outsiders” as Nias or Savosavo. Moreover, it is perhaps not surprising that the diagram shows Kabyle to belong to the large African cluster of marked-S languages characterized by a number of shared properties (see Handschuh 2014: 219). However, the diagram also makes it clear that Kabyle is not very similar to any of the other languages, which is not unexpected given some special features of its case system.

To conclude, I hope to have shown that, (i) the arguments adduced by M&F against treating the Kabyle “states” as an instance of case are based on a simplistic and ill-informed view about what case is and what kinds of case systems exist in the languages of the world; (ii) that factors relevant for the distribution of “states” in Kabyle are recurrent in determining case usage and case alternations crosslinguistically; and (iii) that once identified as an instance of a case system, and a “marked-S” system in particular, the “state” category of

---

7 I am grateful to Corinna Handschuh for generously sharing her database with me.
Kabyle turns out to fit unproblematically into the independently established typology of such languages. These conclusions, in my view, strongly support the treatment of the Berber “state” distinction as case, contra M&F and in accordance with many earlier proposals.

5 “Unrecognized typological category” or conceptual fallacy?

Even if M&F had indeed succeeded in providing a uniform functional account for the distribution of “states” in Kabyle, there remains a question of the typological implications of their contribution. In their conclusion M&F claim that their analysis of Kabyle has straightforward implications for linguistic typology in that they have demonstrated that Kabyle possesses “a new type of morphological coding that has not been recognized in descriptions of individual languages, in theoretical literature, or in typology so far” (Mettouchi & Frajzyngier 2013: 28). I believe that this typological conclusion is too far-fetched and, again, misguided.

The general logic behind M&F’s reasoning, as I understand it, is as follows: (i) they take a particular distribution of grammatical forms in a language they know well; (ii) they argue that this language-particular distribution cannot be fully accounted for by any of the notions existing in current linguistic theory and language typology; (iii) they propose a new notion which, as they argue, is able to better account for this language-particular distribution; (iv) they claim that this new notion is of relevance for typology. This kind of reasoning cannot but strike one as potentially leading to the recognition of an infinite number of “typologically relevant categories” derived from all kinds of language particular distributions of grammatical forms or constructions. First of all, as I have already argued in Section 4, the argument in (ii) is not really valid, since crosslinguistic categories are not as such suitable for a full account of language-particular distributions (see Haspelmath 2010). Second, even if the argument in (ii) is taken to be logically correct, the generalization from language-specific descriptive categories to crosslinguistic comparative concepts (again to use the notions from Haspelmath 2010) cannot be so simple.

The prerequisite for postulating crosslinguistic categories (see, e.g., Dahl 1985; the other commonly used term with a narrower scope is “crosslinguistic gram type”, see Bybee & Dahl 1989) is crosslinguistic comparison establishing that particular functions recognized (or postulated) as universal
are systematically expressed by grammaticalized means in a number of genetically and areally diverse languages. In the domain of dependent marking, such notions as “nominative”, “accusative”, or, more specifically, “marked nominative” can be argued to be crosslinguistic categories, since they – of course, not without variation – recur in a wide variety of diverse languages. Nothing of this kind has been shown by M&F for the “new category” they propose on the basis of Kabyle. Their whole discussion and argumentation pertains strictly to one particular language, Kabyle (even not to the whole language family, Berber), and is entirely based on the language-specific properties of language-specific inflectional forms of this language.

Moreover, the notions in terms of which the generalization pertaining to the Kabyle “states” is formulated by M&F do not seem to be straightforwardly generalizable to other languages. Of course, such concepts as “grammaticalized function” and “value” are conceived of as universal and not related to any particular language; however, they are too general and vague to be straightforwardly applied to particular languages. In order to map such notions on language-specific categories and relations one necessarily has to make appeal to language-particular forms and oppositions, cf. the following passage from Mettochi & Frajzyngier (2013: 7): “A grammaticalized function is discovered through the contrast between one form and other forms within the functional domains in the language”. In my view, this necessarily implies that if the category whose function is to “provide the value for the variable of the function grammaticalized in the preceding constituent” is recognized as a crosslinguistic category, its instantiations in particular languages, critically depending on which functions are grammaticalized in each language and what kinds of “values” these functions admit, would potentially have very little in common – probably, much less in common than case systems in the languages of the world, even under the most “liberal” understanding of case against which M&F argue.

Therefore, I conclude that even if M&F’s account of the distribution of “states” in Kabyle, its weaker points and shortcomings notwithstanding, could be considered an advance in the understanding of the functioning of this particular language-specific category (which, unfortunately, in my view is not the case), the value of this analysis for typology would lie not in the recognition of a “novel type of relationship” but rather in the admittedly more accurate and more comprehensive description of the facts of Kabyle and in highlighting of the language-specific mechanisms and constraints regulating the distribution of “states”. The uniform function proposed by M&F for the “annexed state” in Kabyle, its problematic aspects discussed in Section 3, and general criticism against monosemy analyses (Croft 2001) put aside, can only be viewed as pertaining to the level of language-specific descriptive categories, which, as
argued by Haspelmath (2010), should be strictly kept apart from crosslinguistic comparative concepts. As to the latter, as I hope to have shown in Section 4, “dependent marking” and more particularly “case” is the most adequate comparative concept to which the “state” distinction in Kabyle can be linked.

Acknowledgements: I thank Eitan Grossman and Martin Haspelmath for their moral support, Corinna Handschuh for her very useful comments on the last version of the article and for sharing her database with me, and Frans Plank and three anonymous LT reviewers – especially one of them – for their useful suggestions. This work has been supported by the Russian Science Foundation, grant # 14-18-03270. All faults are mine.

Abbreviations: 1/2/3 = 1st/2nd/3rd person; A = agent; ABS = absolutive; ACC = accusative; ALL = allative; CONN = connective particle; DAT = dative; DEF = definite; DF = default agreement; ERG = ergative; EXPE = expected; F = feminine; GEN = genitive; INF = infinitive; INS = instrumental; LOC = locative; M = masculine; NEG = negation; NOM = nominative; OBL = oblique case; PA = active participle; PFV = perfective; PL = plural; POSS = possessor; PP = passive participle; PRS = present tense; PST = past tense; QUOT = quotative; RLS = realis; SBD = subordinator; SG = singular.

References


Gutova, Jenia. 2011. Ismnnaggar tr Rbbi y yan lmakan (‘God made them meet each other in one place’): Causative constructions in Berber. Manuscript. [https://www.academia.edu/4443976/]
Malchukov, Andrej & Helen de Hoop. 2011. Tense, aspect, and mood based differential case marking. Lingua 121. 35–47.


