RECONSTRUCTION OF MORPHOSYNTACTIC FUNCTION: NONSPATIAL USAGE OF SPATIAL CASE MARKING IN TSEZIC

MICHAEL CYSOUW
Max Planck Institute for Evolutionary Anthropology

DIANA FORKER
Max Planck Institute for Evolutionary Anthropology

The reconstruction of genealogical relationships between languages is traditionally performed through lexical comparison and the establishment of regular sound changes. The historical analysis of other aspects of linguistic structure, like syntactic patterns or the function of grammatical elements, is normally understood to depend on a previously established thorough understanding of the genealogical structure of the family. In this article we show that it is possible to reconstruct historical changes in morphosyntactic function without assuming detailed knowledge about the genealogical developments of the languages involved. Even more surprisingly, we are able to accurately reconstruct the genealogical structure of a language family using only a limited amount of morphosyntactic variation.

As a case study, this article focuses on the nonspatial usage of spatial case marking in the Tsezic languages, a subgroup of Nakh-Daghestanian, spoken in the Caucasus. A new semantic-map-like visualization is developed to represent functional differences in case marking between the various Tsezic languages. Using MAXIMUM PARSIMONY, a reconstruction is proposed for the historical development of the case marking, and the same visualization is used to present the reconstructed historical developments in an insightful manner. Besides various new insights regarding the development of Tsezic case marking, the approach used in this article presents a generally applicable method for the reconstruction of morphosyntactic function.*

Keywords:

1. INTRODUCTION. Reconstruction of proto-languages is an approach that is mainly used in historical-comparative investigations of lexical elements and sound inventories. Only rarely are reconstructions of morphosyntactic properties worked out in any detail (but see Harris & Campbell 1995). One of the reasons that other such aspects of language reconstruction are only rarely developed in any detail is that there seems to be a general assumption that such reconstruction is a 'high-level' problem presupposing detailed knowledge about the historical developments of the sound system. In this article, we argue that this assumption is not necessary. Morphosyntactic function can be reconstructed without making any assumptions about the tree of the languages involved—it can even present independent evidence for their genealogical history. The basic principle of our approach is that (small) differences in the function of particular morphemes can be used to infer historical developments.

Concretely, in this article we reconstruct some aspects of the history of case marking in the Tsezic languages, a subgroup of the Nakh-Daghestanian family of languages, spoken in the Caucasus. Most importantly, we focus not on the reconstruction of the phonological material of the case markers, but on the reconstruction of their morphosyntactic usage. The Tsezic languages have an extensive case-marking system, and various morphemes used to mark case are obviously cognate. The crucial point that makes Tsezic case marking interesting for historical-comparative reconstruction, however, is that these case markers differ as to the kind of functions in which they are used. We use this variation to draw conclusions about historical developments.

In particular, we focus on the nonspatial usage of case markers that originated as spatial cases. Although the Nakh-Daghestanian languages are famous for their rich

* We thank Bernard Comrie, Martin Haspelmath, Nick Evans, and two anonymous referees for helpful suggestions and comments for improving the presentation of this article.

spatial case systems, not a lot is known about the nonspatial meanings of these cases. Previous research has focused mainly on experiencer constructions and other nonnominative subjects (Comrie 2004, Comrie & van den Berg 2006, Ganenkov 2006). Most closely related to our study, Ganenkov 2005 describes in detail the spatial and nonspatial uses of two suffixes in Nakh-Daghestanian and other languages, and Daniel et al. 2007 contains many examples of case marking in ditransitive constructions.

In this article, we first introduce the Tsezic languages and look at their genealogical relationship as established through lexical comparison. Following this, we summarize spatial case marking in Tsezic and introduce the nonspatial functions that we investigated as to their usage of spatial case markers. Next, the method for reconstructing the proto-Tsezic marking of these nonspatial functions is discussed in great detail. Our reconstruction basically tries to minimize the number of changes necessary to account for the living Tsezic languages, an approach known in biological phylogenetics as MAXIMUM PARSIMONY. To interpret the reconstructed developments we then introduce a specific visual display. This display is a variant of the semantic-map approach used to graphically show linguistic variation. The significant improvement with our approach is that the displays are automatically generated, making them easier to compare across languages. Finally, we give an interpretation of the reconstructed developments of nonspatial case marking in Tsezic.

2. THE TSEZIC LANGUAGE FAMILY. The Tsezic languages are a group of closely related languages that form a subgroup of the Nakh-Daghestanian language family. All Tsezic languages are spoken in small mountain villages in southern Daghestan (Russia). Historical-comparative studies on the Tsezic language family can be found in Bokarev 1959 and Alekseev 2003. Opinions about the subgrouping of the Tsezic languages diverge. One of the first researchers of the Tsezic languages, Bokarev (1959), using the comparative method, divides Tsezic into East Tsezic (Hunzib and Bezhta) and West Tsezic (Tsez and Khwarshi) with Hinuq in between the two groups. Differently, van den Berg (1995:5) maintains that the West Tsezic languages comprise Tsez and Hinuq, and the East Tsezic languages Bezhta and Hunzib, while Khwarshi constitutes a separate northern branch. Nikolaev and Starostin (1994), Alekseev (1999:299), and Korjakov (2006: 21) all propose the currently favored subgrouping of East Tsezic (comprising Hunzib and Bezhta) and West Tsezic (comprising Khwarshi, Tsez, and Hinuq). Nikolaev and Starostin and Alekseev used the historical-comparative method, whereas Korjakov applied the lexicostatistical method with the help of the Starostin formula. Thus, there is a clear consensus that Hunzib and Bezhta form one branch and that Tsez and Hinuq form another branch. There is no clear consensus on the placement of Khwarshi, though the preference of the current research is that it should be grouped together with Tsez and Hinuq.

As yet another approximation of the genealogical structure of the Tsezic language family, we performed a quantitative analysis of the lexical data collected by M. Š. Xalilov, as made available in the Intercontinental Dictionary Series (IDS).[1] This resource provides the lexical equivalences of about 1,300 meanings in many different languages. For our purposes, we selected the five Tsezic languages (including dialects) and a selection of other Nakh-Daghestanian languages. For all pairs of these languages we computed the average Levenshtein distance, resulting in approximate pairwise lexical similarities between the languages. Roughly speaking, the Levenshtein distance

[1] Available online at http://lingweb.eva.mpg.de/ids/.

(Levenshtein 1966) between two words is the number of changes in letters/sounds that is needed to change one word into the other.[2] The resulting distance matrix was used to draw a NeighborNet.[3] The network of the Tsezic languages is shown in Figure 1, including two variants of Tsez and two variants of Khwarshi.[4] As expected, the tree clearly shows the eastern subgroup, consisting of Bezhta and Hunzib, and (though somewhat less clear) a subgroup consisting of Hinuq and Tsez. Consistent with the dispute in the literature, Khwarshi is intermediate between these two groups.

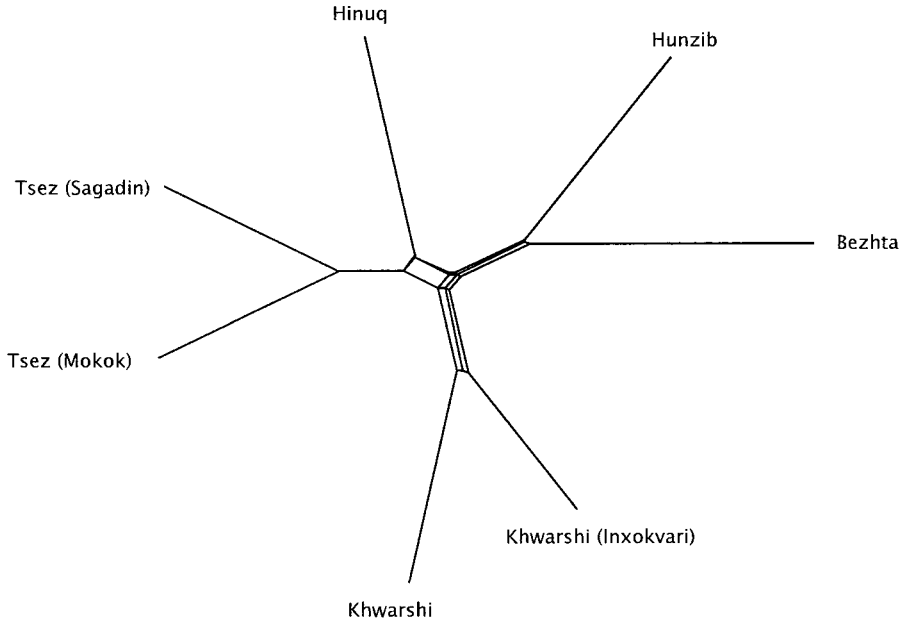

FIGURE 1. Network of the Tsezic languages, based on average Levenshtein distances of the lexical items in the IDS.

Similarities alone cannot provide a starting point (i.e. a root) from where the tree is to be read. This is a principal limitation of subgrouping by pairwise similarities. On such a basis it is only possible to obtain a nested subgrouping of languages. One of

[2] See Heeringa 2004 and Heeringa et al. 2006 for some examples and discussion of using average Levenshtein distance to obtain language similarities.

[3] A NEIGHBORNET is an (unrooted) splits graph based on the NEIGHBOR-JOINING algorithm. The neighbor-joining algorithm was first proposed by Saitou and Nei (1987). The method for drawing SPLITS GRAPHS was introduced by Bandelt and Dress (1992). We used the software SplitsTree (Huson & Bryant 2006) to draw the networks shown here. This software is available online at http://www.splitstree.org/. A splits graph does not force the languages to be placed in a bifurcating tree but allows conflicting evidence to be shown as reticulations (i.e. the little boxes in Fig. 1). The more tree-like the resulting network is (i.e. the thinner the boxes are), the better the evidence that the pairwise similarities make up a bifurcating tree. In the case of the Tsezic languages, the network is very tree-like.

[4] This approach to subgrouping does not provide genealogical evidence, but it gives a good impression of the lexical similarity between the languages. For a more traditional kind of lexicostatistics, the IDS lists would have to be annotated with cognate judgments, and these cognate sets could then be used to infer a genealogy, for example using the assumption that lexical replacement can be modeled by Dollo-parsimony (i.e. cognates can be lost for a particular meaning, but regained only with very low probability). This approach has been used by Gray and Atkinson (2003) for Indo-European. To perform these analyses for Tsezic, however, would amount to a project separate from our current endeavor.

the possibilities to root a tree is to include some languages in the analysis that are known to be distantly related, a so-called OUTGROUP. The branch where the Tsezic family joins these distantly related languages indicates where the root of the family is. We have selected a few Nakh-Daghestanian languages as an outgroup to root the Tsezic family, as shown in Figure 2. At the bottom of the graph, the Tsezic group is clearly separate from the other languages. More importantly, the location where the other Nakh-Daghestanian languages are connected to the Tsezic languages indicates that Khwarshi forms a subgroup with Hinuq and Tsez. But evidence for the position of Khwarshi within the Tsezic family also remains meager using this approach.[5] More in-depth investigation using the traditional historical-comparative method is necessary to resolve this issue.

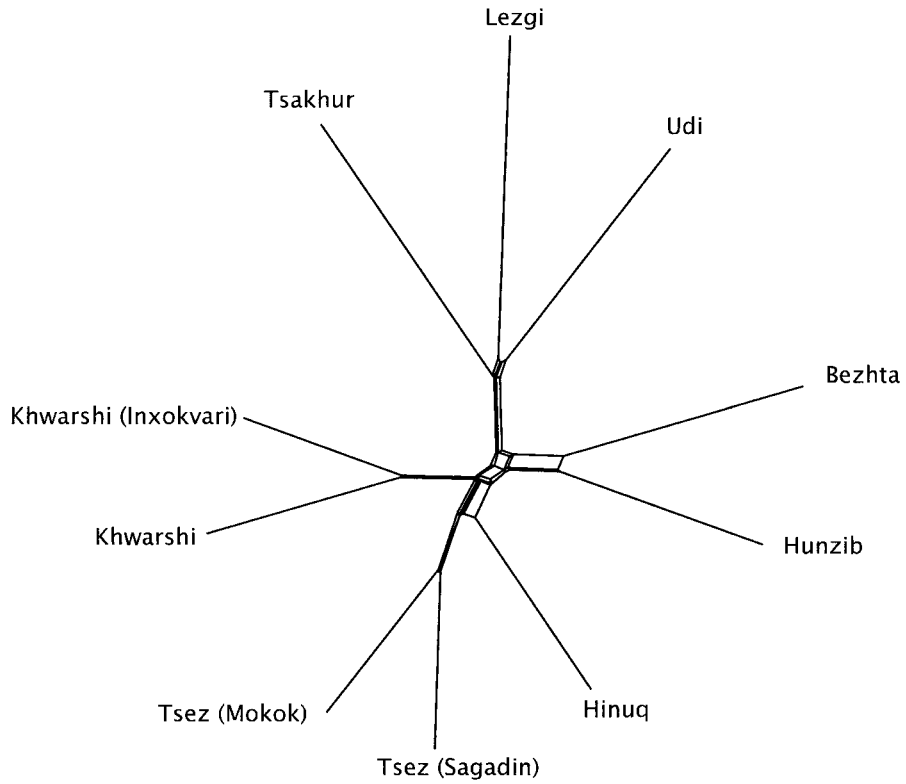

FIGURE 2. Network of selected Nakh-Dagestanian languages, based on average Levenshtein distances from the IDS.

For the current article we assume that Khwarshi belongs in one subgroup together with Tsez and Hinuq. If this assumption would turn out to be incorrect, we would have to slightly revise our case-marking reconstructions for proto-Tsezic. Our reconstructions method, to be described in detail in §4, does not depend on this assumed tree, however, and we would only need to perform one simple recalculation to provide updated reconstructions for a different rooting of the Tsezic family tree.

[5] In the IDS data there is information available about nineteen other Nakh-Dagestanian languages with thirty-six (dialect) variants. The subgrouping of Khwarshi with Hinuq and Tsez is always found, taking any of these languages as outgroup.

3. TSEZIC CASE MARKING.

3.1. SPATIAL CASE MARKING. The inventory of grammatical case markers of the Tsezic languages includes basic cases like absolutive, ergative, instrumental, and genitive.[6] Hinuq also has a dative case. Most famously, however, the Tsezic languages have an exuberant system of spatial case markers, sometimes claimed to result in the largest case-marking systems ever encountered in a human language (but see Comrie & Polinksy 1998 for a rebuttal of this claim). The spatial case marking of the Tsezic languages actually consists of two morphemes, one marker for location and one for direction, that can be combined to form complex categories.[7] Table 1 presents a survey of the spatial case markers in the five Tsezic languages. The markers in this table are organized in columns with regard to their locational element, and in rows with regard

HINUQ	CONT	IN	SUB	SPR	AT	APUD	AD	NEAR
essive	-ɬ	-V/ma	-ƛ	-ƛ'o	-qo	-de	-ho	
lative	-ɬ-er	-V/ma-r	-ƛ-er	-ƛ'o-r	-qo-r	-de-r	-ho-r	
ablative/genitive	-ɬ-es	-V/ma-s	-ƛ-es	-ƛ'o-s	-qo-s	-de-s	-ho-s	
ablative	-ɬ-ezo	-V/ma-zo	-ƛ-ezo	-ƛ'o-zo	-qo-zo	-de-zo	-ho-zo	
directive	-ɬ-edo	-V/ma-do	-ƛ-edo	-ƛ'o-do	-qo-do	-de-do	-ho-do	
TSEZ								
essive	-ɬ	-ā	-ƛ	-ƛ'(o)	-q(o)	-de	-x(o)	
lative	-ɬ-er	-ā-r	-ƛ-er	-ƛ'o-r	-qo-r	-de-r	-xo-r	
ablative	-ɬ-āy	-ā-y	-ƛ-āy	-ƛ'-āy	-q-āy	-d-āy	-x-āy	
versative	-ɬ-xor	-ā-ɣor	-ƛ-xor	-ƛ'-āɣor/ār	-q-āɣor/ār	-d-āɣor/ār	-x-āɣor/ār	
KHWARSHI								
essive	-ɬ	-(m)a	-ƛ	-ƛ'o	-qo		-ho	-ɣo
lative	-ɬ-ul	-(m)a-l	-ƛ-ul	-ƛ'o-l	-qo-l		-ho-l	-ɣo-l
ablative	-ɬ-zi	-(m)a-zi	-ƛ-zi	-ƛ'o-zi	-qo-zi		-ho-zi	-ɣo-zi
versative	-ɬ-ɣul	-(m)a-ɣul	-ƛ-ɣul	-ƛ'o-ɣul	-qo-ɣul		-ho-ɣul	-ɣo-ɣul
translative	-ɬ-ɣužaz	-(m)a-ɣužaz	-ƛ-ɣužaz	-ƛ'o-ɣužaz	-qo-ɣužaz		-ho-ɣužaz	-ɣo-ɣužaz
terminative	-ɬ-q'a	-(m)a-q'a	-ƛ-q'a	-ƛ'o-q'a	-qo-q'a		-ho-q'a	-ɣo-q'a
BEZHTA								
essive	-ɬ	-ʔ	-ƛ	-ƛ'a	-qa	-doy		-ɣa
lative		-ʔ-il		-ƛ'a-l		-doy-l		-ɣa-l
ablative/genitive	-ɬ-so	-ʔ-is	-ƛ-so	-ƛ'a-s	-qa-s	-doy-s		-ɣa-s
ablative	-ɬ-la	-ʔ-la	-ƛ-la	-ƛ'a-la	-qa-la	-doy-la		-ɣa-la
directive	-ɬ-daa	-ʔ-daa	-ƛ-daa	-ƛ'a-daa	-qa-daa	-doy-da		-ɣa-daa
translative	-ɬ-laƛ'ā	-ʔ-(la)ƛ'ā	-ƛ-laƛ'ā	-ƛ'a-laƛ'ā	-qa-laƛ'ā	-doy-		-ɣa-laƛ'ā
HUNZIB								
essive	-ɬ	-V	-ƛ	-ƛ'(o)	-g(o)	-dər		
ablative/genitive	-ɬ-sə	-V-s		-ƛ'o-s	-go-s	-dər-sə		
translative	-ɬ-ƛ'i	-V-ƛ'		-ƛ'o-ƛ'		-dər-ƛ'i		
directive	-ɬ-do	-dɑ-ɑ				-dər-do		

TABLE 1. Spatial case markers in Tsezic.[8]

[6] To be precise, all Tsezic languages (except for Hunzib) have two different genitive markers. The first genitive marker is used when the head noun is in the absolutive case, the second genitive otherwise. The genitive that is discussed in this article, reconstructed for proto-Tsezic as *-*s*, is the former.

[7] In addition to these two parameters of spatial location and direction, Tsez has a third parameter indicating distality, that is, whether the position in question is located near or far from speaker and hearer (see Comrie 1999 for details).

[8] In Khwarshi, all locational markers ending in /o/ have an allomorph ending in /a/. In Bezhta, the IN-Lative and the IN-Ablative/Genitive have allomorphs in which the epenthetic vowel /i/ is either /a/, /e/, or /o/. In Bezhta, the APUD morpheme *-doy-*, as found in the texts that we used, is listed as *-der-* in Kibrik & Testelets 2004. This is possibly a dialectal difference.

to their directional element (the labels used for the columns are discussed later in this section). Empty cells indicate combinations that are not attested in our data, or complete absence of a particular locational marker (viz. in the case of empty columns in the table). The fact that various combinations in Bezhta and Hunzib are not attested is probably due (at least partly) to the limited amount of data at our disposal.

The locational markers in the columns of Table 1 are identified across languages on the basis of their formal similarity. These morphemes appear to be reconstructible rather straightforwardly for proto-Tsezic, as shown in Table 2 (cf. Alekseev 1999, 2003). As an indication of the genealogical relatedness of the locational markers, we use the same glossing across all languages. The actual glosses for the locational markers used in the literature differ slightly from language to language, but we have unified the names of the markers for the sake of clarity and intelligibility. The different glosses are summarized in Table 3.[9]

PROTO-TSEZIC RECONSTRUCTION	APPROXIMATE MEANING IN SPATIAL DOMAIN	GLOSS
*-ɬ-	in contract	CONT
*-ʔa-	in, inside	IN
*-x̌-	under	SUB
*-x̌'o-	on top, above	SPR
*-qo-	attached to	AT
*-de-	near, by (animates)	APUD
*-ho-	near, by (inanimates)	AD
*-ɣo	near, close	NEAR

TABLE 2. Reconstructions of locational markers.

	GLOSSES FOUND IN THE LITERATURE				
OUR GLOSS	HINUQ	TSEZ	KHWARSHI	BEZHTA	HUNZIB
CONT			INTER		
IN					DAT
AT	POSS	POSS	CONT	POSS	AD
APUD					ALL
NEAR			APUD	AD	

TABLE 3. Differences in glossing of locational markers.

The functions of the locational markers in Table 2 are also very similar across the Tsezic languages. There are small differences between the Tsezic languages as to the spatial uses of these markers, but this variation is not discussed further in this article.[10] The core system of locational markers found in all five Tsezic languages comprises the markers CONT, IN, SUB, SPR, and AT. Their spatial meaning can be described roughly as follows:

- CONT ('in contact') indicates a location inside of an amorphous mass (e.g. in water, flour, or ashes), or a location that could be conceptualized as being a kind of mass (e.g. in an avalanche, in a forest, or in leaves). It is also used to mark geographic location (e.g. in a state or province). In Hunzib and Bezhta this case is also used for vertical attachment (e.g. on the wall).

[9] The sources for these glosses are as follows: Hinuq (D. Forker, unpublished material), Tsez (Comrie 2004), Khwarshi (Z. Xalilova, p.c.), Bezhta (Kibrik & Testelets 2004), and Hunzib (van den Berg 1995).

[10] Ideally, an investigation of the spatial marking along the lines of Levinson et al. 2003 would bring out the details of the meaning of the various markers.

- IN ('inside') indicates a location in some kind of container (e.g. a box, belly, or corner), and various 'in' or 'on' locations (e.g. in the village, in the river, on the way, on the field, in the seventh grade).
- SUB ('under') indicates a location under any kind of object (occasionally also behind).
- SPR ('above') indicates a location on a flat surface (e.g. on a bed, table, or square) or on the top of vertical objects (e.g. on a mountain, tree, or staircase). It is also used for locations in/on vehicles (e.g. car, bike, horse, ship).
- AT is used for general location and direction ('at', 'on', 'to', 'in'), but spatial meanings are rare with this suffix. In Hunzib and Bezhta, there are no examples of this case describing spatial location with inanimate nouns.

Furthermore, there are three different markers all indicating some kind of nearness. The precise differences between their spatial usage needs more investigation. Also note that none of the Tsezic languages has reflexes of all three of these markers:

- APUD indicates a location near objects that are almost exclusively animate nouns (this marker is absent in Khwarshi).
- AD indicates a location near objects that are almost exclusively inanimate nouns (this marker is absent in Bezhta and Hunzib).
- NEAR is used for various situations in which an object is near (attested only in Khwarshi and Bezhta).

The directional parts of the spatial case markers (i.e. the rows in Table 1) show much more variation across the Tsezic languages. Only the ESSIVE, LATIVE, and ABLATIVE/GENITIVE are clearly reconstructible, as shown in Table 4. The essive is actually both formally and functionally an unmarked case, and therefore we have decided to gloss it as NULL in our reconstructions. The lative, glossed as LAT, is found in most Tsezic languages, but Hunzib has probably lost it. A trace of the lative is arguably preserved in Hunzib in the suffix of the terminative converb *-čor*,[11] which we have analyzed as consisting of the morphemes *-čo* 'present participle' and *-r* 'lative'. The ablative/genitive directional series is attested only in Hinuq, Bezhta, and Hunzib. In the spatial usage, this suffix clearly has an ablative meaning. But the marking of this ablative is identical to the marking of the genitive as attested in all Tsezic languages, using a suffix reconstructed here as **-s*. For this reason, we gloss this morpheme as GEN.

PROTO-TSEZIC RECONSTRUCTION	APPROXIMATE MEANING IN SPATIAL DOMAIN	GLOSS
*-Ø	essive ('location')	NULL
*-r/l	lative ('orientation toward a goal')	LAT
*-s	ablative/genitive ('movement from, or out of')	GEN

TABLE 4. Reconstructions of directional markers.

Further ablative-like cases are attested in Hinuq, Tsez, Khwarshi, and Bezhta, though only the Hinuq and Khwarshi morphemes appear to be historically related. A VERSATIVE is found in Khwarshi and Tsez, the morphemes of which seem to be related. Its meaning is close to the lative, but not identical because it refers more to movement toward a goal, not orientation. The other three languages, Bezhta, Hunzib, and Hinuq, have a functionally and formally similar case, the DIRECTIVE, which expresses movement in the direction of a goal. Khwarshi, Bezhta, and Hunzib have further cases with transla-

[11] Converbs are nonfinite verb forms whose main function is to mark adverbial subordination (Haspelmath 1995:3).

tive-like meanings, roughly 'along' or 'through'. The morphemes do not seem to be historically related. Finally, Khwarshi is the only language that has a TERMINATIVE. All of these additional directional markers only rarely occur with nonspatial meanings, and are not further considered here.

3.2. NONSPATIAL USAGE OF SPATIAL CASE MARKING. Besides spatial functions, the previously discussed case markers are used in a wide range of nonspatial functions such as (i) temporal and metaphorical location and direction, (ii) marking of arguments of particular constructions, and (iii) marking of nonfinite verb forms in adverbial clauses. We investigated the case marking of thirty-five frequently occurring nonspatial constructions. These constructions were selected because they show spatial case marking in many Tsezic languages (see Table 5 for a summary of the thirty-five constructions). We were particularly interested in looking at which type of spatial case marking is attested in which type of construction in the various Tsezic languages. The case-marked constituents of which we investigate the case marking are surrounded by square brackets in the table.

ROLE	LABEL	CONSTRUCTION
addressees	TALK	talk [to somebody]
	TELL	tell [somebody]
	SHOUT	shout [at somebody]
	SAY	say [to somebody]
	ASK	ask [somebody]
	BEG	beg [somebody]
	EXPLAIN	explain [to somebody]
	TEACH	teach [somebody]
	ORDER	order [somebody]
human objects	MARRY	marry [a man]
	FEAR	fear [somebody]
	BELIEVE	believe [somebody]
	LISTEN	listen [to somebody]
recipients	GIVE (PERMANENTLY)	give X as a gift [to somebody]
	GIVE (TEMPORARILY)	give X to hold [to somebody]
	SHOW	show X [to somebody]
possessors	INALIENABLE POSSESSION	[somebody] has a daughter
	ALIENABLE POSSESSION	[somebody] has money
subjects	AGE	[somebody] is X years old
	NAME	[somebody] is called X
	FIND	[somebody] finds X
objects	EXCHANGE	give X away [for something else]
	PURPOSE/GOAL	go with the goal to get [something]
	LOOK	look [at something]
	METAPHORICAL LOCATION	talk [about something]
causees	CAUSATIVE	cause [somebody] to do X
potential agents	POTENTIAL	[somebody] can do X
	ABLE	[somebody] is able to do X
reasons	NATURAL FORCE	[because of something] X happened
converbs	SIMULTANEOUS	[while X happened] . . .
	TERMINATIVE	[until/before X happened] . . .
	CAUSAL	[because X happened] . . .
	POSTERIOR	[after X happened] . . .
time indications	TIME SPAN	something happened [during period X]
	TIME POINT	something happened [at time X]

TABLE 5. Nonspatial contexts for case marking.

The information on the case marking of these constructions has been gathered by investigating various texts. Many texts have been collected in the field by Zaira Xalilova (Khwarshi) and one of the authors of this article (D. Forker, for Hinuq). The Bezhta texts, which also served as a basis for our investigation, are the memories of Šeyx Ramazan, written down by himself at the end of the last century, translated and edited by Madžid Xalilov, and glossed by one of the authors (D. Forker). As yet all of these texts are unpublished. The Tsez texts that we have used are currently in press (Abdulaev 2009). For Hunzib, the grammar by van den Berg (1995) has been the main source. Additional sources were the series of dictionaries of the Daghestanian languages, edited by the Daghestan Scientific Centre of the Russian Academy of Science (Xalilov 1995, 1999, Xalilov & Isaakov 2001, 2005), and some short grammatical descriptions (Radžabov 1999, Comrie 2004, Kibrik & Testelets 2004). Some of the constructions allow only one kind of case marking, but sometimes there is more than one possibility attested. For this study, only the most typical marking has been considered, though attested variation is incidentally used to help historical-comparative reconstructions (see §4.2). The case marking used in the five Tsezic languages for the thirty-five constructions is summarized in Appendix A.

Most of the contexts in the table speak for themselves. There are, however, a few contexts that are based on special structures attested in the Tsezic languages. For example, the context of the construction MARRY that we used for our investigation is always formulated as a female subject marrying a male object. The reverse situation only rarely occurs in the texts, and when it does, a completely different construction is used. Another special situation is found with the verb 'give'. In all Tsezic languages one of two different cases is used for the recipient of 'give'. One of these case forms indicates that the object that is given is a real gift to the recipient (1a), while the other case form indicates that the object is handed over only temporarily (1b).[12] Interestingly, while all Tsezic languages mark this difference by using case markers on the recipient, each language uses its own selection of cases to make the opposition (see Daniel et al. 2007).

(1) Khwarshi (Z. Xalilova, p.c.)

a. ise heⁿše iłe-l tuƛ-i
he.ERG book she.OBL-LAT give-PST.W
'He gave her the book.' (forever, as a gift)

b. ise heⁿše iłe-ɣo-l tuƛ-i
he.ERG book she.OBL-NEAR-LAT give-PST.W
'He gave her the book.' (only for a certain time)

Another special Tsezic phenomenon is the use of two different constructions to mark potential action. One possibility is to use a potential marker on the verb (we call this construction POTENTIAL). Another possibility is to use an auxiliary, meaning roughly 'be able' (we call this construction ABLE). For both constructions we investigated which case was used for the subject. For example, in Hinuq, as shown in 2, the subject is marked with the general locational case AT-essive in both constructions. As it turns out, all Tsezic languages use the same AT-essive case in both constructions.

[12] The following abbreviations are used in the glossing of the examples in this article: I/II/III/IV/V: gender classes, AT: general location, CAUS: causative, ERG: ergative, ESS: essive, GER: gerund, INF: infinitive, LAT: lative, NEAR: location near, NEG: negation, OBL: oblique, POT: potential, PRET: preterite, PRS: present, PST: past, QUOT: quotative, SPR: super, UNW: unwitnessed, VERS: versative, W: witnessed.

(2) Hinuq (D. Forker, fieldnotes)

a. di-qo buƛe b-u-ɬ-o gom
I.OBL-AT.ESS house(III) III-make-POT-PRS be.NEG
'I cannot build a/the house.'

b. Ayšat-qo k'wezi b-iq-iš b-egi xok'o b-uw-a
name-AT.ESS be.able III-be-PST III-good khinkal(III) III-make-INF
'Ayshat was able to make good khinkal.'

All Tsezic languages have a special construction for saying things like 'the boy's name is John'. This example would be expressed roughly as 'John is at the boy'. An example of this construction in Bezhta is shown in 3. We are interested in the spatial case that is used to mark the thing being named (i.e. the canyon in ex. 3). As it turns out, all Tsezic languages use the superessive case in this construction.

(3) Bezhta (B. Comrie, fieldnotes)
sud wahalliyo kuwa-ƛ'a caaⁿ gey-ƛo nisol-na
why this canyon.OBL-SPR.ESS name be-QUOT ask-PST.UNW
abo-qa öždi
father-AT son.ERG
'The son asked the father, why this canyon had this name.'

There is also a special construction to describe one's age. To express things like 'the boy is ten (years old)', the subject is marked with a spatial case, expressing the sentence roughly as 'ten years is at the boy'. For example, in Hunzib a genitive is used, as shown in 4. Other Tsezic languages use different case marking in this context.

(4) Hunzib (van den Berg 1995:69)
ołu-s oq'e-n rig ƛɨ li
s/he.OBL-GEN four-and ten.OBL year be
'S/he is forty years old.'

The verb meaning 'to find' in Tsezic has an unusual case frame from a western European perspective. The object that is found is marked as absolutive (zero marked and cross-referenced on the verb by gender marking) and the one who finds is marked with a spatial case.[13] For example, in Tsez the finder is marked with a lative case, as shown in 5. As it turns out, most Tsezic languages use a lative case in this construction, except for Hinuq, where a dative case is used.

(5) Tsez (B. Comrie, fieldnotes)
elo yiła-r bercinaw kid y-esu-n
there she.OBL-LAT beautiful girl(II) II-find-PST.UNW
'She found a beautiful girl there.'

Another special Tsezic usage of spatial case marking is exemplified in 6 with a sentence from Tsez. In Tsez, a noun can be marked as the purpose of the sentence by using the superlative case (viz. *qaciƛ'or* in the example). We call this construction PURPOSE/GOAL. Other Tsezic languages use different cases in this context, and we argue that proto-Tsezic probably used a superessive for this function.

(6) Tsez (B. Comrie, fieldnotes)
mamalay ɣudes ciq-a-ɣor qaci-ƛ'o-r b-ik'i-x
rooster daily forest-IN-VERS firewood.OBL-SPR-LAT III-go-GER
zow-n
be.PST-PST.UNW
'Every day the rooster went into the forest for firewood.'

[13] The Tsezic lexical verbs of finding, like *-esu-* in Tsez in 5, are probably better translated as 'be found'.

Finally, Tsezic languages also have a special way to express causatives, exemplified in 7 with an example from Hunzib. The action that is caused, 'kill', is marked with a causative marker, which turns it into a ditransitive verb. The causer of this verb is marked like a regular transitive subject (using the ergative), and the object of the caused verb is marked like a regular transitive object (using the unmarked absolutive, and cross-references on the verb by gender). The interesting aspect for our current purpose is the marking of the causee (the 'father' in the example). In Hunzib, the causee is marked by the general locational marker AT-essive. This marker is found in most Tsezic languages, though in Bezhta an instrumental case marker is used in this construction.

(7) Hunzib (van den Berg 1995:108)

maduhan-li-l	ɑbu-g	sɨ	b-iƛ'e-k'-er
neighbor-OBL-ERG	father-AT.ESS	bear(III)	III-kill-CAUS-PRET

'The neighbor made father kill the bear.'

4. Reconstruction.

4.1. Maximum parsimony. Based on the variation in the usage of the case marking among the Tsezic languages, it is possible to develop hypotheses about the reconstruction of the case usage for proto-Tsezic. Basically, the approach that we use to propose reconstructions is to search for a history that would need the least amount of changes from the proto-stage to the individual languages attested. Of course, it need not be the case that the most parsimonious history is also the real history. Indeed, strictly speaking, our parsimonious reconstruction should be considered only a first guess at the true history. The simplest possible history is the proper starting point of all reconstruction, however, and any nonparsimonious developments (like parallel developments of reversals) are in need of extra argumentation.

To exemplify this kind of argumentation, we assume the tree for the Tsezic languages as described in §2. Then we can reconstruct *AT.NULL as the case marking for the addressee in 'listen [to somebody]' in proto-Tsezic, as shown in Figure 3.[14] This reconstruction is rather straightforward because AT.NULL is the most frequent case for this function, so the changes that are necessary are minimal (there is only one change necessary for this reconstruction). There is no other case that would need fewer, or an equal amount of, changes.

Depending on the distribution of the cases in the tree of languages, sometimes the reconstruction is also possible without one of the cases being numerically predominant. For example, both the cases AT.LAT and NULL.LAT are attested at an equal frequency for the addressee in 'order [somebody]', as shown in Figure 4. But because AT.LAT is attested for Tsez and Hinuq, this case can be confined to the Hinuq/Tsez subgroup, with *NULL.LAT being reconstructed for proto-Tsezic. This reconstruction needs only one change, while a reconstruction of *AT.LAT for proto-Tsezic would minimally need three changes.

[14] In our reconstructions, we use this notation (e.g. 'AT.NULL') to indicate the bimorphemic structure of the case marker used. The first part (AT) is the locational element and the second part (NULL, i.e. 'essive') is the directional element (cf. §3.1). The labels 'AT' and 'NULL' refer to particular morphemes that can be reconstructed for proto-Tsezic (viz. *-*qo*- and *-*Ø*-, respectively; see Table 2). We use the labels to abstract away from the concrete instantiation of these proto-forms in the individual languages. For example, when we say that Khwarshi uses an AT.NULL case, this is an abbreviation for saying Khwarshi uses a case form that is cognate to the proto-Tsezic case *-*qo*-*Ø*.

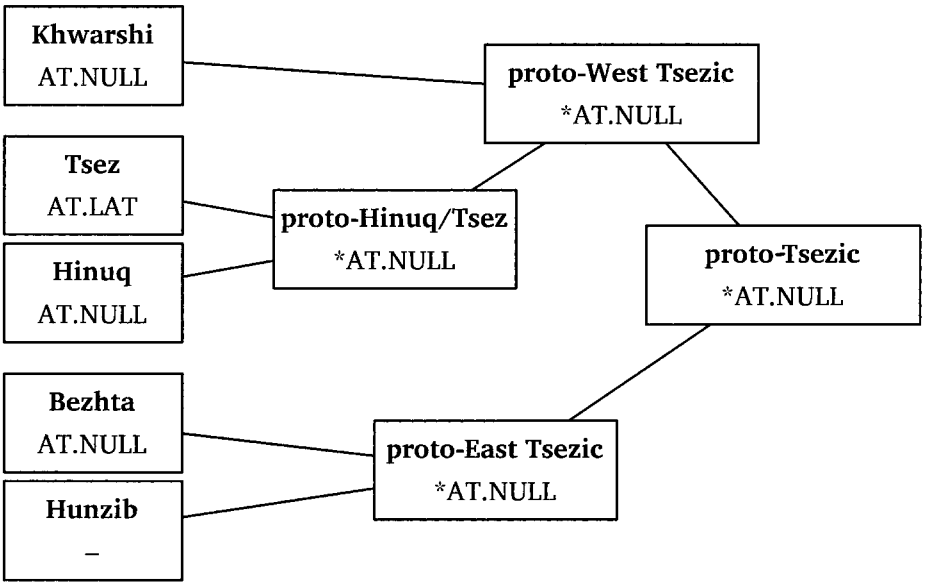

FIGURE 3. Reconstruction of case marking for the addressee of 'listen [to somebody]'.

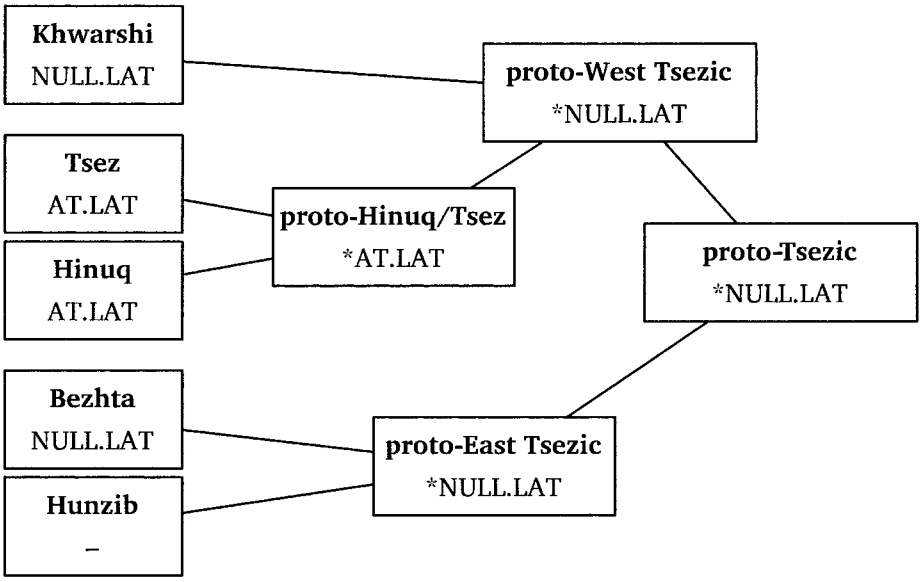

FIGURE 4. Reconstruction of case marking for the addressee of 'order [somebody]'.

Conversely, numerical predominance does not necessarily lead to an unequivocal reconstruction. For example, the case marking of the simultaneous converb (see Figure 5) is SPR.NULL in three out of the five languages. This dominance does not lead to a reconstruction of *SPR.NULL for proto-Tsezic, however, because there is a clear split between *SPR.NULL for West Tsezic and *IN.NULL for East Tsezic. All that can be said is that the second part will very likely be *NULL, but the first part could just as well be *SPR or *IN, since both reconstructions would need one change.

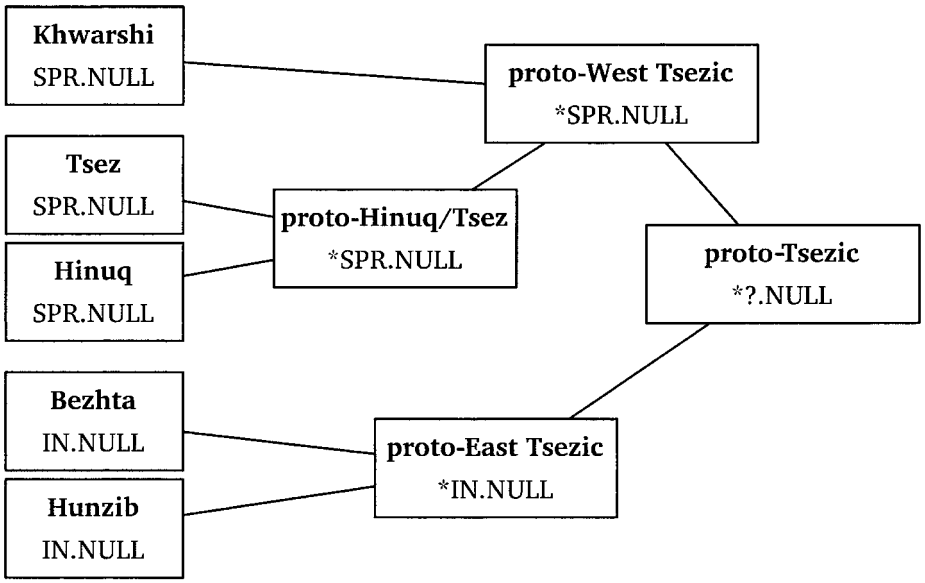

FIGURE 5. Reconstruction of case marking for simultaneous converbs.

The approach of proposing reconstructions through minimization of the number of necessary changes has been developed in the context of biological phylogeny under the name of maximum parsimony. In that context, this approach even goes one step further and searches for the minimal number of necessary changes WITHOUT ASSUMING A TREE. With all examples that we have discussed previously in this section, we have assumed that we know the genealogical tree of the Tsezic languages. The maximum-parsimony approach does not assume this. The trick is basically to try out all possible trees, and search for the optimal reconstruction of each characteristic for each tree. At the end, the total number of changes that are necessary is computed for each tree, and the tree with the lowest total is proposed as the best tree to fit the attested variation.

In biological phylogeny, the pros and cons of maximum parsimony have been discussed extensively (see Felsenstein 2004), and in recent research more advanced methods are normally used. For the current case, however, we considered this somewhat old-fashioned method to be sufficient. Furthermore, it has the benefit of being relatively easy to explain and easy to interpret. See also Dunn et al. 2005 for another example of the utility of maximum parsimony in linguistics.

When we used maximum parsimony on our data for Tsezic nonspatial case marking, there turned out to be just one optimal tree.[15] It is possible, and not even that uncommon, that various trees turn out to be equally parsimonious, so the fact that we find only

[15] For the computation of maximum parsimony, we used the program called 'pars' from the software package 'phylip' by Joe Felsenstein. It is available online at http://evolution.genetics.washington.edu/phylip.html. For the implementation, we coded each part of the bimorphemic case suffixes as separate characteristics (see Appendix A). In a few situations this decision leads to higher counts of the number of changes than actually happened. Specifically, a change of XXX.NULL to NULL.YYY is counted as two changes by the algorithm, though it is probably better analyzed as just being one change from XXX to YYY. This coding decision is not ideal, but it was the only solution we found to consistently deal with both monomorphemic and bimorphemic case suffixes using available software. As far as we can see, the resulting difference in counting does not lead to any different conclusion.

one optimal tree is encouraging. The number of changes and the number of undecidable reconstructions are both low, which likewise suggests that the resulting tree is sensible.[16] The optimal tree is shown in Figure 6. It is an unrooted tree because maximum parsimony (in its basic form) can only detect whether changes are necessary and not in which direction the change happened. But the topology of the tree accurately depicts the assumed genealogical structure of the Tsezic languages as discussed in §2. Also the length of the branches (which is drawn proportionally to the number of changes that are necessary to get from one node to the other) is strikingly similar to the lexical tree. The fact that the morphosyntactic tree matches the lexical tree also substantiates the interpretation that it is genealogical development that is underlying the morphosyntactic variation, and not borrowing or other forces leading to parallel developments.

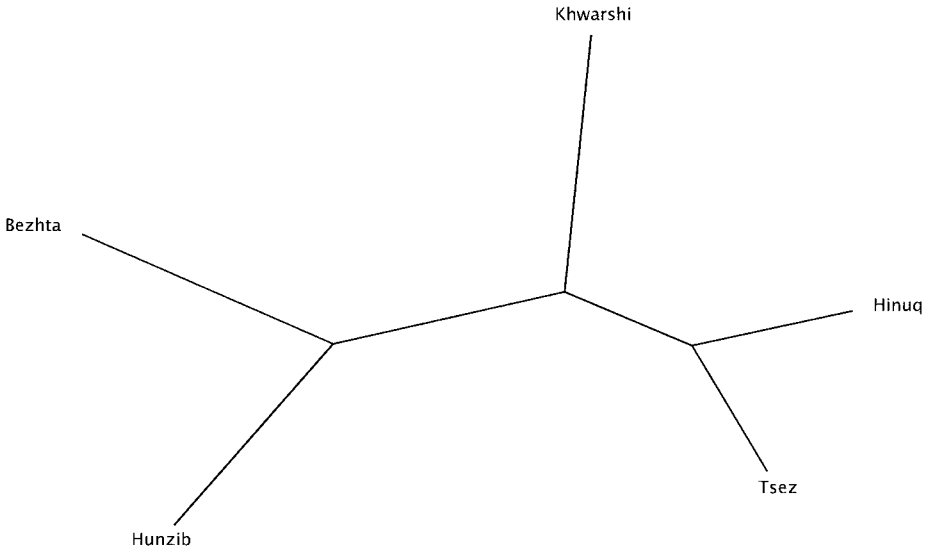

FIGURE 6. Unique optimal (unrooted) tree on the basis of nonspatial case marking (according to maximum parsimony).

The match between the accepted genealogical tree and our tree based exclusively on case marking is a noteworthy result because it indicates that detailed linguistic variation within a narrow domain of linguistic structure can already contain sufficient information to reconstruct the genealogical relationship between languages (at least for this example of a small group of closely related languages). More generally, this suggests that the investigation of the genealogical structure of languages can also be approached from detailed analysis of structural linguistic features, not just from phonological and lexical information. The possibility of using grammatical information for reconstruction is also highlighted by Dunn and colleagues (2005), though their data is of a rather different kind compared to our data. They used highly abstract grammatical characteristics from many different domains. In contrast, we have used fine-grained characteristics from just one small domain of linguistic structure. Our approach can only be used for restricted time depths, however, because we have explicitly used the

[16] The resulting tree needed fifty-four changes in total, which amounts to 11 percent (= 54/7 * 2 * 35) of all vertices, and left twelve undecidable reconstructions, which is 6 percent (= 12/3 * 2 * 35) of all reconstructions for the internal nodes.

formal (but not functional) identification of the case morphemes across languages. In a sense, our approach is more akin to a diachronic Wortfeld investigation in the tradition of Trier 1931.

Besides the optimal tree, the result of the maximum-parsimony analysis also provides the reconstructed status for every intermediate node (there are three intermediate nodes in this tree). With the addition of our assumption (see §2) that the root of the Tsezic tree separates East Tsezic (Bezhta, Hunzib) from West Tsezic (Khwarshi, Hinuq, Tsez), we were also able to provide a reconstruction for proto-Tsezic. All reconstructions are listed in Appendix B.

4.2. Additional diachronic inferences. The advantage of the inferences based on maximum parsimony is that they are independent of the details of the languages under investigation. The same reasoning can also be used for other languages and other data. But there are various details specific to the structure of the Tsezic data that can help to fill in the blanks that remain after maximum parsimony. All of these additional reconstructions (based on the criteria to be discussed in this section) are indicated with italic typeface in Appendix B.

The first additional argument is illustrated with the reconstruction of the recipient in 'show [to somebody]' in Figure 7. Just on the basis of maximum parsimony the first 'locational' part of the case marking is not reconstructible because both *NULL and *AT result in three necessary changes. The second 'directional' part is easily reconstructible as *NULL. This means that we have the options *NULL.NULL or *AT.NULL for proto-Tsezic. The combination *NULL.NULL, however, would mean that the proto-language would have had zero case marking (i.e. absolutive). This seems to be unlikely (because the direct object of 'show' is already marked this way), and therefore we prefer to reconstruct *AT.NULL.[17]

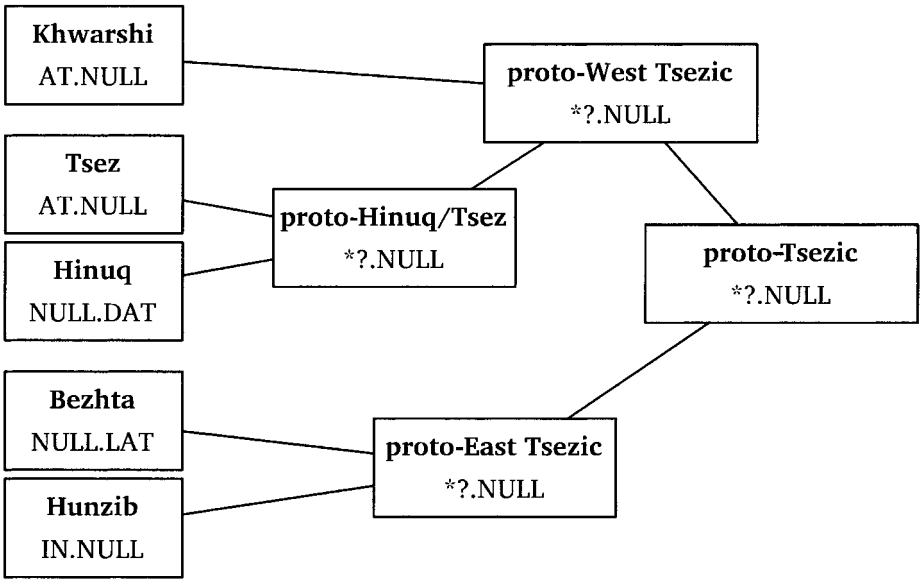

Figure 7. Reconstruction of case marking for the recipient of 'snow [to somebody]'.

[17] This reconstruction depends on the decision to reconstruct both parts of the case suffixes separately (cf. n. 15). When the NULL suffixes are ignored, no case for proto-Tsezic could be reconstructed.

There is another restriction related to the distribution of NULL markers for proto-Tsezic. For spatial case marking it is perfectly possible to combine overt markers for both the locational and the directional part, resulting in case markers like AT.LAT or SPR.LAT. We also find such overt combinations in our nonspatial functions, but they are attested sporadically (namely in twenty-four out of 163 of the attested case markers, which amounts to 15 percent). Furthermore, such overt combinations do not seem to have existed in proto-Tsezic at all. In all reconstructions made by maximum parsimony at least one of the parts is NULL. Because of this, we hypothesize that in proto-Tsezic only single overt case markers were used for nonspatial usage. One of the parts, either the locational or the directional, we reconstruct as NULL. As an example of the impact of this hypothesis, consider the reconstruction of the case for the addressee in 'explain [to somebody]' in Figure 8. The second part of the case marker is clearly reconstructible as *LAT. But for the first part both *NULL and *AT are equally parsimonious, as both need two changes. Given the proposal to ban overt combinations, however, we prefer *NULL.LAT over *AT.LAT.

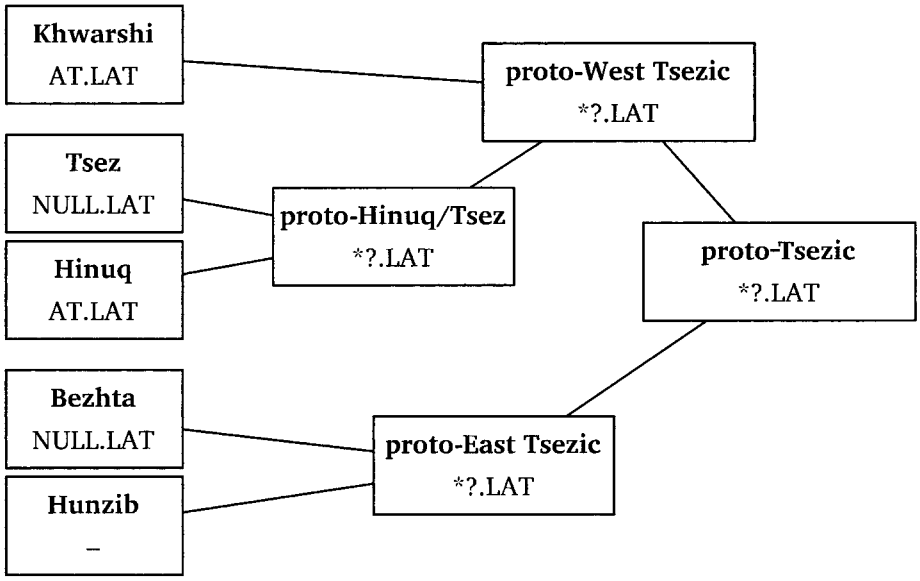

FIGURE 8. Reconstruction of case marking for the addressee of 'explain [to somebody]'.

Finally, all arguments in this article are based on the most common case marking for each function. In a few functions in some languages it is actually possible for other case markers to be used. Such alternative possibilities might help resolve reconstructions that are undecidable by maximum parsimony. In our current data, this argument is only needed to resolve the reconstruction of the case for the subject in the context of the construction AGE, as shown in Figure 9.[18] Proto-West Tsezic is clearly reconstructible as *AT.NULL, but this is completely incompatible with the case marking used in East Tsezic. To be precise, each of the reconstructions *AT.NULL, *NULL.GEN, and *NULL.LAT are equally parsimonious for proto-Tsezic, all requiring four changes. In

[18] All other cases in which there are alternative possibilities are already clearly reconstructible without this extra information. This attested variation also does not contradict the parsimony of any of the reconstructions proposed.

Hunzib, however, the case AT.NULL is also attested marking this function. Because of this alternative possibility, we reconstruct *AT.NULL for proto-Tsezic.

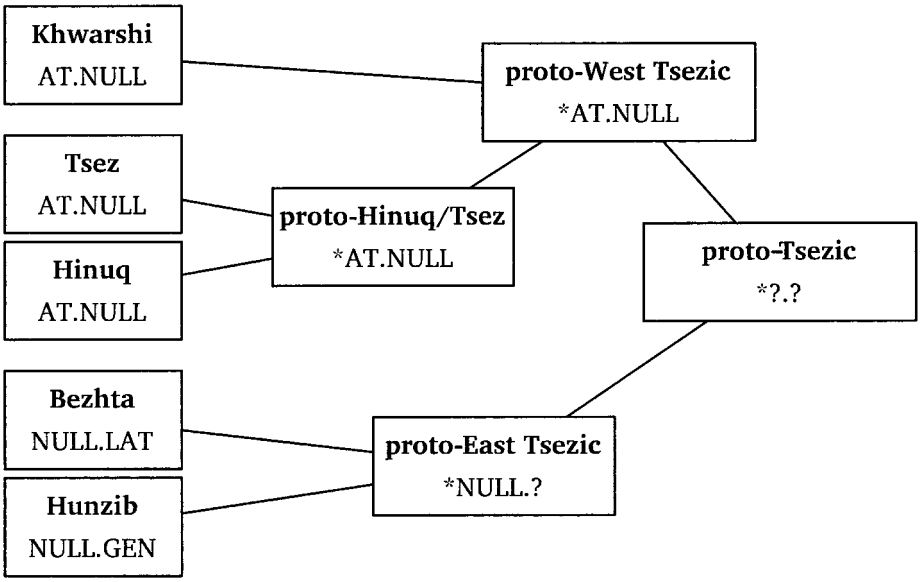

FIGURE 9. Reconstruction of case marking for the subject of the context AGE.

Even when including all of these additional factors, there still remain a few cases that are undecidable for proto-Tsezic. For example, the case for the marking of the addressee of 'teach [somebody]', shown in Figure 10, can be reconstructed by maximum parsimony as either *AT.NULL, *IN.NULL, or *AT.LAT (all of these options require four changes). Even if we discard the option *AT.LAT (because at least one part should be NULL in proto-Tsezic), this still leaves the options *AT.NULL and *IN.NULL. So, we can reconstruct the second part as NULL, but for the first part both AT and IN are equally likely.[19]

5. INTERPRETATION.

5.1. GRAPHICAL DISPLAY. To help humans make sense of large sets of data as presented in Appendices A and B, a good visualization can be of enormous help. In this section, we describe the visualization that we use in the next section to interpret the variation and reconstruction of the case marking in Tsezic. Basically, the visualization is a variation of the semantic-map approach (Haspelmath 2003), consisting of a base display using multidimensional scaling with a contour-line overlay using surface interpolation (cf. Cysouw 2008). It is of central importance to realize, however, that every visualization ignores some (smaller or greater) details present in the data. A visual display never replicates the underlying data one to one, but necessarily (and intentionally) reduces the attested variation in order to show general patterns. Given some caution, however, visualization is a great tool to help humans wade through large seas of data.

[19] From the semantic-map analysis as discussed in §5, it seems most likely that the proto-Tsezic case marker for 'teach [somebody]' was *AT.NULL (see specifically Fig. 14). However, this is an even further stretch of interpretation that we would like to leave undecided here.

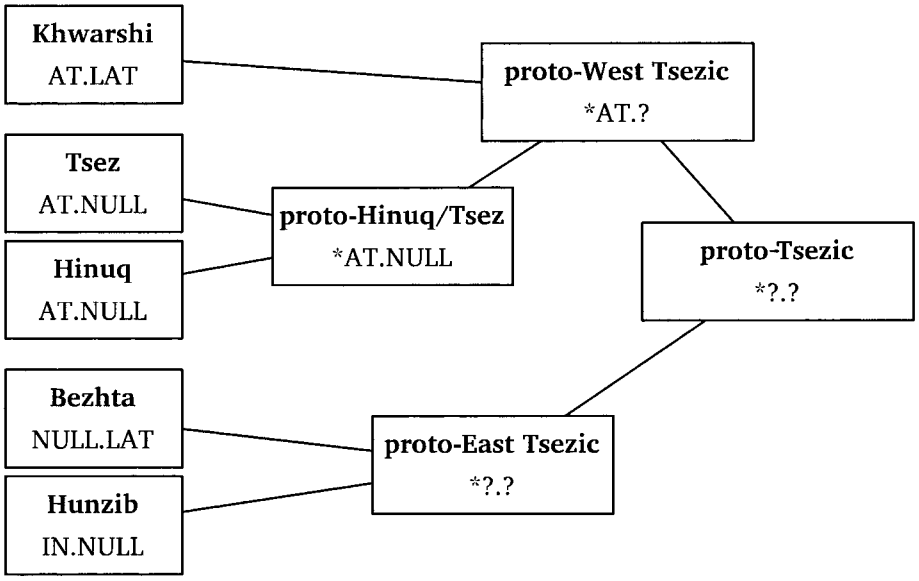

FIGURE 10. Reconstruction of case marking for the addressee of 'teach [somebody]'.

The first step of our visualization is to make a base display of the thirty-five functions that we have investigated. The functions could of course simply be listed alphabetically, but such an order would not help us in the interpretation of the variation. Instead, we emphasize in our display that some functions use similar cases throughout Tsezic, while others are mostly encoded differently. Basically, pairs of functions that are recurrently marked with the same case are shown closer together, while functions that mostly use different cases are shown far apart. In this way, the visual impression of the placement of the functions illustrates 'average' Tsezic case-marking similarity.

Technically, the method that we use to make such a display is multidimensional scaling. To do this, an 'average' Tsezic dissimilarity is computed for each pair of the thirty-five functions. As an example of such a computation, the functions EXPLAIN and TEACH are compared in Table 6. All case marking is split into a locational and a directional part, and each part in each language is compared between the two functions. The differences are summed together (there are three differences in the example) and divided by the number of comparisons made (there are eight comparisons made in the example because the data for Hunzib was not available to us). The dissimilarity between these two functions is then $3/8 \sim 0.38$. With thirty-five functions, there are $35 * 34/2 = 595$ such pairwise comparisons necessary.[20] Based on a table with all of these pairwise comparisons, multidimensional scaling attempts to locate the thirty-five functions in a two-dimensional display in such a way that the distance between each two functions in the display matches the computed dissimilarities as closely as possible. The resulting display is shown in Figure 11.[21]

[20] Actually, there are five pairs of functions that are coded completely identically in all five languages, so in practice there are only thirty different functions displayed on the map, and the number of pairwise dissimilarities is only $30 * 29/2 = 435$.

[21] For all multidimensional scaling in this article, we used the function 'cmdscale' from the statistical environment R (R Development Core Team 2008).

FUNCTION	HINUQ		TSEZ		KHWARSHI		BEZHTA		HUNZIB	
EXPLAIN	AT	LAT	NULL	LAT	AT	LAT	NULL	LAT	—	—
TEACH	AT	NULL	AT	NULL	AT	LAT	NULL	LAT	IN	NULL
differences	0	1	1	1	0	0	0	0	—	—

TABLE 6. Computing 'average' Tsezic dissimilarity between a pair of functions.

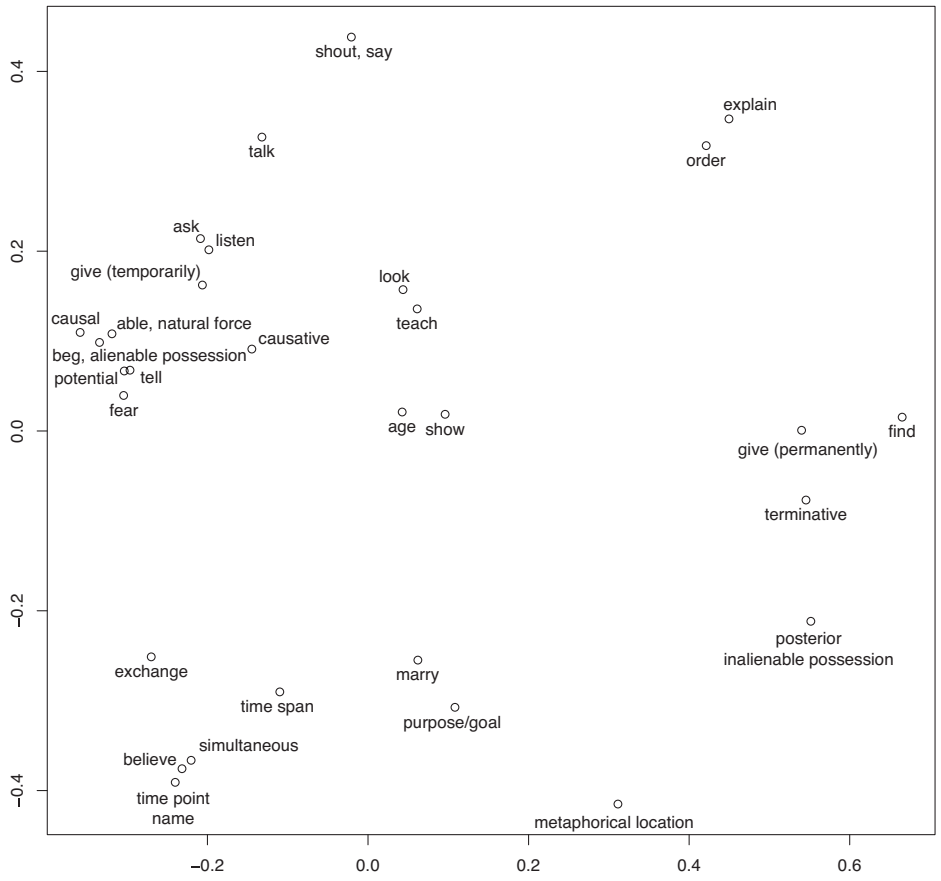

FIGURE 11. Multidimensional scaling of the thirty-five functions according to their case marking in Tsezic.

In most practical applications of multidimensional scaling (and here as well) it is not possible to precisely match the distances in the display to the computed distances. The display of a multidimensional scaling is almost always distorted to a certain extent. The extent to which the distances are correctly displayed is expressed by the explained variance, which is at 56.2 percent in Fig. 11. This is relatively low, indicating that the distances in the display are not really a good visualization of the underlying dissimilarities. In principle, there is no reason to restrict ourselves to only two dimensions. The multidimensional scaling provides various dimensions (viz. up to the number of functions minus one), ordered in diminishing importance. As can be seen from the statistics in Table 7, there are still substantial gains in the cumulative explained variance up to at least the fifth dimension. We use only the first two dimensions here, however, because they allow for an easier graphical display in print format. The restriction to consider only two dimensions is thus exclusively a practical consideration for visual presentation.

DIMENSION	VARIANCE EXPLAINED	CUMULATIVE VARIANCE EXPLAINED
1	35.4%	35.4%
2	20.8%	56.2%
3	12.9%	69.1%
4	8.8%	77.9%
5	7.8%	85.7%
6	3.8%	89.5%
. . .	. . .	. . .

TABLE 7. Explained variance for the first few dimensions of the multidimensional scaling.

On the basis of this two-dimensional display, we show the variation between the languages by means of lines surrounding functions that are encoded identically. This approach follows the tradition of drawing semantic maps (Haspelmath 2003). The distribution of Hinuq directional marking is shown in Figure 12 as an example of such a semantic map. The points represent the location of the functions according to the multidimensional scaling. We have left out the labels of the points in order to reduce clutter in these maps. To identify a function, please refer to the fully annotated base map in Fig. 11.

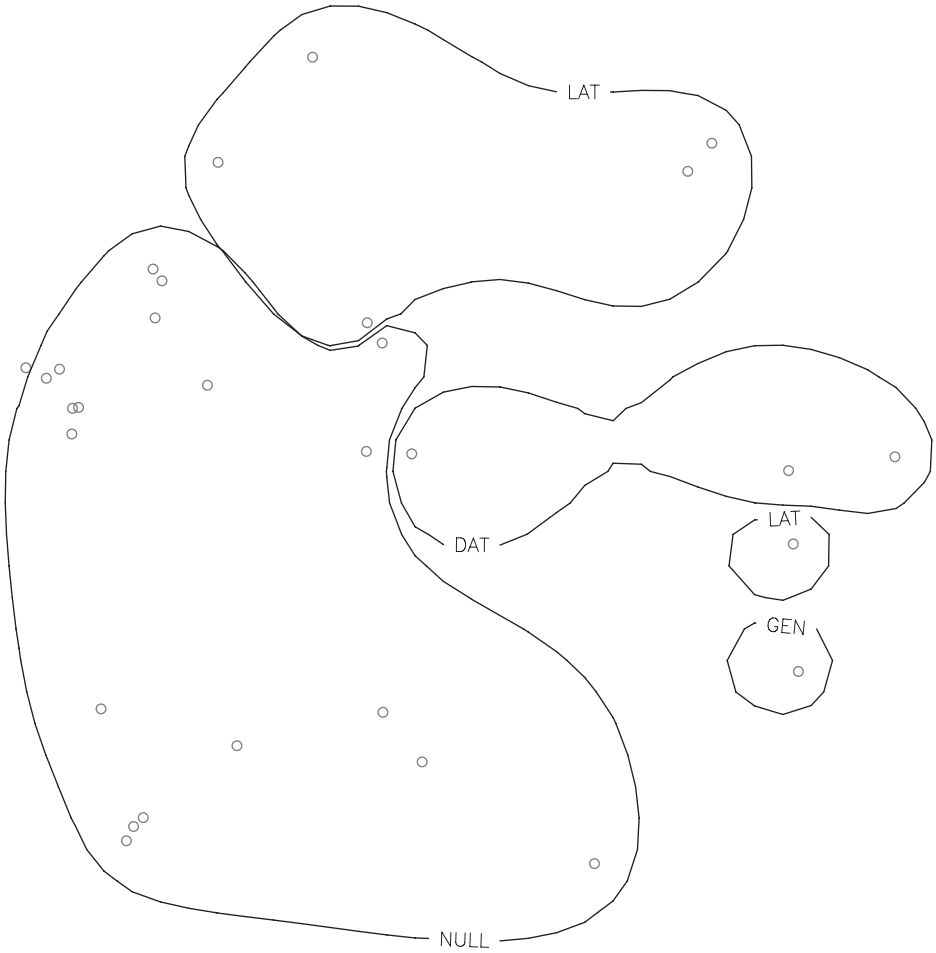

FIGURE 12. Directional marking in Hinuq.

The lines in this map (and all other maps in this article) are not drawn by hand, but automatically computed. In this way, we can ensure that the drawing of the lines follows the same rules in all maps, which is important because we later want to draw conclusions on the basis of a comparison of these displays. The basic method for drawing the lines is as follows. First, for each case marker, consider the functions that are marked by that marker. Assign all of these functions a height of one. All functions that are not marked by that case marker are placed at height zero.[22] Second, interpolate a curved surface ('hills and valleys') to match these elevations. Finally, draw a contour line (technically called 'isohypse') at height 0.5. This contour line then encloses all functions at a height of one.[23]

This approach to drawing semantic maps has a few consequences. First, functions with the same coding are not necessarily grouped within one contour line. For example, in Fig. 12 there are two different 'hills' for lative (LAT) as indicated by the two separate contour lines (one at the top and one mid-right). As we argue later (see §5.2), in this case this is probably the result of the introduction of the DATIVE (DAT) case, splitting an erstwhile continuous range of lative marking. But sometimes such apparent discontinuities are simply the result of the large distance between two functions, which under particular circumstances leads to a 'valley' in the interpolation that might be just lower than the contour line. Such a 'medium valley' will look like a discontinuous collection of functions. Also, the placement of the functions in the two-dimensional map is not always ideal (as argued above), possibly leading to discontinuities. The second effect of the present approach to drawing semantic maps is that sometimes the lines do not perfectly surround functions (see, for example, the left top of the NULL contour line in Fig. 12). In this example, this happens because we have no data on the case marking for this function in Hinuq. In other cases, however, such slight imperfections are simply a side effect of the details of the surface interpolation and the (arbitrary) height of the contour line. In general, not too much importance should be given to such details.

5.2. DIACHRONIC DEVELOPMENTS. Using the visualization as described in the previous section, all five individual languages and the four proto-stages can be mapped out and compared. We have split the presentation of these maps into two separate developments. First, there is a display, shown in Figure 13, of the developments of the second part of the case markers, the 'directional' part. Second, Figure 14 shows the diachronic development of the first part of the case marker, the 'locational' part. The thick one-sided arrows indicate the proposed diachronic developments. The thin two-sided arrows indicate only noncontiguous contours of the same case marker within a single language.

Before we start discussing the details of what possibly happened in the recent history of Tsezic nonspatial case marking, we would like to warn against an unwarranted interpretation. Glancing over the maps shown in Fig. 13 and Fig. 14, it is immediately obvious that the reconstructed stages look much 'smoother' than the attested languages. The maps for the actually attested languages have discontinuities, and there are more

[22] After some experimentation, we decided to place all functions on which we do not have data at a height of 0.3. In this way, information not available is classified neither as absence of case marking, nor as presence. But this is really just an ad hoc solution to deal with incomplete data.

[23] It is not trivial to make such a surface interpolation because the points of the base map are rather unequally distributed. To make such an interpolation, we used a geostatistical technique called 'kriging' as implemented by the function 'krige.conv' in the R-package 'geoR' (Ribeiro & Diggle 2001), with the parameter settings $s^2 = 1$ and $phi = 10$. On this basis, the isohypses were drawn using the 'contour' function.

strangely curly formed contours, and more separate groups. This apparent difference is purely an effect of the process of reconstruction, and it does not say anything about the difference between living languages and extinct languages. The reason for this difference is that it is just almost impossible to reconstruct idiosyncrasies. Reconstructions depend on shared variation, and special characteristics of a single language almost always disappear in the process of reconstruction, which results in more, 'cleaner' looking structures. We strongly believe, however, that earlier stages of today's languages were just as messy as living languages—we are just not able to show this empirically on the basis of the current data.

Investigating the changes in the directional part of the case marking in Fig. 13, we find that most of the functions that we have investigated were originally marked with a (zero) essive marker (NULL) in the proto-language, though its uses are reduced in most living languages. The ablative/genitive (GEN) marking hardly changes throughout Tsezic history (marking inalienable possession, metaphorical location, and the posterior converb). Only in Hinuq and Tsez is its usage slightly restricted. Furthermore, there are various idiosyncratic developments, like the occurrence of instrumental case marking in Khwarshi for TIME SPAN, an ablative in Tsez for METAPHORICAL LOCATION, an instrumental marker in Bezhta for CAUSATIVE, or a genitive in Hunzib for AGE.

The most interesting development is the spread of the lative. The (overt) lative marker (LAT) originally had only a restricted number of uses, namely: 'give permanently [to somebody]', 'explain [to somebody]', 'order [somebody]', '[somebody] finds X', and the terminative converb 'until'. These functions make sense, at least partly, as metaphorical extensions of the spatial meaning of the lative. The spatial meaning of the lative can be described roughly as marking the goal of a movement in which this goal is affected by the movement. Most clearly, this describes the meaning 'give permanently [to somebody]' in which the recipient is like an affected goal of the moving gift. Extended into the domain of communication, this also holds for the recipient of an explanation or an order. In general, recipient marking belongs to the most typical functions of the lative crosslinguistically, but addressee and terminative are also quite common uses of the lative (Rice & Kabata 2007). The fact that the one who finds something is also marked with the lative may be due to its similarity to a recipient.

In most Tsezic languages the distribution of LAT has been extended relative to this original distribution (with the exception being Hunzib). Going from proto-Tsezic to proto-West Tsezic, LAT extends across the top of the map, now including the addressees of some verbs of communication (SAY, SHOUT, but also LOOK). This extension is amplified in Khwarshi and proto-Hinuq/Tsez, and made even stronger in Tsez. The new uses extend the addressee marking to an increasing number of verbs of communication, and to objects of perception verbs. In Tsez the extension of LAT is also seen at the bottom of the map, marking PURPOSE/GOAL, one of the most typical functions of the lative crosslinguistically (Rice & Kabata 2007). This general tendency for LAT to extend its domain is counteracted in Hinuq by the introduction of a dative (DAT), which actually splits up the domain of lative and takes over some of its original uses (viz. FIND, GIVE (PERMANENTLY), and also SHOW).

In the eastern branch of the family, there is no change moving from proto-Tsezic to proto-East Tsezic. The reason for this is that Bezhta and Hunzib differ so strongly that their reconstructed shared ancestor has to be very close to proto-Tsezic (at least with regard to directional case marking). In Bezhta again we observe LAT extending its functions, though in a wholly different direction from what is seen in West Tsezic (viz.

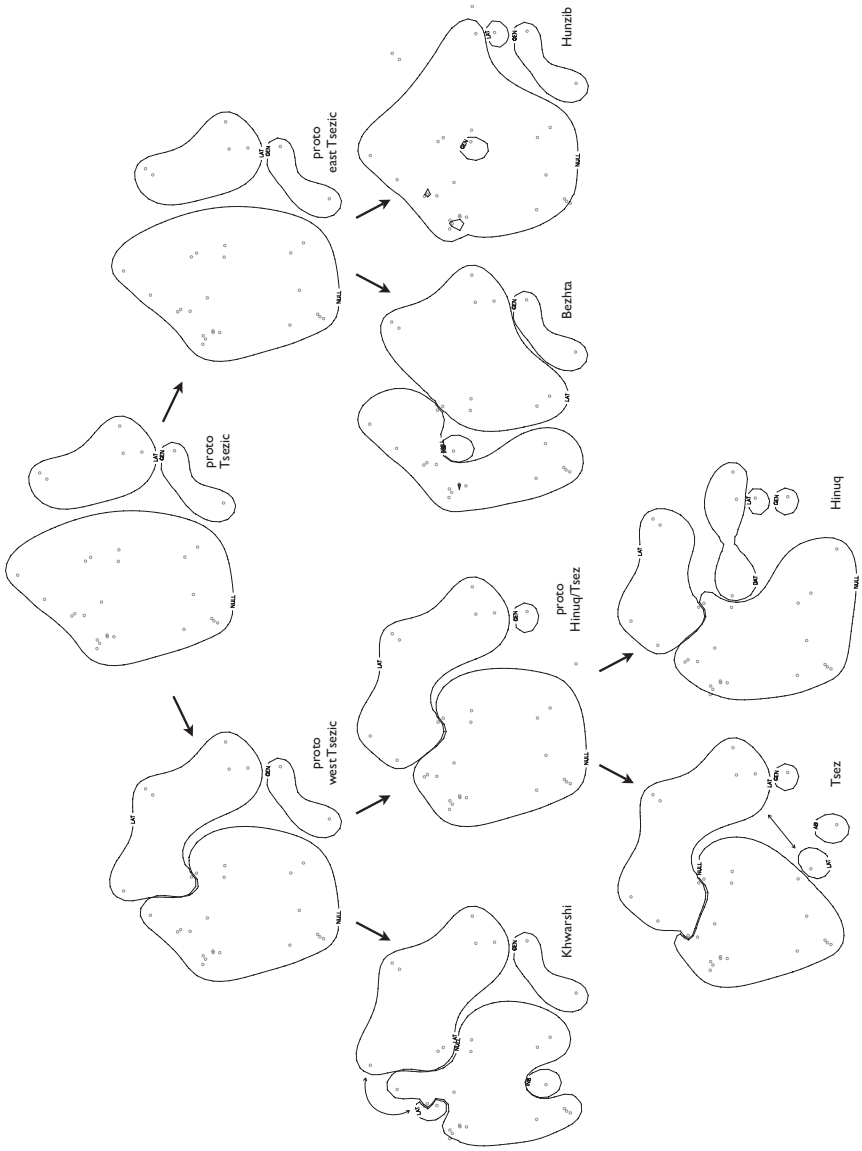

FIGURE 13. Development of directional case marking.

TEACH, SHOW, AGE, MARRY, PURPOSE/GOAL). Uniquely among the Tsezic languages, the developments in Hunzib take a completely different direction. Here the essive (NULL) extends almost across the complete map. The lative has completely disappeared, leaving just a frozen trace in the ending of the terminative converb.

The diachronic developments of the locational part of the case marking, shown in Fig. 14, show a similar trade-off between two large groups, namely AT and NULL. The uses of AT in proto-Tsezic can be divided into several clusters, marking (i) temporary possessors (viz. ALIENABLE POSSESSION, GIVE (TEMPORARILY), AGE), (ii) addressees (SHOUT, SAY, TELL, TALK, ASK), (iii) nonagentive subjects (CAUSATIVE, CAUSAL CONVERB, NATURAL FORCE, POTENTIAL, ABLE), and (iv) objects of perception (LISTEN, LOOK). First, the marking of temporary possessors with the spatial suffix AT is quite straightforwardly a metaphorical extension of the spatial usage. An object possessed by somebody can be conceptualized as just being located in the neighborhood of this person. Regarding the second group, the semantic parallels between recipient and addressee have been widely noted (e.g. Newman 1996). Members of the third group, nonagentive subjects (e.g. causees, involuntary agents, potential agents), are in many Daghestanian languages marked with a spatial marker that puts them in opposition to real agents, which are marked by the ergative. The Tsezic languages are not the only ones that use a recipient-like case for these functions. Other languages with a similar pattern are Akhwvakh, Karata, Godoberi, and Lak (Ganenkov 2006). Concerning the fourth group, we do not have any coherent account as to why objects of perception also end up being marked by AT. Finally, SHOW is also typically marked with AT. For this marking there are two explanations available. First, the verb 'show' is a derived causative verb in all Tsezic languages. Literally, it means 'cause somebody to see'. Thus the indirect object of SHOW is a kind of causee. Second, the indirect object of SHOW can be interpreted as a recipient or addressee.

The locational markers SUB, SPR, and CONT can also be reconstructed for proto-Tsezic. The use of CONT for TIME SPAN is straightforward considering that time can be easily conceptualized as a kind of mass. In contrast, SPR is used for various apparently unrelated functions in proto-Tsezic (TIME POINT, NAME, BELIEVE, PURPOSE/GOAL). Likewise, it is unclear to us why SUB is used for the object of EXCHANGE in proto-Tsezic. The metaphors underlying these extensions of spatial marking are either lost in history or were rather random in the first place.

The changes from proto-Tsezic to proto-West and proto-East Tsezic are small. First, in proto-West Tsezic the AD marker was introduced for MARRY probably just by expanding this marker's spatial meaning of nearness. There is no dedicated verb 'to marry' in the Tsezic languages; instead, various verbs like 'give', 'come', or 'go' are used with special case marking of the object. Also in proto-West Tsezic, SPR slightly extends its uses to include the case marking of SIMULTANEOUS CONVERB, which is related to the usage of SPR for TIME POINT. In the developments from proto-West Tsezic to Khwarshi and proto-Hinuq/Tsez, the usage of AT extends slightly. But more significant is that in Khwarshi the usage of AD extends into functions where Hinuq and Tsez use SPR (e.g. PURPOSE/GOAL).

Second, in proto-East Tsezic the IN marker is introduced for the SIMULTANEOUS CONVERB. In both Bezhta and Hunzib the usage of IN spreads to various other functions. This leads to a unique (and somewhat irregular looking) distribution of the location marker, especially in Hunzib. In contrast to the West Tsezic languages, the functions of AT are reduced in the eastern branch of the Tsezic tree.

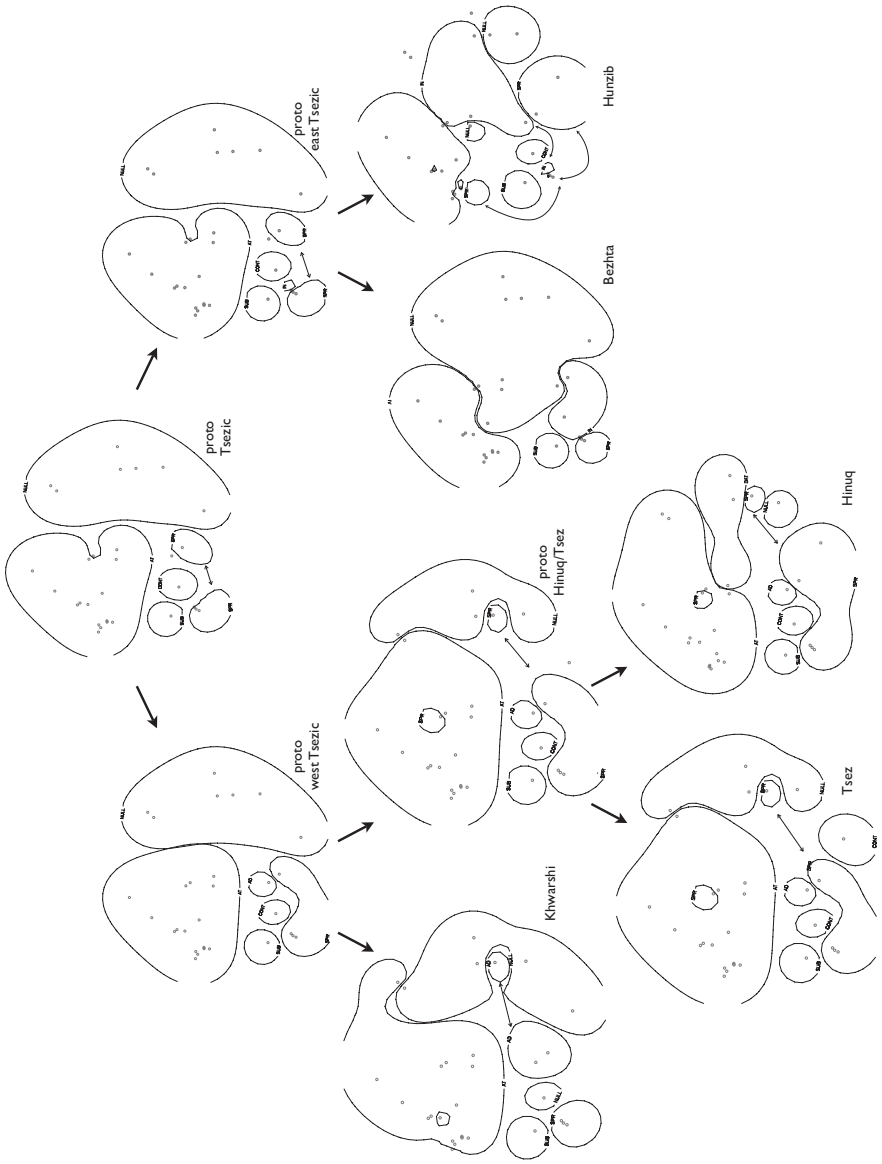

FIGURE 14. Development of locational case marking.

6. CONCLUSION. Based on the internal variation within the Tsezic language family, we have reconstructed the case marking for thirty-five nonspatial functions for proto-Tsezic. We were able to reconstruct the case marking for all but three of these functions in proto-Tsezic. For the remaining three it was impossible to determine which of the available options would be the best reconstruction. A summary of our proposed reconstructions is shown in Figure 15. The dots in this figure indicate the thirty-five functions (see Fig. 11 for details). The lines indicate which marker we reconstruct for the enclosed functions (see §3.1 for explanation of the labels).

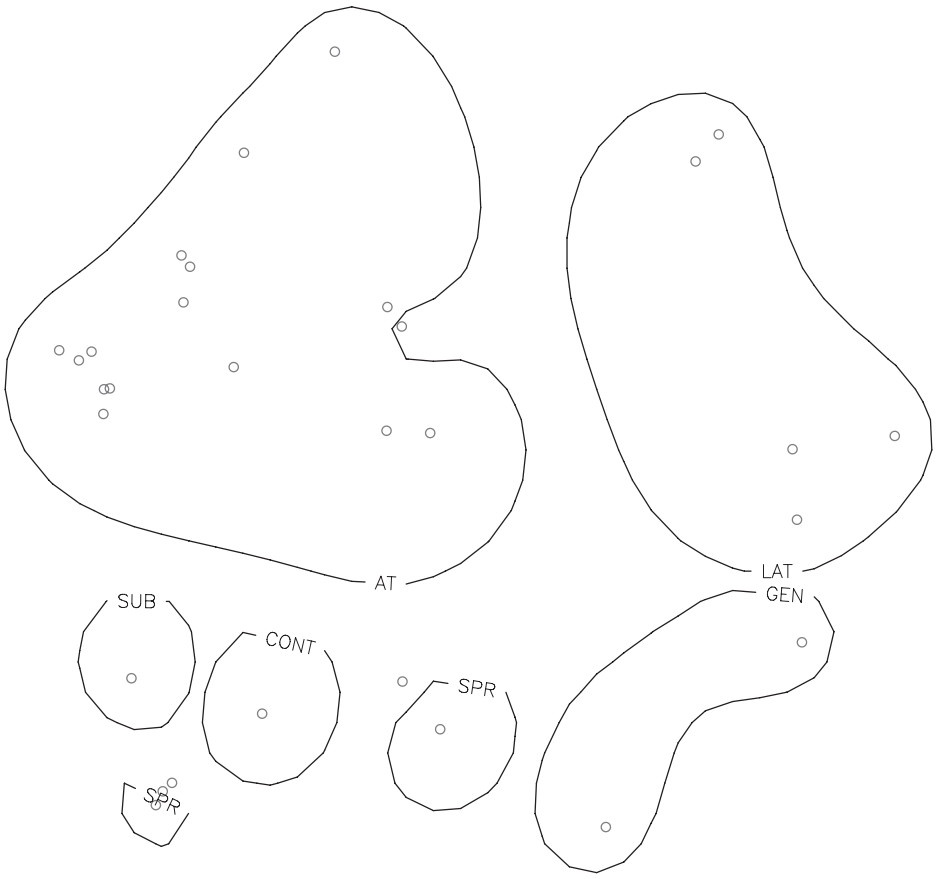

FIGURE 15. Summary of nonspatial case marking for proto-Tsezic.

Specifically, we have argued that nonspatial case marking in proto-Tsezic was monomorphemic. This is significant because spatial case marking is mostly bimorphemic in Tsezic, and also various nonspatial case markers in individual Tsezic languages are bimorphemic, consisting of a locational and a directional part. Our reconstruction suggests, however, that in the original (pre-proto-Tsezic) transfer of spatial case markers to nonspatial functions only one of the two spatial morphemes was used (either the locational part or the directional part).

The lative marker (LAT) was metaphorically extended from its original meaning 'goal of movement' to affected recipients and terminative converbs. The general loca-

tion marker (AT) was extended to meanings of temporary possession, less affected recipients, and addressees. Various other case markers were transferred to nonspatial contexts in proto-Tsezic (CONT, SUB, SPR, GEN), though the semantic or functional rationale is not always completely clear for these developments. From this starting point as reconstructed for proto-Tsezic, different individual developments led to the case-marking structures of the presently living Tsezic languages. Although some vestiges of the original pattern are often still discernible, the idiosyncratic developments mostly obscured general tendencies that originally shaped the system.

With this exercise in reconstructing morphosyntactic function of case markers, we also want to suggest that much more can be concluded from the minutiae of family-internal morphosyntactic variation. The sometimes mind-boggling small differences between closely related languages offer a unique possibility to learn more about the historical processes that shaped these languages. This insight is of course widespread in the historical-comparative reconstruction of the lexicon and of sound patterns, but we think that the same basic approach (i.e. using the minutiae of variation to reconstruct history) can also be fruitfully used to reconstruct morphosyntax. And, in turn, such morphosyntactic reconstructions can be used as independent arguments for the reconstruction of the genealogical structure of a group of languages (cf. Fig. 6).

The approach that we have used in this article can without adaptation be applied to other languages and other morphosyntactic domains. The only necessary premise is that the morphemes used in the various languages are known to be cognate. In our example, we started from the known Tsezic cognacy of the case markers (see Tables 2 and 4). The different distribution of the same morpheme (historically speaking) in different languages can then be interpreted as a result of historical processes. This approach is thus not applicable to discovering previously unknown genealogical relationships, and it is not suitable for the purely typological comparison of functional domains (i.e. comparison without identity of form across languages). Yet the method as presented in this article offers the prospect for unprecedented detailed investigation of the diachronic developments within groups of related languages.

Appendix A: Nonspatial case marking in tsezic

Dashes indicate either that the data is missing, or that a completely different construction is used, not involving spatial case marking.

	HINUQ		TSEZ		KHWARSHI		BEZHTA		HUNZIB	
FUNCTION	LOC	DIR	LOC	DIR	LOC	DIR	LOC	DIR	LOC	DIR
ABLE	AT	NULL	AT	NULL	AT	NULL	AT	NULL	AT	NULL
AGE	AT	NULL	AT	NULL	AT	NULL	NULL	LAT	NULL	GEN
ALIENABLE POSSESSION	AT	NULL	AT	NULL	AT	NULL	AT	NULL	—	—
ASK	AT	NULL	AT	NULL	AT	LAT	AT	NULL	AT	NULL
BEG	AT	NULL	AT	NULL	AT	NULL	AT	NULL	—	—
BELIEVE	SPR	NULL	SPR	NULL	SPR	NULL	SPR	NULL	SPR	NULL
CAUSAL	—	—	AT	NULL	—	—	AT	NULL	AT	NULL
CAUSATIVE	AT	NULL	AT	NULL	AT	NULL	NULL	INSTR	AT	NULL
EXCHANGE	SUB	NULL	SUB	NULL	SUB	NULL	SUB	NULL	SUB	NULL
EXPLAIN	AT	LAT	NULL	LAT	AT	LAT	NULL	LAT	—	—
FEAR	AT	NULL	AT	NULL	AT	NULL	AT	NULL	SPR	NULL
FIND	DAT	DAT	NULL	LAT	NULL	LAT	NULL	LAT	—	—
GIVE (PERMANENTLY)	DAT	DAT	NULL	LAT	NULL	LAT	NULL	LAT	IN	NULL
GIVE (TEMPORARILY)	AT	NULL	AT	NULL	NEAR	LAT	AT	NULL	AT	NULL

INALIENABLE POSSESSION	NULL	GEN	NULL	GEN	NULL	GEN	NULL	GEN	NULL	GEN
LISTEN	AT	NULL	AT	LAT	AT	NULL	AT	NULL	—	—
LOOK	SPR	LAT	SPR	LAT	AT	LAT	AT	NULL	AT	NULL
MARRY	AD	NULL	AD	NULL	AD	NULL	NULL	LAT	IN	NULL
METAPHORICAL LOCATION	SPR	NULL	CONT	ABL	NULL	GEN	NULL	GEN	SPR	GEN
NAME	SPR	NULL	SPR	NULL	SPR	NULL	SPR	NULL	—	—
NATURAL FORCE	AT	NULL	AT	NULL	AT	NULL	AT	NULL	AT	NULL
ORDER	AT	LAT	AT	LAT	NULL	LAT	NULL	LAT	—	—
POSTERIOR	NULL	GEN	NULL	GEN	NULL	GEN	NULL	GEN	NULL	GEN
POTENTIAL	AT	NULL	AT	NULL	AT	NULL	—	—	—	—
PURPOSE/GOAL	SPR	NULL	SPR	LAT	AD	NULL	IN	LAT	SPR	NULL
SAY	AT	LAT	AT	LAT	AT	LAT	AT	NULL	AT	NULL
SHOUT	AT	LAT	AT	LAT	AT	LAT	AT	NULL	AT	NULL
SHOW	DAT	DAT	AT	NULL	AT	NULL	NULL	LAT	IN	NULL
SIMULTANEOUS	SPR	NULL	SPR	NULL	SPR	NULL	IN	NULL	IN	NULL
TALK	AT	LAT	AT	LAT	AT	NULL	AT	NULL	AT	NULL
TEACH	AT	NULL	AT	NULL	AT	LAT	NULL	LAT	IN	NULL
TELL	AT	NULL	AT	NULL	AT	NULL	AT	NULL	IN	NULL
TERMINATIVE	SPR	LAT	SPR	LAT	AD	LAT	NULL	LAT	NULL	LAT
TIME POINT	SPR	NULL	SPR	NULL	SPR	NULL	SPR	NULL	—	—
TIME SPAN	CONT	NULL	CONT	NULL	NULL	INSTR	IN	NULL	CONT	NULL

APPENDIX B: RECONSTRUCTIONS OF NONSPATIAL CASE MARKING

Inferences using maximum parsimony (§4.1) are shown in regular typeface. Additional ad hoc inferences (§4.2) are shown with an italic typeface.

	PROTO-HINUQ/TSEZ		PROTO-WEST TSEZIC		PROTO-EAST TSEZIC		PROTO-TSEZIC	
FUNCTION	LOC	DIR	LOC	DIR	LOC	DIR	LOC	DIR
ABLE	AT	NULL	AT	NULL	AT	NULL	AT	NULL
AGE	AT	NULL	AT	NULL	*AT*	*NULL*	*AT*	*NULL*
ALIENABLE POSSESSION	AT	NULL	AT	NULL	AT	NULL	AT	NULL
ASK	AT	NULL	AT	NULL	AT	NULL	AT	NULL
BEG	AT	NULL	AT	NULL	AT	NULL	AT	NULL
BELIEVE	SPR	NULL	SPR	NULL	SPR	NULL	SPR	NULL
CAUSAL	AT	NULL	AT	NULL	AT	NULL	AT	NULL
CAUSATIVE	AT	NULL	AT	NULL	AT	NULL	AT	NULL
EXCHANGE	SUB	NULL	SUB	NULL	SUB	NULL	SUB	NULL
EXPLAIN	*NULL*	LAT	*NULL*	LAT	*NULL*	LAT	*NULL*	LAT
FEAR	AT	NULL	AT	NULL	AT	NULL	AT	NULL
FIND	NULL	LAT	NULL	LAT	NULL	LAT	NULL	LAT
GIVE (PERMANENTLY)	NULL	LAT	NULL	LAT	NULL	LAT	NULL	LAT
GIVE (TEMPORARILY)	AT	NULL	AT	NULL	AT	NULL	AT	NULL
INALIENABLE POSSESSION	NULL	GEN	NULL	GEN	NULL	GEN	NULL	GEN
LISTEN	AT	NULL	AT	NULL	AT	NULL	AT	NULL
LOOK	SPR	LAT	AT	LAT	AT	NULL	AT	*NULL*
MARRY	AD	NULL	AD	NULL	—	NULL	—	NULL
METAPHORICAL LOCATION	—	—	*NULL*	GEN	*NULL*	GEN	*NULL*	GEN
NAME	SPR	NULL	SPR	NULL	SPR	NULL	SPR	NULL
NATURAL FORCE	AT	NULL	AT	NULL	AT	NULL	AT	NULL
ORDER	AT	LAT	NULL	LAT	NULL	LAT	NULL	LAT
POSTERIOR	NULL	GEN	NULL	GEN	NULL	GEN	NULL	GEN
POTENTIAL	AT	NULL	AT	NULL	—	—	AT	NULL
PURPOSE/GOAL	SPR	NULL	SPR	NULL	SPR	NULL	SPR	NULL
SAY	AT	LAT	AT	LAT	AT	NULL	AT	*NULL*
SHOUT	AT	LAT	AT	LAT	AT	NULL	AT	*NULL*
SHOW	*AT*	NULL	*AT*	NULL	*AT*	NULL	*AT*	NULL

SIMULTANEOUS	SPR	NULL	SPR	NULL	IN	NULL	—	NULL
TALK	AT	LAT	AT	NULL	AT	NULL	AT	NULL
TEACH	AT	NULL	AT	*NULL*	—	*NULL*	—	*NULL*
TELL	AT	NULL	AT	NULL	AT	NULL	AT	NULL
TERMINATIVE	SPR	LAT	*NULL*	LAT	*NULL*	LAT	NULL	LAT
TIME POINT	SPR	NULL	SPR	NULL	SPR	NULL	SPR	NULL
TIME SPAN	CONT	NULL	CONT	NULL	CONT	NULL	CONT	NULL

REFERENCES

ABDULAEV, ARSEN. In press. *Cezjas folklor*. Maxačkala:.

ALEKSEEV, MIXAIL E. 1999. Cezkie jazyki. *Jazyki mira: Kavkazskie jazyki*, ed. by Mixail E. Alekseev, 299–303. Moscow: Academia.

ALEKSEEV, MIXAIL E. 2003. *Sravnitel'no-istoričeskaja morfologija avaro-andijskix jazykov*. Moscow: Nauka.

BANDELT, HANS-JORGEN, and ANDREAS W. M. DRESS. 1992. Split decomposition: A new and useful approach to phylogenetic analysis of distance data. *Molecular Phylogenetics and Evolution* 1.242–52.

BOKAREV, EVGENIJ A. 1959. *Cezskie (didojskie) jazyki Dagestana*. Moscow: AN SSSR.

COMRIE, BERNARD. 1999. Spatial cases in Daghestanian languages. *Sprachtypologie und Universalienforschung* 52.108–17.

COMRIE, BERNARD. 2004. Oblique-case subjects in Tsez. *Non-nominative subjects*, ed. by Peri Bhaskararao and Karumuri Venkata Subbarao, 113–27. Amsterdam: John Benjamins.

COMRIE, BERNARD, and MARIA POLINSKY. 1998. The great Daghestanian case hoax. *Case, typology and grammar*, ed. by Anna Siewierska and Jae Jung Song, 95–114. Amsterdam: John Benjamins.

COMRIE, BERNARD, and HELMA VAN DEN BERG. 2006. Experiencer constructions in Daghestanian languages. *Semantic role universals and argument linking: Theoretical, typological, and psycholinguistic perspectives*, ed. by Ina Bornkessel, Matthias Schlesewsky, and Bernard Comrie, 127–54. Berlin: Mouton de Gruyter.

CYSOUW, MICHAEL. 2008. Generalizing scales. *Scales* (Linguistische Arbeits Berichte 86), ed. by Marc Richards and Andrej Malchukov, 379–96. Leipzig: Institut für Linguistik, Universität Leipzig.

DANIEL, MICHAEL; ZAIRA XALILOVA; and ZARINA MOLOCHIEVA. 2007. Ditransitives in Nakh-Daghestanian. Paper presented at the Conference on Ditransitive Constructions, Leipzig.

DUNN, MICHAEL; ANGELA TERRILL; GER REESINK; ROBERT A. FOLEY; and STEPHEN C. LEVINSON. 2005. Structural phylogenetics and the reconstruction of ancient language history. *Science* 309.2072–75.

FELSENSTEIN, JOSEPH. 2004. *Inferring phylogenies*. Sunderland, MA: Sinauer.

GANENKOV, DIMITRY. 2005. *Kontaktnyje lokalizacii v naxsko-dagestanskix jazykax*. Moscow: University of Moscow dissertation.

GANENKOV, DIMITRY. 2006. Experiencer coding in Nakh-Daghestanian. *Case, valency and transitivity*, ed. by Leonid Kulikov, Andrej Malchukov, and Peter de Swart, 179–202. Amsterdam: John Benjamins.

GRAY, RUSSELL D., and QUENTIN D. ATKINSON. 2003. Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature* 426.435–39.

HARRIS, ALICE C., and LYLE CAMPBELL. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.

HASPELMATH, MARTIN. 1995. The converb as a cross-linguistically valid category. *Converbs in cross-linguistic perspective*, ed. by Martin Haspelmath and Ekkehard König, 1–55. Berlin: Mouton de Gruyter.

HASPELMATH, MARTIN. 2003. The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison. *The new psychology of language: Cognitive and functional approaches to language structure*, vol. 2, ed. by Michael Tomasello, 211–42. Mahwah, NJ: Lawrence Erlbaum.

HEERINGA, WILBERT. 2004. *Measuring dialect pronunciation differences using Levenshtein distance*. Groningen: Rijksuniversiteit Groningen dissertation.

HEERINGA, WILBERT; PETER KLEIWEG; CHARLOTTE GOOSKENS; and JOHN NERBONNE. 2006. Evaluation of string distance algorithms for dialectology. *Proceedings of the Workshop on Linguistic Distances*, 51–62. Sydney: Association for Computational Linguistics.
HUSON, DANIEL H., and DAVID BRYANT. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* 23.254–67.
KIBRIK, ALEKSANDR E., and YAKOV G. TESTELETS. 2004. Bezhta. *The indigenous languages of the Caucasus*, ed. by Michael Job, 217–98. New York: Caravan.
KORJAKOV, J. B. 2006. *Atlas kavkazskix jazykov*. Moscow: Pilgrim.
LEVENSHTEIN, V. I. 1966. Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady* 10.707–10.
LEVINSON, STEPHEN C.; SÉRGIO MEIRA; and THE LANGUAGE AND COGNITION GROUP. 2003. 'Natural concepts' in the spatial topological domain—adpositional meanings in cross-linguistic perspective: An exercise in semantic typology. *Language* 79.485–516.
NEWMAN, JOHN. 1996. *Give: A cognitive linguistic study*. Berlin: Mouton.
NIKOLAEV, SERGEY L., and SERGEY A. STAROSTIN. 1994. *A North Caucasian etymological dictionary*. Moscow: Asterisk.
R DEVELOPMENT CORE TEAM. 2008. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
RADŽABOV, RAMAZAN. 1999. *Syntaksis csezkogo jazyka*. Moskow: Moskovskij gosudarstvennyj lingvisti eskij universitet, Centr jazykov i kultur Severnoj Evrazii.
RIBEIRO, PAULO JUSTINIANO, JR., and PETER J. DIGGLE. 2001. geoR: A package for geostatistical analysis. *R News* 1.15–18.
RICE, SALLY, and KAORI KABATA. 2007. Crosslinguistic grammaticalization patterns of the ALLATIVE. *Linguistic Typology* 11.451–514.
SAITOU, NARUYA, and MASATOSHI NEI. 1987. The neighbour-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4.406–25.
TRIER, JOST. 1931. *Der deutsche Wortschatz im Sinnbezirk des Verstandes*. Heidelberg: Winter.
VAN DEN BERG, HELMA. 1995. *A grammar of Hunzib (with texts and lexicon)*. Munich: LINCOM Europa.
XALILOV, MADŽID Š. 1995. *Bežtinsko-Rysskij slovar'*. Maxačkala: Institut JaLI DNC RAN.
XALILOV, MADŽID Š. 1999. *Cezsko-Rysskij slovar'*. Maxačkala: Institut JaLI DNC RAN.
XALILOV, MADŽID Š., and I. A. ISAAKOV. 2001. *Gunzibsko-Russkij slovar'*. Maxačkala: Institut JaLI DNC RAN.
XALILOV, MADŽID Š., and I. A. ISAAKOV. 2005. *Ginuxsko-Russkij slovar'*. Maxačkala: Institut JaLI DNC RAN.

Max Planck Institute for Evolutionary Anthropology
Deutscher Platz 6
04103 Leipzig
Germany
[cysow@eva.mpg.de]
[forker@eva.mpg.de]

[Received 20 November 2008; accepted 9 April 2009]