Article

The widespread converb suffix *z

a relique morpheme of Nostratic dating

Arnaud Fournet

Abstract: The paper deals with a particular verbal form that will be called *converb*. It begins with a definition of converbs. Then the paper attempts a comparative survey of the suffixal material that exists in a number of languages and that is used for one of the converbial constructions, which expresses simultaneity. The survey does not aim at being exhaustive but it will be shown that a whole set of languages belonging to different linguistic families throughout Eurasia that are not necessarily considered to be genetically related do share the same morpheme for converbs. This situation clearly raises the issue of these languages being ultimately related and this converb suffix being a relique morpheme of very high dating.

Keywords: Converb, Nostratic, Macro-Comparison, Eurasia.

1. *Introduction and purposes*

The paper deals with a particular verbal form that will be called *converb*. This verbal form needs to be distinguished from other nonfinite derivatives and for the sake of clarity it will be best illustrated with examples. Hungarian fulfills this purpose:

(1) Example of Hungarian *Finite Verb*:

A lány (az[1] ember) sír.

[1]a [2]lány ([1]az [3]ember) [4]sír
[1]Definite Article [2]'girl' ([3]'man') [4]'to cry'
'The girl (the man) cries (is crying).'

(2) Example of Hungarian *Infinitive*, in that syntactical construction the relationship between the Verb and the subject is intermediated by another unit:

A lány (az ember) sírni kezdett.

[1]a [2]lány ([1]az [3]ember) [4]sír-[5]ni [6]kezd-[7]ett
[5]Infinitive [6]'to begin'-[7]Past.P3sg
'The girl (the man) began to cry.'

(3) Example of Hungarian *Present Participle*, in that syntactical construction the Verb behaves as as an Adjectival describer independent from the rest of the sentence:

A síró lány (ember) bejött[2] a szobába.

---

[1] The Definite Article is <az>, instead of <a>, when followed by a noun with an initial vowel.

[1]a [4]sír-[8]ó [2]lány ([3]ember) [9]be-[10]jö-[7]tt [1]a [11]szobá-[12]ba
[8]Present Participle [9]'into' [10]'to enter' [11]'room' [12]Illative
'The crying girl (man) entered the room.'

(4) Example of Hungarian *Converb*, in that syntactical construction the Verb is a circumstancial describer, independently cooccuring with the main event:

A lány (az ember) sírva bejött a szobába.

[1]a [2]lány ([1]az [3]ember) [4]sír-[13]va [9]be-[10]jö-[7]tt [1]a [11]szobá-[12]ba
[13]Converb
'The girl (the man) entered the room crying.'

(5) Example of Hungarian *Nomen Actionis* or *Nominalized verb*:

A lány (az ember) sírása ingerel engem.

[1]a [2]lány ([1]az [3]ember) [4]sír-[14]ás-[15]a [16]ingerel [17]engem
[14]Nominalizer [15]P3sg Genitive [16]'to irritate' [17]'P1sg Accusative'
'The crying of the girl (of the man) irritates me.'

The term *converb* originates in the Altaicist grammatical tradition where it has been introduced as a substitute for other terms like *adverbial participle*, *conjunctive participle*, *gerund* or *gerundive* (French *gérondif*) to describe a number of non-finite verbal forms. In the Uralicist tradition converbs have sometimes been called *verbal adverbs* (in German *Verbaladverb*), as in Wiedemann (1884: 176-179). My purpose in this paper is not to discuss the legitimacy of this neologism, nor is it to discuss whether *converbs* should be considered *verbal adverbs*, as in Haspelmath (1994: 153; 1995a: 3-4; 1996: 50) or van der Auwera (1998b: 276), nor whether *converbs* are not longer verbal forms, in spite of sharing apparent roots with verbs. But see examples (10) and (13).

The purpose of the paper is a comparative survey of the suffixal material that exists in a number of languages and is used to express the *converbial* construction, especially the one for simultaneously occuring verbs. The present survey does not aim at being exhaustive but it will be shown that a whole set of languages belonging to different linguistic families throughout Eurasia that are not necessarily considered to be genetically related do share the same morpheme for this converb. After the description of the individual languages a cladistic branching of Northern Eurasian languages and families is established and then compared with the points of view about the genetic relationships usually presented or debated.

2. *Linguistic survey of Northern Eurasian converb formatives*

As much as it is possible languages are described from West to East, starting with Basque.

2.1 *Basque (usually considered isolated)*

Basque is usually considered to be a linguistic isolate. For example a former specialist of that language, Larry Trask, once asserted: "Basque is a genetically isolated language: there is not the slightest shred of evidence that it is related to any other living language." (1997:35).

As regards *converbial* constructions Basque has the following morpheme <(e)z>, phonetically [(e)s], suffixed to the verb root:

---

[2] *jött be* is also acceptable. *Be* as a preverbal formative is obviously the same etymological "morpheme" as the (pro-)nominal Illative case marker -*ba*, used in different morphological and segmental environments. They will be dealt with as two separate units as their distributional properties are different.

(6) *Eri izanez ez etorri du*. (simplified from Lafitte 1980:228)

[1]eri [2]izan-[3]ez [4]ez [5]etorri [6]du

[1]'ill' [2]'to be'-[3]Converb [4]'not' [5]'to come' [6]Auxiliary-Present-P3sg
'Being ill he will/does not come.'

It can be noted that Basque accepts to some extent - with the help of words like *gero*(z) 'as', *geroztik* 'since'- that the grammatical subject of the converb may not be the same as that of the main verb:

(7) *Aitak erranez geroz ez etorri du*. (adapted from Lafitte 1980:228-9)

[7]aita-[8]k [9]erran-[3]ez [10]geroz [4]ez [5]etorri [6]du
[7]'father' [8]'P2sg' [9]'to say' [10]'as, because'
'As your father affirms it he will/does not come.'

So-called strong verbs like *izan* 'to be' and *erran* 'to say' directly affix *-ez*, the other verbs, the majority of them in fact, affix *-tu-z*: *beha-tuz* '(while) seeing', *handi-tuz* '(while) growing', etc. So far Basque has never been proposed to be a member of Nostratic.

2.2 *Etruscan (possibly Nostratic)*

As is well known Etruscan was the language of the Etruscan people who once inhabited the area in northwest central Italy between the Arno, the Tiber and the Tyrrhenian Sea. It was written in a set of alphabets derived from a Greek prototype and there is therefore little apparent difficulty in reading the language although it is considerably more difficult to understand it. Written Etruscan is attested from 700 BCE to AD 50. Because it is now a dead language Etruscan remains incompletely understood and interpretating of Etruscan inscriptions and texts remains conjectural to some extent. In addition a few Etruscan texts come from other areas of Italy, especially from Campania, Emilia and from Corsica, and isolated examples are known from Provence, Tunisia, Greece and Egypt.

Ultimately Etruscan is primarily known thanks to inscriptions. They number about 9000 according to Rix (2008: 141) and more than 13 000 according to Bonfante (1994: 437), depending on what is called an "inscription". We are lucky enough to have one instance of *converb*, which is formed with the suffix *-asa*:

(8) *Clenar ci acanasa elsśi zilachnu*. (TLE 169) (Bonfante-Bonfante 2002:103)

[1]clen-[2]ar [3]ci [4]acan-[5]asa [6]els-[7]śi [8]zilachnu
[1]'son'-[2]Plural [3]'three' [4]'to procreate'-[5]Converb [6]'two'-[7]'cardinal' [8]'praetor'
'Having had three sons [he was] twice praetor.'

Etruscan has been increasingly considered to be some relative of Proto-Indo-European in view of a number of lexical and grammatical features. The inclusion of Etruscan into the Nostratic framework is accepted in Bomhard (2008), even though he does not deal with the present converb formative.

2.3 *Lezgi (a North-Eastern Caucasic language)*

Lezgi is spoken in southern Daghestan and northern Azerbaidzhan in earthern Caucasic. In his monograph on Lezgi grammar Haspelmath (1993:1) accepts the genetic affiliation of Lezgi within an North-Eastern Caucasic group, that can also be called the Nakho-Daghestanian family. He further emphasizes that this group should not be confused with Kartvelian or Abkhaz-Adighean, this latter being also known as North-Western Caucasic. Lezghi provides examples of the converb formative *-z*:

(9) *Cükweraz sa žehil q̄wez akuna*. (simplified from Haspelmath 1993:370)

[1]Cükwer-[2]az [3]sa [4]žehil [5]q̄we-[6]z [7]aku-[8]na
[1]'Cükwer'-[2]Dative [3]'one' [4]'young man' [5]'to come'-[6]Converb [7]'to see'-[8]Aorist
'Cükwer saw a young man coming.'

I have followed the original parsing and terminology. The converb suffix is *-z* with strong verbs and *-iz* with weak verbs (Haspelmath 1993:130). A more complex example is this one:

(10) *Gah qʰwer awaz gah q'il galtadiz raxana*. (simplified from Haspelmath 1993:333)

[9]gah [10]qʰwer [11]awa-[6]z [9]gah [12]q'il [13]galtad-[6]iz [14]raxa-[8]na
[9]'now' [10]'to smile' [11]'to be in' [12]'head' [13]'to shake' [14]'to talk'
'He was talking, now with a smile now shaking his head.'

Haspelmath (1993) also indicates that the converb *-(i)z* can be combined with other morphemes to produce units like *z(a)maz*, *n(a)maz*, *amaz*, etc. Incidentally it can be noted that in (10) converbs have grammatical objects, which precludes that they could be considered to be adverbs. Some people have proposed that Northern Caucasic could be related to Indo-European. This is mostly the case of Colarusso. Bomhard (2008:530) considers that the issue is more about areal contacts than genetic relationship: "Colarusso has not sufficiently demonstrated that there was a genetic relationship between Proto-Indo-European and Northwest Caucasian. What Colarusso has shown, however, and quite convincingly, is that there was contact between Proto-Indo-European and Northwest Caucasian."

2.4 *Georgian (a Kartvelian language)*

Georgian belongs to the so-called Southern Caucasic branch, also known as Kartvelian. In spite of being generally held to be Nostratic, it does not provide any example of the converb formative *-z*. The numerous participles of Kartvelian do not seem to include this suffix.

2.5 *Turkish (a Turkic language of the Altaic group)*

Turkish (of Turkey or türkçe) does not seem to have this suffix *-z*. Its equivalent is *-p*, which is used to chained together a string of cooccuring verbal actions.

(11) *otur-pı je-pi çay iç*. (translates the Manchu example in §.2.8)

'to sit'-Converb 'to eat'-Converb 'tea' 'to drink'
'sit down, eat and drink tea' (Li 2010:369)

This form can be found in Manchu as well *-fi*, with variant form *-mpi*, *-pi* (Li 2010:369), as will appear below in the paragraph 2.8.

2.6 *Moksha-Mordvin (a Uralic language)*

Moksha is a Uralic language, spoken in the Russian Federation. Together with Erzya it constitutes the two main branches of Mordvin. Many Mokshan speakers live in Mordovia, a province located 500 km to the south-east of Moscow. It must nevertheless be noted that on the whole only one third of Mordvins live in Mordovia and that Mordvins are one of the most scattered "Nationalities" of the Russian Federation. Moksha provides examples of the converb formative *-(e)z*:

(12) авардез корхтай.

[1]avard-[2]ez [3]kořta-[4]j
[1]'to weep' [2]Converb [3]'to speak' [4]Present-P3sg
'He (or she)'s speaking, weeping (at the same time)'

Moksha has a rich set of "participles" but this formation *-(e)z* can never determine a noun, nor be predicate. It is actually not a Participle, not a nominal form of a verb, but it is a circumstancial verbal form that can only appear subordinate to a Predicative verb. *Cf.* Zaicz (1998:205), who mainly deals with Erzya, the sister language of Moksha, and Aliamkin (2000:178-181). When the action is either simultaneous or past, then the negative segment is /af/ (Present) or /ašəz/ (Past): /af pičədəz/ <аф пичедез> 'without worrying'. There exists an interesting reduplicated formation, with a durative value: /kštiz kšti/ (lit.) 'he (or she) dansingly danses /kšti/'. Aliamkin (2000:179) also cites a combination with /tijez/ '(while) doing', as in /avardez tijez/ (lit.) '(while) weeping doing', with an intensive value. Note that this converb can have a separate subject, which appears in that construction with the Nominative, which is reserved for the subject of the main verb:

(13) шись нинге ашезь стя.

^{5}šəj-^{6}s ^{7}niŋge ^{8}aš-^{2}əz ^{9}stea
^{5}'sun' ^{6}'Definite Article+Nom. ^{7}'yet' ^{8}'Negative particle' ^{9}'to rise, get up'
'[with] the sun having not yet risen'

This construction is semantically similar to Latin *ablativus absolutus*. Uralic is generally accepted to have fairly close connections with Indo-European languages and Altaic languages, as noted in Bomhard (2008:230): "A number of scholars have claimed that Indo-European and Uralic are more closely related to each other than either of them is to any other language or language family, while others have claimed that Uralic and Altaic are particularly close, even going so far as to set up a Ural-Altaic language family." These issues are discussed at the end of the paper.

2.7 *Mongolian Halh (an Altaic language)*

Mongolian is a family of languages and the Halh dialect of Mongolia provides examples of the converb formative *-ž*:

(14) Хоолоо идэж, пиво уусан. (Gaunt-Bayarmandakh 2004:87)

^{1}ho:lo ^{2}id-^{3}əž ^{4}pivo ^{5}u:s-^{6}ən
^{1}'meal' ^{2}'to eat'-^{3}Converb ^{4}'beer' ^{5}'to drink'-^{6}Past-P1sg
'I ate and drank some beer at the same time.'

The converb *-ž* expresses simultaneous occurence, and if replaced by *-d* it then expresses succession: 'I first ate and afterwards had some beer.' These converbial formatives are a typical feature of Mongolian languages as mentioned in the Introduction above.

2.8 *Manchu (a language of the Altaic group)*

Manchu is usually considered to be a branch of the Tungus-Manchu family. It does not have examples of the Converb formative *-ž/-z* but Manchu has *-fi*, with variant form *-mpi*, *-pi* (Li 2010:369). A suffix that also exists in Turkic as noted before.

(15) *te-fi je-fu cai omi*. (Li 2010:369)

'to sit'-Converb 'to eat'-Converb 'tea' 'to drink'
'sit down, eat and drink tea' (gloss in Li 2010:369)

From the perspective of this converb it would therefore appear that Manchu is closer to Turkic than to the rest of Tungusic-Manchu. This morpheme may constitute a major isogloss within the controversial group called Altaic.

2.9 *Udihe (a Tungusic language of the Altaic branch)*

Udihe is a language belonging to the southern branch of the Tungus-Manchu family. It has examples of the Converb formative -*s*:

(16) *Kept'e-si taŋi:-ni*. (Nikolaeva-Tolskaya 2001:741)

'to lie'-Converb 'to read'-Present-P3sg
'He is reading lying down.'

In this language the converb formative -*z* seems to be devoiced into -*s*.

2.10 *Korean (a language with unclear relationship with the Altaic group)*

Korean is a very debated possible member of the Altaic group. It can be noted that it has examples of neither the converb formative -*ž*/-*z* nor the Turkic-Manchu *p(i)*. What is most similar in function to the converbial suffixes surveyed in the paper is a suffix -*go*,

(17) *meog-go cha hanjan*.

'to eat'-'and' 'tea' 'to drink'
'eat and drink tea'

The meaning of the suffix -*go* seems closer to a sort of conjunction.

2.11 *Japanese (a language with unclear relationship with the Altaic group)*

Just like Korean Japanese is a very debated possible member of the Altaic group. It can be noted that it has examples of neither the converb formative -*ž*/-*z* nor the Turkic-Manchu *p(i)*. What is most similar in function to the converbial suffixes surveyed is a suffix -*te*, -*de*. Japanese also has an adverbial suffix for adjectives -*ku*, which has a phonetic similarity with Korean -*go*.

(18) *Watashi-wa suwat-te ocha nomu*.

'P1Sg'-'Thematic particle' 'to sit'-'Durative' 'tea' 'to drink'
'I'm sitting and drinking tea.'

(19) *Watashi-wa hayaku aruku*.

'P1Sg'-'Thematic particle' 'fast' 'to walk'
'I walk fast.'

The adverbial *hayaku* is one of the form of the Adjective *hayai* 'early, fast'. In all cases from the precise point of view studied in the paper Japanese does not appear to share anything with other Eurasian languages.

3. *A cladistic approach of Northern Eurasian languages*

The first part of the paper described the *converb* formatives for cooccurence in a number of languages. On the whole they can be divided into four groups:

1. Basque, Etruscan, North-East Caucasic, Uralic, Mongolian, Tungusic (without Manchu) with explicit traces of *z*,

2. Kartvelian, with no traces of *z*, *p* nor *ku*,

3. Turkic and Manchu, with another formative *p,

4. Korean and Japanese, with no traces of *z or *p, and possibly another formative *ku.

It is now interesting to compare this apparently simplistic "Stammbaum" with more elaborate classifications such as that of Nostratic developed by Bomhard (2008:28; 2010:21). A first point is the position of Afrasian: there is no converb suffix in that branch of Nostratic. This is coherent with what Bomhard (2008:521) states: "There is a growing feeling among several scholars studying Nostratic that Proto-Afrasian may be a sister language of Proto-Nostratic rather than a descendant language." Another point is the position of Japanese and Korean, as outliers. This again agrees with Bomhard (2010:2): "Other languages may belong as well, such as, for example, Korean and Japonic (Japanese-Ryukyuan). But much work needs to be done before these two groups can be convincingly shown to be related to Altaic (itself quite controversial), as is often assumed, let alone Nostratic." From the perspective of the present paper this cautious approach seems indeed supported. Kartvelian is also considered by Bomhard (2008:28) to be an early splitter within the Nostratic family. This point of view is supported by the absence of the widespread converb formative *z in that Southern Caucasic group.

To a large extent, but with reservations that need to be discussed, the languages or families where *z is attested coincide with what Bomhard (2010:2) calls Eurasiatic, originally a word introduced by Greenberg: "Eurasiatic, in turn, includes the following: Tyrrhenian (Etruscan, Raetic, and Lemnian), Indo-European, Uralic-Yukaghir, Altaic (Tungus, Mongolian, and Turkic), Chukchi-Kamchatkan, Gilyak (also called Nivkh), and Eskimo-Aleut." Chukchi-Kamchatkan, Gilyak (also called Nivkh), have not been studied in this paper for lack of relevant data. Eskimo-Aleut's polysynthetic structure does not have converbs: one of the morphemes used to connect verbs is *lu*, which appears to be a kind of conjunction with no relation with any of *z, *p or *ku. In that respect Eskimo-Aleut appears to be an outlier. At the same time the simplistic branching based on the converb morphemes for cooccurence reveals four major differences:

- Basque would appear to be more related to Eurasiatic languages than is usually assumed.

- North-Eastern Caucasic should be included in Nostratic, and the so-called Dene-Caucasian, that is supposed to include Basque, Northern Caucasic among others is to be considered a fictitious entity.

- Altaic would be confirmed to be a non-existing node, resulting mostly from areal factors. In that respect Turkic and Manchu appear to be outliers, related to neither Eurasiatic nor Korean-Japonic. Manchu should be separated from the rest of Tungusic.

- Proto-Indo-European (PIE) does not have the converb *z on the surface. This situation is interesting because it sheds light on the prehistory of PIE. Contrary to Kartvelian where there is no morpheme *z, retrievable for any function, or Afrasian, which does not have this kind of morpheme, PIE underwent a change in the function of *z. It is well-known that PIE and many ancient languages, like Latin or Greek, opposed two aspects for verbs: completed and incompleted. But there exists a third term, the so-called sigmatic aorist, which is neutral when it comes to aspect. This form adds a *s*-suffix unto the root. My analysis is that the aorist originates in a converbial form that has been promoted into a main form. This promotion nevertheless left a grammatical or semantic scar: the aorist is aspectually neutral, just as the *z*-converb for cooccurence is or used to be. It can be noted that the aorist is not the only form that has participial or converbial features: the suffix for P3Pl *-nt* is extraordinarily identical to the Present Participle *-nt*. From the present survey I would conclude that this formal identity is not a chance coincidence. This is the same morpheme. A set of verbal forms of PIE originates in converbs.

Incidentally it can be observed that some of the languages that do not have the converb *z seem to prefer coordination rather than subordination for verbs: Turkic, Manchu, Korean and Japanese. This difference may also have a genetic relevance.

4. *Conclusion*

The paper began with a survey of *Converb* formatives in a number of languages, with a special focus on a particular one *z, which is widespread throughout Eurasia. Simplistic as it may seem at first sight the groups based on this single converb coincide for the most part with genetic classifications based on considerably more information and data. It seems probable that this converb is very ancient and has a genetic relevance as a relique morpheme identifying a set of Nostratic languages with stronger links. Among the conclusions that can be reached on that basis two main points need further examination: Basque seems to have more Nostratic material than is usually hypothesized and the PIE verbal system involves a subset of forms, that were originally converbs (*s, *nt).

References

Aliamkin, N.S.

    2000   *Mokshen Kjal Morfologija*. Saransk: Krasnyj Oktiabrj.

Bomhard, Allan R.

    2008   *Reconstructing Proto-Nostratic*, *Comparative Phonology, Morphology, And Vocabulary*, Vol. 1 & 2. Leiden: Brill.

    2010   The current status of Nostratic studies. *The Macro-Comparative Journal*, *Vol. 1.1.*

Bonfante, Guiliano; Bonfante, Larissa

    2002   *The Etruscan Language. An Introduction*. Manchester University Press.

Gaunt, John; L. Bayarmandakh

    2004   *Modern Mongolian: A Course-Book*. London-New-York: RoutledgeCurzon.

Haspelmath, Martin

    1993   *A grammar of Lezghian*. Berlin-New York: Walter de Gruyter.

Lafitte, Pierre

    1980   *Grammaire basque (navarro-labourdin littéraire)*. Donostia: Elkar.

Nikolaeva, Irina; Tolskaya Maria

    2001   *A grammar of Udihe*. Berlin-New York: Walter de Gruyter.

Li, Gertraude Roth

    2010   *Manchu: A Textbook for Reading Documents*. Manoa: University of Hawai.

Trask, R. Larry

    1997   *The History of Basque*. London/New York: Routledge.

Wiedemann, Ferdinand Johann

    1884   *Grammatik der syrjänischen Sprache: mit Berücksichtigung ihrer Dialekte und des wotjakischen*. Sankt Petersburg: Kaiserliche Akademie der Wissenschaften.