

The state of the art of Uralic studies: tradition vs innovation

edited by
Angela Marcantonio



Collana Convegni 41

STUDI UMANISTICI
Serie Philologica

The state of the art of Uralic studies: tradition vs innovation

Proceedings of the 'Padua Uralic seminar'
University of Padua, November 11-12, 2016

edited by
Angela Marcantonio



SAPIENZA
UNIVERSITÀ EDITRICE

2018

Copyright © 2018

Sapienza Università Editrice

Piazzale Aldo Moro 5 – 00185 Roma

www.editricesapienza.it

editrice.sapienza@uniroma1.it

Iscrizione Registro Operatori Comunicazione n. 11420

ISBN 978-88-9377-066-8

Publicato ad aprile 2018



Quest'opera è distribuita
con licenza Creative Commons 3.0
diffusa in modalità *open access*.

In copertina: *The Uralic Languages Map*.

Index

Preface	vii
Introduction (<i>Angela Marcantonio</i>)	1
A 'steppe nomadic culture' vs a 'forest language': Modern identity dissonance in the history of the Magyars <i>Giuseppe Cossuto</i>	21
Information Structuring and typology: Finnic and Samic word order revisited through the prism of orality <i>M.M. Jocelyne Fernandez-Vest</i>	33
Issues of comparative Uralic and Altaic Studies (4): On the origin of the Uralic comparative marker <i>Juha Janhunen</i>	49
Revisiting the theory of the Hungarian vs Chuvash lexical parallels <i>László Marác</i>	59
Are the Hungarians Ugric? <i>Borbála Obrusánszky</i>	87
The impossibility of the evolutionary metaphor: Neogrammarians, family trees and linguistic affinity <i>Péter Pomozi</i>	107
Differential Object marking in Eastern Mari and Permian: A look from the field <i>Natalia Serdobolskaya</i>	129

The 'impossible' Comparison between the Creatures and Figures of the Hungarian Folklore and the Finnish Mythology <i>Elisa Zanchetta</i>	157
Contributors and abstracts	167

Preface

This volume contains the Proceedings of the ‘Uralic Studies’ Seminar, titled: *The State of the Art of Uralic Studies: Tradition vs Innovation*, held in Padua (~ Padova, Italy), November 12-13, 2016. The seminar was part of an ongoing series of seminars, conferences and workshops, organized by the Department of ‘Studi Linguistici e Letterari (DiSLL)’ of Padua University, precisely by the ‘Chair of Hungarian Language and Literature’, directed by Professor Cinzia Franchi. The topics of the various past (and future) seminars always revolve, obviously, around ‘Hungarian’ as a pivotal theme, however much diverse they may be.

The 2016 Seminar has been organized by Cinzia Franchi – supported by her Department, Head Department Professor Anna Bettoni – in cooperation with Angela Marcantonio, Professor of ‘Linguistics, Finno-Ugric & Hungarian Studies’ at Sapienza University of Rome – with the support of her Department of ‘Scienze Documentarie, Linguistico-filologiche e Geografiche’, under the direction of Professor Giovanni Solimine, as well as the ‘Facoltà di Lettere e Filosofia’, under the direction of the ‘Presidente’, Professor Stefano Asperti.

We wish to thank the speakers, who have willingly and generously accepted our invitation, as well as the various institutions that have offered us this great opportunity, supporting and sponsoring our ‘joint venture’, that is, as mentioned, the University of Padua and Rome ‘La Sapienza’ (and related Departments / Faculties), as well as CISUECO, that is: ‘Centro Internazionale di Studi Ungheresi e sulla Europa Centro Orientale’.

To give a detailed and precise record of the time, the venue, and the talks that took place, and other relevant items of information, we include the original program of the seminar itself.

We hope that the topics and related debates of our Seminar will raise much interest and further stimulating debates and research regarding Hungarian and the other Uralic languages. We also hope that this series of seminars and conferences may continue (perhaps by organizing, again as a joint venture, another Uralic seminar in the near future), with the purpose of bringing together scholars approaching Uralic studies from different angles and perspectives, against the background of a fully interdisciplinary approach.

Cinzia Franchi

Program of the seminar

FRIDAY 11

I SESSION – 10.00-13.30

Welcoming notes

SERGIO BOZZOLA

(Deputy) Head Department, University of Padua

CINZIA FRANCHI

University of Padua

Introduction: overview and purpose of the conference

ANGELA MARCANTONIO

University of Rome “La Sapienza”

The state of the art of Uralic studies: tradition vs innovation (an overview)

BORBÁLA OBRUSÁNZKY

Budapest: Károli Gáspár University

Are the Hungarians Ugrians?

GIUSEPPE COSSUTO

Cluj/Kolozsvár: The Institute of Turkish and Central Asian Studies - Babeş-Bolyai University

A ‘steppe nomadic culture’ vs a ‘forest language’: a modern identity dissonance in the history of the magyars

12.45 QUESTION TIME

13.30 LUNCH BREAK

II SESSION – 15.00-19.00**ELISA ZANCHETTA**

PhD, University of Florence

*The impossible comparison between the creatures and figures of the Hungarian folklore and the Finnish mythology***NATALIA SERDOBOLSKAYA**

Moscow: Russian State University for the Humanities

*Differential Object marking in Permic, Mari and Erzya-Mordvin***16.15-16.45 QUESTION TIME****M.M. JOCELYNE FERNANDEZ-VEST**

Paris: CNRS; Université Sorbonne Nouvelle

*Information Structuring and typology: Finnic and Samic word order revisited through the prism of orality***17.15-17.30 COFFEE BREAK****TAPANI SALMINEN**

University of Helsinki

*The background of grammatical parallels between Finnic and Samoyed***18.15-19.00 QUESTION TIME****SATURDAY 12****III SESSION – 10.00-13.30****LÁSZLÓ MARÁCZ**

University of Amsterdam; L.N. Gumilyov Eurasian National University

*West Old Turkic as a phantom-category: the Hungarian substrate of Chuwash***JUHA JANHUNEN**

University of Helsinki

*Proto-Uralic derivational issues***PÉTER POMOZI**

Budapest: Eötvös Loránd University

*The impossibility of the evolutionary metaphor: Newgrammarians, family trees and linguistic affinity***12.15-13.00 QUESTION TIME****THE ORGANIZERS**

Cinzia Franchi

Angela Marcantonio

SATURDAY 12**III SESSION – 10.00-13.30****LÁSZLÓ MARÁCZ**

University of Amsterdam;
L. N. Gumilyov Eurasian National University
*West Old Turkic as a phantom-category:
the Hungarian substrate of Chuvash*

JUHA JANHUNEN

University of Helsinki
Proto-Uralic derivational issues

PÉTER POMOZI

Budapest: Eötvös Loránd University
*The impossibility of the evolutionary metaphor:
Newgrammarians, family trees and linguistic affinity*

12.15-13.00 QUESTION TIME

ORGANIZZAZIONE

Cinzia Franchi (cinzia.franchi@unipd.it)
Angela Marcantonio (angela.marcantonio@uniroma1.it)

SEGRETERIA ORGANIZZATIVA CONVEGNI

Luisa Sibillio – Emilia Eleno
tel. 049/827.4870
convegni.dislil@unipd.it

ELABORAZIONE GRAFICA – NADIA RADOVICH

**CENTROSTAMPA
PALAZZO MALDURA**

*The state of the art
of Uralic studies:
tradition vs innovation*



Università di Padova
11-12 November 2016

Sala Merio Scattola – Palazzo Maldura
piazza Gianfranco Folena, 1 - 35137 Padova

Valido ai fini dell'acquisizione dei 3CFU - Altre attività

FRIDAY 11**I SESSION – 10.00-13.30***Welcoming notes***SERGIO BOZZOLA**

(Deputy) Head Department, Università di Padova

CINZIA FRANCHI

Università di Padova
*Introduction: overview and purpose
of the conference*

ANGELA MARCANTONIO

Università di Roma "La Sapienza"
*The state of the art of Uralic studies:
tradition vs innovation (an overview)*

BORBÁLA OBRUSÁNSZKY

Budapest: Károli Gáspár University
Are the Hungarians Ugrians?

GIUSEPPE COSSUTO

Cluj/Kolozsvár: The Institute of Turkish and Central Asian
Studies- Babeş-Bolyai University
*A 'steppe nomadic culture' vs a 'forest language':
a modern identity dissonance in the history
of the magyars*

12.45 QUESTION TIME

13.30 LUNCH BREAK

II SESSION – 15.00-19.00**ELISA ZANCHETTA**

PhD, Università di Firenze
*The impossible comparison between
the creatures and figures of the Hungarian folklore
and the Finnish mythology*

NATALIA SERDOBOLSKAYA

Moscow: Russian State University for the Humanities
*Differential Object marking in Permian,
Mari and Erzya-Mordvin*

16.15-16.45 QUESTION TIME

M.M JOCELYNE FERNANDEZ-VEST

Paris: CNRS; Université Sorbonne Nouvelle
*Information Structuring and typology:
Finnic and Samic word order revisited
through the prism of orality*

17.15-17.30 COFFEE BREAK

TAPANI SALMINEN

University of Helsinki
*The background of grammatical parallels
between Finnic and Samoyed*

18.15-19.00 QUESTION TIME

Introduction

1. The state of the art

1.1. Nowadays there is plenty of interesting (mainly synchronic) studies on the Uralic (U) languages, based on modern linguistic theories and approaches, including: the sounds of the U languages from the point of view of modern phonological and prosodic studies; new approaches to morphology and morpho-phonology; argument structure, word order and information structuring; differential object marking; U languages and linguistic universals; U languages from a cognitive perspective; U languages and 'Tense, Mood & Aspect' structure (so-called TAM); 'corpus linguistics' approach, non-finite subordinate construction (NFC) / patterns and 'sententiality' (a set of characteristic features possessed by NFCs), etc. Not to mention the recent results obtained in the field of 'language-in-contact', through the adoption of new methods of analysis (see below), etc. It is certainly the case that several linguistic phenomena (believed to be specifically U) have become of general, theoretical interest, and one could apply to the U languages (and related research) what Helasvuo & Campbell (2006: 2 ff.) have stated with regard specifically to Finnish:

... we find numerous instances of Finnish examples playing significant roles in theoretical discussions of vowel harmony, gemination, meter, codeswitching, child language acquisition, language contact, loanwords, language change, word order, theoretical morphology, computational linguistics, morphological processing, case [...], possession, anaphora, metaphor [...], null subjects, and typological issues of many sorts

There are also recent volumes addressing the historico-political and socio-economical status of the U languages and people, such as Taagepera (1999): *The Finno-Ugric Republics and the Russian State*, or Nanovszky (ed. 2004): *The Finno-Ugric World*, volumes that also provide geographical information about the U languages / peoples and related contact with U and non-U people. However, this recent waves of exciting research do not typically relate to, or touch upon, the topic of the origin and classification of the U languages, topic that, in fact, has remained on a much more 'traditional' footing: the conventional U family tree model – whose validity is generally taken for granted – keeps being re-proposed and promoted both within specialistic studies and on the world stage, essentially unchanged, as if it were more of a 'dogma', than a model of analysis. As a matter of fact, the U theory is now about 200 years old, and alternative models (be they just more or less deep revisions, or rejection of the model altogether) have sprung up numerous through the years, as shown in the following article by Salminen (1999): *Problems in the taxonomy of the U languages in the light of modern comparative studies* (see also Marcantonio (2002; chapter 2.) for an outline of various, alternative models). However, regrettably, these 'dissonant voices' do not appear to have had much of an impact on more traditional views.

1.2. A similar situation can be found in the field of 'language-in-contact', including 'diachronic contact and change'. There are many studies that address the topic of contacts within U and between U and non-U languages, as mentioned (see an overview in Laakso (2010b)), such as the contacts between the various U languages with the neighbouring Germanic, Slavic, Baltic languages (see for example Joki (1973)), or the (assumed) contacts at the level of proto-languages, that is, the Indo-European vs the U proto-language (see, for example, Koivulehto (1991 & 1992/3)). Unlike these more traditional studies, the recent studies tend to adopt innovative methods of research, working with notions such as: (sound) variation (as against the traditional notion of 'regularity' of sound change and substitution), code-mixing / switching, creole languages, differential loan sources, areal distribution of variants, 'whole-sale-borrowing', substratum interference, etc. (see Metsmägi, Sedrik & Uusküla (2014), or the *Atlas Linguarum Fennicarum* (ALFE)). However, these new concepts and methods appear to have had little impact on some, major areas of more traditional U studies dealing with

the topic of ‘borrowing’. For example, it could be argued that the standard classification of Hungarian as a ‘pure / proto-typical’ U language – as it were – is an instance of lack of interaction between tradition and innovation within this field (but see Marácz (2012) and in this volume). As a matter of fact, Hungarian is recognized as an ‘isolate’ language within the family (“*egyedülálló*”, according to Laakso (2010a)), having no close relatives, not even the Khanty (/Ostyak) and Mansi (/Vogul) language, all grouped together under the ‘Ugric’ node within the conventional tree diagram – no explanation has been provided thus far as to the cause(s) of this ‘isolation’ (as far as I know). That Hungarian is indeed ‘peculiar’, is testified (also) by the fact that Hungarian (in the same way as the Ob-Ugric languages) is missing from Janhunen’s (1981 & 1998) and Sammallahti’s (1988) systematic comparison with the other U languages, for the purpose of reconstructing proto-U – see also Csúcs (2008) regarding the lack of reconstruction of the Ugric node, and, therefore, of the U node. In other words, Janhunen’s (1981) rigorous reconstruction of U (at the basis of any subsequent reconstructions), is only a ‘partial’ reconstruction, an “approximation” (Janhunen 1998: 461), because it has been achieved without taking Hungarian (and Ob-Ugric) into systematic consideration. As the author himself observes (1998: 461), including these languages into the process of reconstructing would trigger indeed “considerable taxonomic and reconstructional problems”. As to the issue of the cause(s) of this ‘isolation’ – hardly ever properly addressed in the literature, as mentioned – among the possible explanations, the following one could be put forward: the Hungarian language is isolated because it displays a deep, pervasive ‘Turkic component’ (whatever the origin of this component may be). As a matter of fact, Hungarian and the Turkic languages share, among others, the following traits: a) a statistically significant number of (basic) lexicon, including verbs and adjectives (typically high in the scale of resistance to borrowing); b) a statistically significant number of derivational morphemes¹, morphology being also typically resistant to borrowing (although derivational morphology is believed to be less resistant than

¹ Hungarian does not share functional morphology with Turkic; however, this fact is not relevant for the issue under discussion, since Hungarian does not share functional morphology with any U language either (see Korhonen (1996), Marcantonio (2002: chapter 8.)). The ‘complex’ case suffixes of Hungarian are relatively recent formations, and its ‘simple’ suffixes, such as the accusative *-t*, are not typically in line with the traditional U reconstruction.

functional morphology²; see Comrie (2008:15 -16) and Thomason & Kaufman (1988:74-5)). In addition to the shared, single (derivational) suffixes, Hungarian and Turkic share the process of ‘word formation’, that is, the same internal morphological structure of the word: numerous Turkic and corresponding Hungarian words consist of a corresponding stem, a corresponding (sequence of) suffixes, and even, the same path of sound development for some of these suffixes, as shown in Marcantonio (2016 & 2017a). Not to count shared isomorphism (possessive, post-positional constructions) and shared, wide spread Eurasiatic (~ Eurasian) typological features. In other words, Hungarian and (Old) Turkic display statistically significant instances of what could be called ‘global /full-scale correspondence’ – a state of the art surprisingly *not* shared between Hungarian and the U languages. In this regard one may compare Lehtisalo (1936) for a list of all the (supposed) U suffixes, and Róna-Tas & Berta (1011) for a comprehensive account of the lexical and morphological similarities and /or correspondences occurring between Hungarian and Turkic. This state of the art has since long been known, precisely, since the times of the ‘Ugric-Turkic battle’ (for which see below, as well as Marcantonio, Nummenaho & Salvagni (2001)), but nothing has come out of it in terms of re-thinking, re-visiting these intricate issues. Not even the publication of the (quoted) volumes by Róna-Tas & Berta (1011), that clearly document how deeply rooted³ the Turkic component is in Hungarian, appears to have triggered renewed debates about the ‘real’ nature of Hungarian, as far as I know (for example, the standard interpretation of the ‘pure’ U nature of Hungarian has been recently re-proposed in Bence (2014)). Indeed, if and when the presence of this Turkic component in Hungarian is mentioned in the literature, the conventional interpretation is still put forward, according to which there are, as expected, many loan words in Hungarian, from many languages, as a ‘normal’ result of ‘normal’ contacts with

² As Comrie (2008:15 & 16) puts it, the received wisdom is that “agglutinative morphology is more borrowable than fusional [...]. While details are intricate and often idiosyncratic, a general pattern emerges whereby inflectional morphology certainly can be borrowed and certainly can participate in language mixing; however, nevertheless there is a tendency for such inflectional morphology to be agglutinative rather than fusional”. Similarly, according to Thomason & Kaufman (1988:74-5), despite the fact that morphological borrowing is de facto quite complex, it is received wisdom that derivation is borrowed more frequently than inflection.

³ RT&B research draws from a long standing, high level tradition of Hungaro-Turkic studies, as well as from their own original research

various languages / peoples, including Turkic. However, the relevant fact that the level of borrowing from Turkic, is, in fact, ‘not-normal’ (because, as mentioned, basic lexicon and (derivational) morphology has been ‘borrowed’ too), is usually brushed under the carpet. If the Turkic component present in Hungarian – a clear instance of ‘intensive’, rather than ‘normal’ borrowing – were investigated within the framework of current ‘language-in-contact’ methods and notions (investigation that has not yet been carried out, as far as I know), Hungarian would probably be classified as some sort of ‘mixed’⁴ language⁵ (see foot note (4) and (5)). In other words, there seems to be (a majority) consensus in the literature that intensive borrowing – as evidenced, in particular, through the borrowing of (complete) morphological paradigms⁶ – is typically the result *not* of ‘normal’ contacts, ‘normal language transmission’, but rather the result of ‘non-normal contacts’, ‘non-normal transmission’,

⁴ The definition of ‘mixed’ language is rather controversial. Here there is no room, and is not relevant either, to enter into this debate. A simple, practical guideline, in my opinion, is the definition proposed by Comrie (2010: 26-7): “Languages that combine elements of different genealogical origins to such an extent that at least it becomes questionable whether the result can be genealogically assigned unequivocally to one of its roots rather than the other”. In practice, every language is ‘mixed’ to a certain extent, and every linguistic element (lexical, grammatical, etc.) can in principle be borrowed, given the right circumstances. Thus, it is rather an issue of ‘degree’ of mixture that is at stake, and, in fact, not all linguists accept the distinction between ‘normal borrowing’ / ‘ordinary contact’ and ‘intense, special contact / borrowing’, the latter being indeed typical of mixed languages (as discussed in the text). Whatever the case, by the definition ‘mixed’ here, I loosely refer to a language that combines material from two (or more) existing languages.

⁵ Some linguists have asked themselves whether Mari (spoken in the middle stretch of the Volga River, etc.) could be classified as a ‘mixed’ (/creole) language. Indeed, Mari and Turkic peoples have been living side by side for long time, the Volga-Kama region representing a *Sprachbund* area inhabited by various, different languages / peoples (the Russian occupation of the Volga region took place since the 16th century (see Pomozi (2004)). This classification, however, has been generally rejected, for example by Itkonen (2010).

⁶ This analysis, in my opinion, applies to Hungarian even if there is no shared functional morphology / shared morphological paradigms between Hungarian and Turkic (as pointed out in foot note (1)). As a matter of fact, it could be argued that in the agglutinative Eurasianic languages derivational morphology plays a more important role than functional morphology – for the task of assessing genetic relations – due to the following reason: whilst case suffixes are typically independent formation of the single languages – being mainly the (unstable) results of loose processes of grammaticalization and exaptation – derivational morphology / word formation represents the oldest levels of language, being an integral part of the internal, morphological structure of the word. Similar views are expressed in Robbeets (2010: 82), who states that derivational morphology “belongs to the older strata of a language and provides rather reliable evidence to demonstrate common ancestorship”.

that is, basically, language mixing (see for example Gardani *et al.* 2015: 10). Thus, within this new framework of analysis, a possible interpretation of the nature of Hungarian could be that it is not a 'pure U' language, containing a certain, 'normal' amount of loanwords (whatever the definition of normal in this context may be), but rather a 'mixed' language, mixed with Turkic⁷ (and other Asiatic) languages – assuming, of course, that the classification of Hungarian as a U language is considered sound and valid.

1.3. On a similar note, it could be argued that the lack of innovation, or just revision(s), within the framework of the traditional U paradigm, could hamper progress in the field of 'typology' and 'language universals' (for which see Comrie (1989) and Croft (1990)). For example, the U languages are characterized by the presence of an intricate bundle of 'exotic' typological features and structures, that are nowadays the object of intense investigation, and through innovative methods of analysis. Some of these features are wide spread, such as: a) absence of grammatical gender; b) use of suffixes and postpositions; c) agglutination; d) absence of definite article (with the exception of Hungarian, where the article is a recent development, and Mordvin, which has an enclitic article); e) *determinant* → *determinatum* word order; f) use of grammatical singular in connection with a noun preceded by numerals or other quantifiers, etc. Other features – of structural / pragmatic type – are shared among several languages only, such as: g) lack of a proper genitive suffix: the notion of possession is implemented by the adoption of a possessive / genitive construction, whereby the possessor is (typically) unmarked and the possessive relation is (typically) rendered explicit by marking the possessed element with the 3rd person possessive suffix, this being the case for Hungarian, Mansi and Khanty⁸ (Abondolo 1998: 32, Sipőcz 2006: 300, etc.); h) the adoption of a specific marker (generally, but misleadingly, called 'accusative'), to mark the definite / topical object

⁷ This analysis does not take into count the subsequent mixing of the early Magyars with the (Indo-)European languages /peoples they were surrounded by after settling in the Carpathian Basin (for which see Berend *et al.* (2013)). At this point, their ancient tribal structure was eradicated and replaced by a western, Christian State – fact that surely must have had an enormous impact on the original, overall structure of the language and its subsequent development.

⁸ Other U languages / dialects instead, such as Tundra Nenets and Finnish, have a distinctive genitive suffix, *-h* and *-n* respectively.

only, as in some dialects of Mansi (/Vogul, see Marcantonio 1994), or in Finnish, where the accusative/genitive *-n* is used for definite, singular objects, and *-Ø* marking, + the plural suffix *-t*, for definite, plural objects (Marcantonio 1988). This basic uniformity, basic coherence of the typological / pragmatic structure (despite existing variation, for which see discussion below), is another remarkable network of U traits that are, in fact, of great interest on the linguistic stage, worldwide (see Marcantonio 2017b). However, these traits are also present in the so-called 'Altaic' languages, some of them, again, wide spread (features (a)-(f)), others, again, present in some languages: for example, the traits (g) and (h) – marking of the definite/topical object only and genitival /possessive construction – are present in the Turkic languages. There are several, recent publications that argue in favour of the existence of what Janhunen (2001:213, 2007, 2009:71, 2012 & 2014) has labelled as the: "Ural-Altai complex", "Ural-Altai areal context", or the "single original area of Ural-Altai typology". In other words: the typologically compact (despite normal variation) Eurasiatic area, criss-crossed by various structural, pragmatic, typological, as well as (some) morphological isoglosses. Nevertheless, the features in question are still mainly believed and purported as being 'unique' to the U languages, thus giving way to a distortion, a misleading representation of language universals and typology both within the Eurasiatic area, and in more general terms. This is particularly evident in the case of Hungarian (once again), since, as discussed above, Hungarian shares a thick bundle of isoglosses with Turkic, arguably thicker than with any other U languages. This also means that these 'U features' cannot really be predicted on the basis of the (assumed) genetic relatedness of the U languages.

1.4. The aim of the Padua seminar, and, consequently, of this volume (as the 'Proceedings' of the seminar) is that to bring together linguists working on the U languages from different perspectives, with the purpose of increasing the exchange of ideas and fostering mutual, beneficial influences on each other field and methods of study. This, in turn, could increase knowledge, or awareness regarding the current 'state of the art of U studies', not just among specialists, but also, and mainly, among linguists and general public outside the field. Surely, the 'image' of the U languages, and, perhaps, of the Altaic languages too, or even of the 'Ural-Altai linguistic belt' (see above), would bene-

fit not just from the results of current, vibrant synchronic research, but also from overcoming the stagnation of the standard U theory, through a reassessment of the correlations, the isoglosses occurring across the whole of the Eurasiatic linguistic area, as shown above.

This enterprise is, of course, just an attempt, a small step in this direction, for such a vast, fascinating topic.

2. Introducing the articles

In this session I shall introduce the articles contained in the volume. Firstly, I shall briefly present them inserted within the context of the purpose and scope of the seminar /volume (par. 2.1.), without however following the order of presentation of the respective talks at the time of the seminar. Then (in par. 2.2.) I shall report the abstracts of the talks (/ articles).

2.1. As the reader can see from the *Index* of the volume, there are well four articles dedicated (mainly) to Hungarian. This was *not* planned by the organizers of the conference (even if Hungarian is the pivotal theme, as mentioned in the *Preface*), being instead a mere coincidence – a coincidence that, however, significantly and faithfully reflects the fact (highlighted above) that Hungarian is a problematic language within U – but, perhaps, not within the Eurasiatic belt. László Maráczy' article (*Revisiting the theory of the Hungarian vs Chuvash lexical parallels*) deals with the long-standing issue of the (assumed) 'borrowing' from Turkic into Hungarian. The author rejects the conventional analysis (for which see discussion above), and proposes an alternative model of explanation, bypassing the strait jacket of the conventional U and Altaic theory, and looking instead at some specific isoglosses that encompass the languages of the Eurasiatic belt, whilst also adopting the notion of substratum interference and *Sprachbund*. Borbála Obrusánszky and Giuseppe Cossuto, in their article (*Are the Hungarian Ugric? and A 'steppe nomadic culture' vs a 'forest language': Modern identity dissonance in the history of the Magyars*, respectively), address the issue of the isolated position of Hungarian from a historian point of view – the historical back ground, if available (as in this case), being an integral part of historical linguistics research, a vital component for 'reconstructing' the origin and development of languages and peoples. Both authors carry out a detailed analysis of the early historical sources that refer to the Hungarians (/ early Magyars), and /or other languages and peoples now classified as U.

On the basis of the emerging historical picture, both authors (independently from one another) re-affirm a scenario that has been known all along, that is: the evidence from the historical sources (however much unclear they may be at times) contradicts the classification and the origin of the Magyars as U peoples. As a matter of fact, the following holds. First, all the relevant sources unanimously refer to /describe the Hungarians as ‘nomadic peoples’ of the (Eurasian) steppe. Second, there is nowhere in these records any mention, any reference or clues to U peoples /languages, or any U, material and / or spiritual ‘culture’. Third, the first inhabitants of the Eurasian steppe we know of are the local nomadic peoples, typically referred to as Skytians, and often mentioned in connection with the Huns. Last, but not least, there is no historical, cultural, ethnographic or other types of connection between the Hungarians and the Siberian Ob-Ugric people (Mansi /Vogul and Khanty / Ostyak), with whom Hungarian is associated within the traditional family tree – indeed, the Ob-Ugric languages are not particularly close to Hungarian, apart from sharing with it some structural, typological, ‘Eurasian’ features, as discussed above. Thus, the question might be asked: should the evidence from history ‘trump’ the linguistic model? If not, how can the contradiction be resolved? Elisa Zanchetta (with her *The ‘impossible’ comparison between the creatures and figures of the Hungarian folklore and the Finnish mythology*), adds the final touch to the above depicted scenario, from the point of view of mythological research: there is no close, ‘unique’ connection between the Finnish (/Finnic) mythological world and the Hungarian mythological world, apart from rather universal recurrent themes, present in many cultures / mythologies of the world. Thus, Hungarian turns out to be an ‘isolate’ also with regard to folklore and mythology. All in all, these articles, consistently and independently from one another, highlight, once again, the contradictions, the ‘dissonance’ existing between the tenets and predictions of the linguistic model, and the actual historical, cultural and ethnographic (as well as archaeological) knowledge we have regarding the early Magyars.

Three of the talks / articles contributing to this volume revolve, more or less directly, round the issue of the soundness, or otherwise, of the comparative method, and the results it yields. In fact, according to some scholars, the method is basically reliable, whilst for others it is not – both in general terms and with specific regard to the U family. Péter Pomozi (in his *The impossibility of the evolutionary metaphor: Neo-grammarians, family trees and linguistic affinity*), highlights the intrinsic inadequacy of the

Darwinian, evolutionary model to faithfully capture real relations among real languages, particularly because of the following two factors: a) the indisputable tautological character of the comparative method; b) the fundamental ‘variable nature’ of language, a ‘fact’ that is at odds with the *Lautgesetz* principle – the founding principle of the comparative method and overall model. In contrast, Juha Janhunen (in his *Issues of comparative Uralic and Altaic Studies (4): On the origin of the Uralic comparative marker*), obviously considering the process of reconstruction and the results it yield as basically sound, proposes a new reconstructed item – in line with traditional practice – to add to the corpus of U reconstructed suffixes (the comparative marker **-mpA*). However, an interesting element of ‘innovation’ is introduced into this analysis and reconstruction with regard to traditional U reconstructions: a plausible typological parallel is proposed from the realm of the Turkic languages, this in turn reinforcing Janhunen’s and other scholars’ claim of the existence of the Eurasiatic typological belt, complex. Similarly, Tapani Salminen (in his *The background of grammatical parallels between Finnic and Samoyed*), introduced an innovative perspective in his presentation⁹ regarding the long standing issue of the internal taxonomy of the U tree – in particular, the position of Samoyedic. According to the author, “there are non-trivial parallels in the grammatical structure of the Finnic and Samoyed languages”, including *Object Marking* and *Non-finite Subordinate Patterns* (for which see above, as well as the contribution by Serdobolskaya and Fernandez-Vest, respectively, below). Thus, Salminen tried to determine whether these features, insofar as they cannot be explained away as straightforward retention, are based on language contact or convergent development.

Finally, the talks /articles by Natalia Serdobolskaya and M.M. Jocelyne Fernandez-Vest (*Differential object marking in Eastern Mari and Permian: A look from the field*, and *Information structuring and typology: Finnic and Samic word order revisited through the prism of orality*; respectively) deal with two phenomena – ‘differential object marking’ and ‘information structuring’, respectively – that are generally considered to be specific U traits, and whose study is currently very popular and conducive of innovative results (as mentioned). In particular, Serdobolskaya examines the various ways in which the direct object can be marked across

⁹ Regrettably, Prof. Salminen’s article does not feature in this volume, having not reached the editors within the established deadline. For the same reason, the related abstract is missing in the next paragraph.

some U languages (according to the presence *vs* absence of certain parameters), pointing out to similarities and differences. Fernandez-Vest, highlighting the nature of what she calls the “information grammar of orality”, as opposed to the grammar of “normalized languages” (those with a long(er) written tradition), observes the rarity of non-finite constructions in Sami, whose register is still mainly oral, as opposed to Finnish. As to the relevance of these phenomena (and related studies) to the theme of our seminar/volume, it can be observed that they cannot be ‘predicted’ from the ‘genetic’ classification of the languages under investigation as U, for two reasons: a) these phenomena are present also beyond the U area; b) they may represent (also) the manifestation of universal structures of language, including basic cognitive structures. Thus, for example, differential object marking is present in Turkic, whilst the rarity of non-finite constructions in Sami appears to contradict the claim that these constructions are genuine, ancient U constructions – not to count the fact that, once again, they occur also beyond the U area.

2.2. In this session the abstracts of the talks/article are listed, according to the alphabetical order of the authors’ family name.

Cossuto, Giuseppe; Cluj/Kolozsvár: The Institute of Turkish and Central Asian Studies; Babeş-Bolyai University

A ‘steppe nomadic culture’ vs a ‘forest language’: Modern identity dissonance in the history of the Magyars

With the title: *Gesta Hungarorum* scholars refer to two Hungarian *Chronicles*, that are, however, written in Latin. The *Chronicles* narrate the history of the Hungarian peoples and their old standing connection with eastern, Asiatic peoples, such as the Huns and the Schyrians. The *Chronicles* also report important items of information relating to the legends of the Hungarians, as well as their desire, their aspiration to trace back their own origin to much older populations. In this presentation, whilst trying to disentangle what is ‘history’ from what is ‘legend’ in the events reported in the two *Chronicles*, the author addresses the issue of the origin of the Magyars (before they arrived and settled in the Carpathian Basin). This analysis is carried out from the Hungarians’ own perspective, that is: the perspective of peoples that originally lived and firmly belonged to the socio-cultural, economic, military system of the nomads, to then find themselves firmly located and settled in Western Europe – fully immersed in the Western Euro-

pean culture as well as socio-economic and military system at the very time of the writing of the *Chronicles*. Despite numerous attempts, there does not seem to be (yet) a credible, evidence-based explanation of *why, when* and *how* the early magyars would have left their original, U *habitat* and way of life to embrace the nomadic way of life of the steppe peoples: a modern identity dissonance in the history of the Magyars.

Fernandez-Vest, M.M.Jocelyne; CNRS; Université Sorbonne Nouvelle
Information Structuring and typology: Finnic and Samic word order re-visited through the prism of orality

Fernandez-Vest's demonstration (a contribution to the Information Grammar of Orality (IGOR) theory), is based on two types of constructions: 'Detachment Constructions' (DEC) and 'Non Finite Constructions' (NFC). With an 'Information Structuring' (IS) methodology that lays stress on the Rheme – the pivot of two binary information strategies involving *Initial* and *Final Detachments*, Theme-Rheme and Rheme-Mneme – a typology of Question-Answer pairs can challenge the traditional word order classifications. Discourse configurational languages alternately employ the possibilities of fronting oblique arguments and detaching them, which partly contradicts the hypothesis of syntactic rigidity as favouring Detachments. Even in languages possessing a Detachment option, the written style tends to systematically reduce the spoken segmentation to a condensed utterance, whatever morphological complexity this might imply. The comparison in Sami-Finnish translations of two current NFCs, temporal and referative constructions, shows that NFCs have a higher frequency in Finnish. The IS articulation of the Sami sentence is generally binary, whereas the synthetic structure of the Finnish sentence is enunciatively opaque: a temporal NFC, which functions as an adverbial clause rather than a separate Theme-clause, seems to imply a cognitive reductionism typical of normalized language. This could partly explain the rarity of this 'Uralic' (so to say) construction in Sami, a language whose register is still mainly oral.

Janhunen, Juha; University of Helsinki

Issues of comparative Uralic and Altaic Studies (4): On the origin of the Uralic comparative marker

This article examines the issue of the origin of the comparative marker **-mpA*, as attested in three U branches: Finnic, Samic and Hungarian (for example Hungarian *-bb* and Finnish *-mpi* ~ *-mpa*).

This marker is conventionally assumed to be based on a ‘moderative’ meaning, as still present in Samoyedic, or, alternatively, on a ‘contrastive’ meaning, as attested in a few pronominal derivatives in Finnic (see for example Finnish *jompi kumpi* ‘either one’). As a new possibility, the author takes up a forgotten proposal of G.J. Ramstedt, according to which this form may actually be based on the nominalization of translative verbs in **-m-*, as formed derivationally from nominal roots. This possibility appears to be supported by an interesting typological parallel from the realm of the Turkic languages. In addition, still according to the author, the issue is connected with the structure of roots and stems in proto-U, especially with the question concerning the formation of the so-called ‘consonant stems’.

Marác*z*, László; University of Amsterdam; Astana: Gumilyov Eurasian National University.

Revisiting the theory of the Hungarian vs Chuvash lexical parallels

Hungarian *vs* Chuvash lexical similarities (and / or correspondences) have since long been the object of intense investigation, these parallels being part of a broader pattern that opposes the isoglosses /r/ *vs* /z/ in Hungarian, Chuvash (West Old Turkic), and (at times) Mongolian on the one hand, and Common Turkic (East Old Turkic) on the other hand, respectively: the so-called rhotacism *vs* zetacism isogloss. Within the framework of the standard U theory, the Hungarian words having a counterpart in Chuvash (and /or Mongolian) are classified as ‘(normal) borrowing’ from either Chuvash, or one of its hypothesized ancestor languages, such as Volga Bulgar (West Old Turkic). The possibility of a different interpretation of the facts, including a different direction of borrowing, that is, from Hungarian into Chuvash – has simply never been taken into serious consideration. However, the standard analysis is not void of problems. For example, it assumes the existence of two, sharply different branches of Turkic (as already hinted at): West Old Turkic *vs* East Old Turkic, assumption whose supporting evidence, however, does not stand up to scrutiny. In this article, the author argues that the rhotacism isogloss observed in Chuvash is the outcome of substratum influence from Hungarian, substratum influence that materialized within the context of intensive and layered processes of contact between the different languages / language families co-existing in the Volga Bends region (between the 4th/ 5th and the 10th / 11th centuries). If this

analysis proves correct, it will have a major impact on the U and Altaic studies, because: a) the West Old Turkic branch (however much postulated through a long, high-quality tradition of Hungaro-Turkic studies), turns out to be a 'phantom-category'; b) the instances of rhotacism vs zetacism are to be interpreted as 'Eurasianic isoglosses', going back to the Early Central Asian linguistic area, that includes also the early Magyars. In turn, this will give further support to the thesis of the existence of the 'unique Eurasianic typological complex' (as discussed above).

Obrusánszky, Borbála; Budapest: Károli Gáspár University
Are the Hungarians Ugric?

In the 19th century some western European scholars established a new linguistic model of language formation and diffusion, and grouped (several) languages of the world into the nowadays well-known language families / family trees. However, the Hungarians were adversely affected by this: the newly established Finno-Ugric, then Uralic family tree model (supposedly a 'scientific' model) prescribes for the Hungarians an origin and a language classification that sharply contradict their ethnography, anthropology, as well as the narrative of their own ancient *Chronicles* – not to count the evidence provided by the ancient historical sources referring to them. In order to promote, and even enforce, what has now become the standard U theory, the Habsburg House, in 1858, forcefully changed the structure of the Hungarian Academy of Science, by removing Hungarian linguists and replacing them with (mainly) German linguists, who would consent to the Habsburg's new language policy. By doing so, they also cut off the Hungarians from their own real ancestors, the Central Asia peoples (such as the Scythians and the Huns), and 'strait-jacketed' them into an alien community. Their pre-history had to be rewritten. This means that the Hungarians had no connection of any sort with the peoples of Siberia, specifically the Khanty or Mansi (the Ob-Ugric peoples), as featured instead in the traditional family tree diagram. This being the state of the art, the author argues that the U 'commonwealth' never existed. As to the Magyars, they shared the way of life of the steppe peoples – the Scythians, Huns, Turks, etc. – who eventually settled down also in the area of the Ural-mountains (a mineral-rich area), in this way influencing the local communities.

Pomozi, Péter; Budapest: Eötvös Loránd University

The impossibility of the evolutionary metaphor: Neo-grammarians, family trees and linguistic affinity

The historical-comparative method is based on some Neogrammarian postulates that are, however, not supported by the linguistic evidence. These postulates have by now become ‘dogmas’, that, according to the author, should be dismissed. For example, the excessive emphasis attributed to phonetic / phonological changes, and their (alleged) regularity – *Lautgesetz* – has led the evolutionist model astray: on the one hand, similar sound shifts are known as occurring in a huge number of languages, on the other hand, phonological variables can and do coexist in both space and time. However, the Neogrammarian model does not take these facts into account – particularly the reality of the fundamental ‘variable nature’ of languages – although plenty of counter evidence to the conventional model has been put forward by other schools of linguistics, such as the so-called school of philosophical linguistics, dialectology, and, more recently, cognitive linguistics and socio-linguistics. Furthermore, researches on language families have been built to the analogue of the human family relationships, the Darwinian model of evolution, whose scope of applicability to languages has been overestimated. As a matter of fact, the evolutionary metaphor is intrinsically unable to describe every type of relations occurring between really existing languages, considering the historical, socio-cultural circumstances and events the languages constantly undergo to. However, may be the most problematic weakness of the Neogrammarian model is its indisputable tautology, as the author will spell out in his essay.

Serdobolskaya, Natalia; Moscow: Russian State University for the Humanities

Differential object marking in Eastern Mari and Permian: A look from the field

One of the characteristic features of the U languages is the phenomenon of so-called *Differential Object Marking* (DOM). In the Balto-Finnic languages the distribution of the markers of the direct object is, mostly, *verb-oriented*, that is, based on verbal categories such as tense, aspect, mood or polarity. Some reference grammars of Udmurt and Mordvin also consider the verbal categories among the basic factors that regulate the marking of the object. However, this is not necessarily the case in all the dialects of these languages. In this article, the author is

going to present new data, drawn from her own fieldwork research in the Permic languages (Pechora dialect of Komi-Zyrian, Beserman dialect of Udmurt), Eastern Mari and Erzya-Mordvin (Shoksha dialect). She argues that in these idioms the system of DOM is different from the Balto-Finnic type, being, mostly, *object-oriented*, that is, based on properties of the object, such as its animacy, its referential status, and the overall information structure of the sentence, whilst the verbal categories play a secondary role. Many scholars reconstruct for proto-U a model of DOM with 'definiteness' as the most relevant factor controlling it; however, as mentioned, the systems of DOM can be very different in the modern U languages (even within the dialects of one and the same language, as in Permic), whereby, in many of them definiteness is relevant only partially. The detailed analysis of DOM systems in the languages in question leads to the conclusion that specific semantic, morpho-syntactic and prosodic features proper of DOM must be considered as innovative features. It can be speculated that such features (that is, the specific forms of DOM) are subject to recent and very fast language changes and, therefore, can only be used with caution in reconstructing a proto-U DOM system – not to count the fact that this phenomenon is present also beyond the U family, for example in the Turkic languages.

Zanchetta, Elisa; University of Padua (PhD)

The 'impossible' comparison between the creatures and figures of the Hungarian folklore and the Finnish mythology

Unlike the other U peoples, the Hungarians did not develop a folk belief rich in mythical beings, this trait being even more evident when we compare the Hungarian mythology with the complex, rich Finnish / Finnic mythology – the few existing overlaps can be easily explained as recurrent, universals themes of the mythology of the world. It is this very 'impossible' comparison between the Hungarian and Finnish mythological world that the author is going to explore in her contribution to the seminar/volume, arguing for the need to try to 'reconstruct' the Hungarian mythological world. The study will point out to some contradictions that are still unresolved at the current stage of research within U studies, since Hungarian appears to be an 'isolate' within the family from the perspective of its mythological world, too. Certainly, it can be easily objected that the introduction of the Christian faith among the Hungarians (after their settling in the Carpathian Basin, and oth-

er factors), must have eradicated most of their early ‘pagan’ beliefs. Whatever the case, the reality is that the available Hungarian folklore material is rather scanty. Thus, as anticipated, the aim of this article is to emphasize the necessity for further research in order to ‘reconstruct’ ancient Hungarian folk beliefs, drawing also from the language of children and folk tales. According to the author, in fact, Hungarian mythology, or, better, what can be ‘reconstructed’ about it, will contribute significantly to philological, anthropological and linguistic research, that, in turn, will help to clarify the position of Hungarian within the U family.

Angela Marcantonio

References

- ABONDOLO, D. (ed.). 1998. *The Uralic Languages*. London: Routledge.
- ABONDOLO, D. Introduction. In D. Abondolo (ed.). 1-42.
- BENCE, G.I. 2014. Legújabb viták az altaji nyelvek rokonságáról. *Nyelvtudományi Közlemények*, 112: 1-5.
- BEREND, N. et al. 2013: *Central Europe in the High Middle Ages: Bohemia, Hungary, and Poland, c. 900 - c. 1300*. Cambridge University Press.
- COMRIE, B. 1989. *Language Universals and Linguistic Typology*. University of Chicago Press.
- COMRIE, B. 2008. Inflectional morphology and language contact, with special reference to mixed languages. In P. Siemund & N. Kintana (eds.), *Language Contact and Contact Languages*. Amsterdam: Benjamins. 15-32.
- COMRIE, B. 2010. The role of verbal morphology in establishing genealogical relations among languages. In L. Johanson & M. Robbeets (eds.), *Trans-eurasian Verbal Morphology in a Comparative Perspective: Genealogy, Contact, Chance*. Turcologica 78. Wiesbaden: Harrassowitz. 21-33.
- CROFT, W. 1990. *Typology and Universals*. Cambridge University Press.
- CSÚCS, S. 2008. Hozzászólások Angela Marcantonio előadásához. In K. Czeglédi (ed.), *Hozzászólások Angela Marcantonio 2007. október 26-i budapesti előadásához*. *Eleink*, 1/13: 59-72.
- GARDANI, F., ARKADIEV, P. & AMIRIDZE, N. 2015. Borrowed morphology: An overview. In F. Gardani, P. Arkadiev & N. Amiridze (eds.), *Borrowed Morphology*. Berlin: Mouton de Gruyter, 1-26.
- HELASVUO, M.-L. & CAMPBELL, L. 2006. Introduction: Grammar from the human perspective. In M.-L. Helasvuo & L. Campbell (eds.), *Grammar from the Human Perspective: Case, Space and Person in Finnish*. Amsterdam: Benjamins. 1-25.

- HONTI, L. (ed.), 2010. *A nyelvrokonságról. Az török, sumer, és egyéb áfium ellen való orvosság.* Budapest: Tinta Kiadó.
- ITKONEN, E. 2010. A családfa és /vagy a nyelvek közti kapcsolatok a modern összehasonlító nyelvtudományban: a finnugor nyelvek rekonstrukciójának néhány problémája. In L. Honti (ed.). 271-278.
- JANHUNEN, J. 1981. Uralilaisen kantakielen sanastosta. *Mémoires de la Société Finno-Ougrienne (Suomalais-Ugrilaisen Seuran Toimituksia)*, 77: 219–274.
- JANHUNEN, J. 1998. Samoyedic. In D. Abondolo (ed.). 457-79.
- JANHUNEN, J. 2001. Indo-Uralic and Ural-Altai: On the diachronic implications of areal typology. In Ch. Carpelan, A. Parpola & P. Koskikallio (eds.), *Early Contacts between Uralic and Indo-European: Linguistic and Archaeological Considerations*. Helsinki: Mémoires de la Société Finno-Ougrienne (Suomalais-Ugrilaisen Seuran Toimituksia), 242. 207-219.
- JANHUNEN, J. 2007. Typological expansion in the Ural-Altai belt. *Incontri Linguistici* 30: 71-83.
- JANHUNEN, J. 2009. Proto-Uralic – what, where, and when? in J. Ylikoski (ed.), *The Quasiquicentennial of the Finno-Ugrian Society*. Mémoires de la Société Finno-Ougrienne (Suomalais-Ugrilaisen Seuran Toimituksia), 258. 57-78.
- JANHUNEN, J. 2012. Non borrowed, non-cognate parallels in bound morphology: Aspects of the phenomenon of shared drift with Eurasian examples. In L. Johanson & M. Robbeets (eds.), *Copies Versus Cognates in Bound Morphology*. Leiden / Boston: Brill. 23-46.
- JANHUNEN, J. 2014. Ural-Altai: The polygenetic origins of nominal morphology in the Transeurasian zone. In M. Robbeets & W. Bisang (eds.), *Paradigm Change in the Transeurasian Languages and Beyond*. Studies in Language Companion Series 161. Amsterdam: Benjamins. 311-335.
- JOKI, A.J. 1973. *Uralier und Indogermanen. Die älteren Berührungen zwischen den uralischen und indogermanischen Sprachen.* Mémoires de la Société Finno-Ougrienne (Suomalais-Ugrilaisen Seuran Toimituksia) 151.
- KOIVULEHTO, J. 1991. *Uralische Evidenz für die Laryngaltheorie.* Wien: Sitzungsberichte der Österreichischen Akademie der Wissenschaften, Philosophisch-historische Klasse 566.
- KOIVULEHTO, J. 1992/93. *Suomen kielen indo-eurooppalaiset lainasanat ja indo-eurooppalais-uralilainen alkusukuongelma.* Helsinki: Academia Scientiarum Fennica Vuosikirja.
- KORHONEN, M. 1996. *Typological and Historical Studies in Language. A memorial Volume published on the 60th anniversary of his birth.* Edited by T. Salminen. Helsinki: Mémoires de la Société Finno-Ougrienne (Suomalais-Ugrilaisen Seuran Toimituksia) 223.
- LAAKSO, J. 2010a. Mítoszok a finn nyelv ősiségéről. In L. Honti (ed.). 319-32.
- LAAKSO, J. 2010b. Contact and the Finno-Ugric Languages. In R. Hickey (ed.), *The Handbook of Language Contact*. Oxford: Wiley-Blackwell. 598-617.
- LEHTISALO, T. 1936. *Über die primären ururalischen Ableitungssuffixe.* Mémoires de la Société Finno-Ougrienne (Suomalais-Ugrilaisen Seuran Toimituksia), 72.

- MARÁ CZ, L. 2012: Toward a theory of 'Ural-Altai' isoglosses: A research agenda for a Eurasian linguistic area. *Turkic World Almanac*. Astana: Turkic Academy of Science. 93–102.
- MARCANTONIO, A. 1988. On the case of the Object in Finnish: a typological, diachronic and comparative analysis. *Finnisch-Ugrische Forschungen*, 48: 130-170.
- MARCANTONIO, A. 1994. Double conjugation *vs* double marking for the object in some dialects of Vogul. *Études Finno-Ougriennes*, 25: 19-42.
- MARCANTONIO, A. 2002. *The Uralic Language Family: Facts, Myths and Statistics*. Oxford/Boston: Blackwell.
- MARCANTONIO, A. 2016. Magyar-türk nyelvi korrelációk: kölcsönzés, genetikai örökség vagy valami más? Paper presented at the conference: *A nyelvtörténeti kutatások újabb eredményei IX, 2016. április 27-28*. University of Szeged, Hungary. [Abstract & power point presentation available in: <http://sites.google.com/site/angelamarcantonio/Home>].
- MARCANTONIO, A. 2017a. Magyar-uráli és magyar-török morfológiai megfelelések a tudomány jelenlegi állása szerint. In Ferenczné Szócs, É., Gazda, J. & Szabó, E. (eds.), *Kőrösi Csoma Sándor Gyökereink*. Kovászna: Kőrösi Csoma Sándor Közművelődési Egyesület. 85-102.
- MARCANTONIO, A. 2017b. Finnish and Uralic studies in Finland and elsewhere: their contribution to linguistic knowledge. In U. Tuomarla, I. Piechnik & B. Bíró (eds.), *Finland Suomi 100: Language, Culture, History*. Kraków: Jagiellonian Library & University of Helsinki. 71-83.
- MARCANTONIO, A., NUMMENHAHO, P. & SALVAGNI, M. 2001. The 'Ugric-Turkic battle': a critical review. *Linguistica Uralica* 2: 81-102.
- METSMÄGI, I., SEDRIK, M. & UUSKÜLA, M. (eds.) 2014. *Uralic and Indo-European language contacts. Proceedings of the Uralic and Indo-European Language Contact Conference: Uralic 2013*. Tallinn: The Institute of Estonian language, 21st November, 2013. *Journal of Estonian and Finno-Ugric Linguistics (Eesti ja soome-ugri keeleteaduse ajakiri; special issue)*.
- NANOVFSZKY, Gy. (ed.) 2004. *The Finno-Ugric World*. Budapest: Teleki László Foundation.
- POMOZI, P. 2004. The origin of the Mari peoples. In Gy. Nanovfszky (ed.). 457-468.
- ROBBEETS, M. 2010. Trans-Eurasian: Can verbal morphology end the controversy? In L. Johanson & M. Robbeets (eds.), *Transeurasian Verbal Morphology in a Comparative Perspective: Genealogy, Contact, Chance*. Turcologica 78. Wiesbaden: Harrassowitz. 81-114.
- RÓNA-TAS, A. & BERTA, Á. 1011. *West Old Turkic: Turkic Loan Words in Hungarian*, I-II. Turcologica 84. Wiesbaden: Harrassowitz.
- SALMINEN, T. 1999. Problems in the taxonomy of the Uralic languages in the light of modern comparative studies. Paper presented at the seminar: *Tradition und Innovation in der modernen Komparatistik*, Finnisch-Ugrisches Seminar der Georg-August-Universität; Göttingen, 19-21 November, 1999. <http://www.helsinki.fi/~tasalmin/kuzn.html>.

- SAMMALLAHTI, P. 1988. Historical phonology of the Uralic languages (with special reference to Samoyed, Ugric and Permic). In D. Sinor (ed.), *The Uralic Languages. Description, History and Foreign Influences*. Handbook of Uralic Studies I. Leiden: E.J. Brill. 478-554.
- SIPŐCZ, K. 2006. A magyar mint uráli nyelv. In F. Kiefer (ed.), *Magyar nyelv*. Budapest: Akadémiai Kiadó. 288-314.
- TAAGEPERA, R. 1999. *The Finno-Ugric Republics and the Russian State*. London: Routledge.
- THOMASON, S. & KAUFMAN, T. 1988. *Language Contact, Creolization, and Genetic linguistics*. Berkeley: University of California Press.

A 'steppe nomadic culture' vs a 'forest language': Modern identity dissonance in the history of the Magyars

Giuseppe Cossuto

1. Introduction

Herodotus (c. 484-425 BC), the father of historiography, has been the first historian to leave written records concerning populations that have been associated with Uralic (U) tribes and peoples (Rawlinson 1909: IV/16-36). According to his *Histories*, in the lands belonging to the *Sauromatae – non-Scythians* people that nevertheless practised a steppe people lifestyle – in the territories located north of the Tanais (Don) River, the *Budini* peoples also lived. To the north-east of these people, the peoples named *Thyssagetae* lived, and, next to the latter, there were the *Iyrcae* peoples (Rawlinson, 1909: IV/ 21-22). The Iyrcae inhabited thick forests and were hunters, but their system of practising hunting was very peculiar: they hunted on a horse, together with dogs, and using bow and arrows; compare the following quote from Rawlinson (1909: IV/ 21-22):

When one crosses the Tanais, one is no longer in Scythia; the first region one crosses is that of the Sauromatae, that, beginning at the upper end of the Palus Maeotis, stretches northward a distance of fifteen days' journey, inhabiting a country which is entirely bare of trees, whether wild or cultivated. Above them, possessing the second region, dwell the Budini, whose territory is thickly wooded with trees of every kind. Beyond the Budini, as one goes northward, first there is a desert, seven days' journey across; after which, if one inclines somewhat to the east, the Thyssagetae are reached, a numerous nation quite distinct from any other, and living by the chase. Adjoining them, and within the limits of the same region, are the people who bear the name of Iyrcae; they also support themselves by hunting, which they practise

in the following manner: the hunter climbs a tree, the whole country abounding in wood, and there sets himself in ambush; he has a dog at hand, and a horse, trained to lie down upon its belly, and thus make itself low; the hunter keeps watch, and when he sees his game, lets fly an arrow; then mounting his horse, he gives the beast chase, his dog following hard all the while. Beyond these people, a little to the east, dwells a distinct tribe of Scyths, who revolted once from the Royal Scythians, and migrated into these parts

The ethnic name *Iyrcae* can be easily connected, because of superficial similarity, with the ethnic and /or toponymic names *Ugra ~ Yugra*, that is, the Yugra region. It is in fact well known that, throughout the centuries, many Hungarian scholars have tried to link the people more or less contemporary to the early Magyars with those listed by Herodotus, particularly with those that “inhabited the forest” (Kristó 1996: 7 ff.; see also the article by Obrusánszky in this volume). However, Herodotus himself (Rawlinson, 1909: IV/110-117), highlights that the Budini¹ live in the territory ‘belonging to the Sauromatae’², an observation that reveals an important fact: peoples of different culture and economical attitude may coexist within a political confederation ruled by steppe peoples; in turn, this means that their culture is not build up of hermetically sealed compartments, as the Sauromates’ ‘myth of origin’ itself demonstrates. This myth, preserved by Herodotus himself, reports that the Sauromates were peoples consisting of the union between the Scythians and the Amazons, and that the Sauromates spoke the same language of the Scythians, but in an adulterated form, because the Amazons never learn to speak the language of their Scythian men.

2. ‘Steppe people’ language vs ‘forest people’ language

The Scythians (Cossuto: 2012: 25 ff.) are ancient nomadic warriors and breeders, considered by the Hungarian authors of the oldest

¹ We have scarce information about these people; however it is known that they had red hairs and blue eyes, lived also around great, fortified wooden cities (inhabited by the *Geloni* peoples), and had a nomadic life-style, based on the breeding of reindeer in the thick forest areas surrounding nowadays Samara, a city located on the Volga region; see Cossuto (2012: 16 ff.).

² The Sauromates practised a nomadic steppe culture, moving about from the Black Sea eastward up to and beyond the Volga region. They gradually dominated a huge territorial area extending from the 7th century BC until the 3rd century AD; see Cossuto (2012: 37 ff.).

Hungarian *Chronicae* as their own ancestors – these authors being fully converted to Catholicism and well immersed in the cultural sphere of Western Europe civilization. Indeed, the author of the *Gesta Hungarorum*, the so-called 'Anonymous Magister P.' (12th century), as well as Simon de Kéza, the literate clerk at the Hungaro-Cuman Court of King Ladislaus (1282-85; see Tempesti (1969: 12 ff.)), as well as Marci de Kalt (Mark Kalti, second half of the 14th century (see Tempesti, *ibidem*) link their contemporary magyar peoples not to the forest hunters Budini, Thyssagetes or Iyrces, but to the Scythians, the steppe peoples (see Rady & Veszprémy: 2010). The reference to *Scythia* and to a Scythian way of life (later on emphasized as 'Hunnic way of life' in Simon de Kéza's work, a historical and mythological reference at the same time) became an 'identity *leitmotiv*' for the Magyars in the centuries to come. As a matter of fact, when the Hungarian language was classified as belonging to the 'forest people' language group (that is, the Ugric branch of the traditional family tree), a lot of passionate discussions took place amongst scholars (see Marcantonio, Nummenaho & Salvagni (2001: 81-102), Marcantonio 2002: 44 ff.). The affinity between the languages of Lapps, Finns and Magyars was proposed, amongst others, by the Jesuit priest János Sajnovics in his work written in 1770: *Demonstratio idioma Ungarorum et Lapponum idem esse* (see Cossuto (2009: 233-34); for a critical analysis see Marcantonio (2002: 45 & 2004)). Sometimes, the two opposite models (steppe people *vs* forest people) have influenced the analysis of the history of the Magyars, causing misleading interpretations of the historical sources as well as the Magyars' own tradition and sense of identity. In any case, we must not forget that, according to Hungarian traditional historiography the *Ungar-s* (Kristó 1996: 15 ff.) of the *Home Conquest* period (that is, the conquest of Pannonia, or, better, the 'return' to Pannonia³), were truly steppe people, coming and going around in the Steppe – indeed, this was their self-identification and as such they were identified by the various, different peoples that had contacts with them. Inevitably, scholars who adopt the forest people language model, should also be able to explain, with convincing evidence, *how*, *when* and *why* the switch from the forest people way of life to the equestrian, nomadic way of life took place. In other words, why the proto-Magyars would have decided to live their own (suppos-

³ The word *reditus* is emphasised by Simon de Kéza to remark the Hun-Hungarian identification.

edly) original habitat, to move miles away and adopt a totally different socio-economical system? When would have they done this, and how did they implement this change? This is, of course, a very important point, that will be taken up again and discussed below. Meanwhile, depriving the Magyars of their Schyitian origin (be this an historical fact or a myth, or a combination of the two) means that the legitimation of their first dynasty would be compromised too, not only within the western European system, but within the wider Eurasiatic scenery.

3. The Hun / Onogur tradition

The Arpadians⁴ were interested in establishing for themselves a genealogy more ancient than that of the Scythians, universally accepted as *nobilissima gens*. They believed that their dynasty originated directly from Attila, and, before him, from a far away past, going back all the way to the biblical *Japhet*. This remained the case also after the abandonment of the steppe people value system, since the time of the founding of the Hungarian Kingdom on behalf of King (/Saint) Stephan, in the year 1000 or 1001 AD (Stephan ruled between 997 and 1038). The Attila's ancestry was, at the time, the only possible one, and the only one able to match the 'Onogur tradition' with the Biblical one and with the tradition of the erudite and literate people of Western Europe (see Kristó 1996: 75). The analysis of the name *Attila* has been put forward by Pritsak (1956: 404-419; 1982: 444), who claims that this name could be a political title, meaning 'the oceanic, universal [ruler]'. Clearly, this consideration puts Attila on a higher position with respect to his own ancestors and these ancestors' conquests. As Pritsak (1965: 389) puts it:

The Hungarian-Hun tradition (Attila) is not the product of bookish inventions by humanists, but was based on the continuous traditions of the *Onogur-Bulgar* component of the would-be Hungarians, who stayed in Pannonia till 805 and called themselves *Onogur-Bulgars*, the direct descendants of *Irnich*, the youngest son of Attila

Kristó (1996: 81) agrees with Rona-Tas' (1988: 128) and Pritsak's claims regarding the 'Magyar-Hun legitimation' issue; nevertheless Kristó (1996: 81) supports the 'Ugric' linguistic model. Kristó (*ibidem*) also sug-

⁴ The dynasty descending from Árpád (probably 895-907) ruled the Kingdom of Hungary from the 11th century until 1301.

gests a possible time and location for the (assumed) transition from the 'forest' to the 'steppe' way of life. In other words, Kristó assumes that the proto-Magyars became steppe peoples after their encounter with the Bulgars. He also proposes that the 'still Finno-Ugric' proto-Magyars acquired their legitimacy myth (as discussed above) during their encounter with the so-called Volga Bulgars (that, in turn, followed the steppe traditions of the Huns). However, the assumption of a proto-community that, for no obvious reasons, leaves its own original environment and moves for thousands of miles to become a proper steppe people, needs some sort of evidence – evidence that is not provided either by Kristó or by any other scholars – as far as I know. Modern 'fix identity indicators' for peoples, such the language, are not necessarily correlated – in the steppe world – with the socio-economical structure of the peoples in question, because the steppe peoples are continuously exposed and permeated by every possible change determined by the inclusion (or, exclusion) of cultural and materials elements of every type. This, in turn, generates other countless problems.

4. The legitimation of the Árpád dynasty

The specific case of the legitimation of the Árpád dynasty through Attila's son *Ernak* (Ἡρνάχ), raises again a fundamental question in connection with the 'Ugric-Turkic battle', disconnecting the linguistic identity of the dynasty from their original socio-economic structure. Ernak has often been identified with the mighty *Irnik* in the *Nominalia of the Bulgarian Khans*⁵, and he is considered to be the descendant of *Dulo*, noble steppe aristocrat and leader of the Bulgars, starting approximately from 437 AD (see Cossuto (2014: 39-43); Pritsak (1982: 428-476) and Golden (1992: 88-90)). Ernak was the leader of various populations living around the Black Sea (after the death of his brother *Ellak*, a son of Kreka-Akryan, a Hunnic wife of Attila), and ruler of the *Akatzir-s* (Ἀκάτζιροι) and other population (see below). Ellak (whose name seems to be a noble title, too, according to Pritsak (1982)), was the legitimate ruler during the bloody battle of *Nedao*, a river whose exact location is still uncertain, although probably in Pannonia (454 AD).

⁵ *The Nominalia of the Bulgarian Khans* is a short, Old Slavonic manuscript dated around the 9th -11th centuries, containing the names of the presumed first (proto-) Bulgar kings.

Following Attila's death, the various components of the Hun Confederation (mainly Ostrogoths and Gepids), unwilling to accept the rule of the Attila dynasty and the Huns, succeeded in catalysing the Hunnic forces and others loyal groups. Meanwhile, the Gepids of Ardaric (around 455 AD) demanded and obtained from the Eastern Roman Emperor Marcianus the region of Transylvania (Jordanes (*Getica* XII) calls this territory *Dacia*). The Ostrogoths of Walemir demanded the Pannonia region from the same Emperor, whilst other, sizeable groups of Huns, with their allies the Sauromates (the steppe people cited by Herodotus; see above) and the *Cemandres*, settled themselves around *Castrum Martena* (Marburg / Maribor, on the Drava river, in nowadays Slovenia⁶). The surviving heirs of Attila, such as Ernak, loosing the territories that they considered as their own legitimate land (see below), had been forced to take refuge in the area of the mouth of the Danube, a place that Jordanes (*Getica* LII) remarked as being their ancestral land. Jordanes (*ibidem*) also reports an interesting Hun place name: *Hunnivar*, more exactly:

... eas parte Schytiae peteret, quas Danubii amnis fluenta praetermeant, quae lingua sua *Hunnivar* appellant⁷

Pritsak (1982: 429) identifies the word '*Vr*' with the river Dnieper, in nowadays Ukraine. Ernak, after obtaining the authorization of Marcianus, settled in *Scythia Minor*, *grossomodo* nowadays Dobrogea, between Bulgaria and Romania. Under the rule of Ernak, the Huns became again very powerful and dangerous. *Hormudak*, another of Attila's sons, conquered *Serdica*, nowadays *Sofia*, at the time the capital of the *Dacia Mediterranea* province, but the territory that the Huns would have liked to recover at the time is huge, and their enemies were too many. Ernak succeeded in sending his cousins *Emnezar* and *Uzindur* to *Dacia Ripensis* (south of the Danube, with capital *Ratiaria*, c. 30 miles from Vidin – north-western Bulgaria), with the purpose of joining the Huns of *Uto* and *Iscaimus*. The tribes belonging to these two chiefs, after some time, became sedentary peoples, whilst also taking tribal names of Latin origin: *Sacromontisi* and *Fossatisii* (Jordanes, *Getica* L). *Fossatum* means 'a ditch used for fortification', from which also the old Romanian word

⁶ This identification has been proposed by Bartolini (1991: 167, note 113).

⁷ "...Reaching those lands of Schytia, located beyond the Danube, that they in their language call *Hunnivar*".

for 'auto-protected village', *fsat-sat*, derives (Russu 1981: 228). As to the Huns that remained in Pannonia, a single tribal name has remained, *Sadagi*. These various remaining tribes were persecuted by the Ostrogoths. *Dintzik* (~ *Dengizik*; see Pritsak 1982), yet another son of Attila, in order to help them (perhaps ruling the tribes in question together with *Ernak*), assembled a very powerful Hun army composed of the *Ultinzur*, *Angiscir*, *Bittugur* and *Bardor* tribes, and attacked the Goths in their land (the territory reclaimed by Huns). However, they were defeated by the Ostrogoths. After their victory, the Ostrogoths started to oppress the *Scire-s*, a Schyrtian tribe of difficult ethnic identification, from whom Flavius Odoacer (son of *Edeko*, an important Hunnic noble man of the Attila's entourage), will emerge – a loyal ambassador of Attila (see Reynolds & Lopez (1946: 36-53); Martindale *et al.* (1980: 791-793); Macbain (1983: 323-327); Scrofani (2013) and Cossuto (2009: 79-80)). On the year 466, *Ernak* and *Dengizik* together sent ambassadors to Constantinople asking for a commercial emporium for the Huns and the Byzantines. Probably they asked to have access to grazing lands, too, because a war had broken up between them and the tribe of the *Saraugur-s*. The *Saraugurs* (Σαράγουροι) were an *Ogur* confederation (originating from Western Siberia and the Kazakh steppe), that was pushed by the *Sabir-s* up to the Northern Caucasus – they have been considered as one of the first *Ogur* Confederation that penetrated into the Pontic Steppe. The *Saraugurs* attacked the *Akatzirs* and the other tribes of the Hunnic Confederation. On the year 463 their ambassadors were received by the Byzantine Emperor together with those of the *Ogurs* and *Onugurs*. On the year 469 they asked and received protection from Byzantium. At the end of the 5th century, the *Saraugurs*, *Kutrigur-s*, *Utigur-s* and *Onogurs* still formed separate confederations in the steppe north of the Black Sea, whilst the *Akatzirs* disappeared, absorbed mainly by these very confederations (see Golden (1992: 92-97) regarding the language(s) spoken by these confederations). Because of this movements of peoples, *Dengizik* was forced to pass the frozen Danube to drive away the Goth farmers and to occupy their lands. But the Byzantines rounded up and organized the Goths and defeated *Dengizik*, forcing the surviving, hungry and impoverished Huns to break up and disperse themselves into little powerless groups, at first. Afterwards, thanks to the military operations of a Byzantine general of Hunnic origin, *Kalkal*, the Byzantines exterminated the Huns, bringing the head of *Dengizik* in triumph to Constantinople, in 469 AD (see Cossuto 1999: 78-79).

Ernak remained alone, and there aren't safe literary sources regarding his fate, at this point. Nevertheless, the memory of his figure was preserved by the Huns that survived, and he is identified as the man that would bring back Attila's Huns to the place where they lived before, namely Pannonia II (that is, the Roman administrative region), according to the shamanistic prophecy recorded by Priscus of Panion. Maenchen-Helfen (1973: 89) is of the opinion that this Roman territory was given to *Ruga*⁸, the Hun ruler belonging to the Attila dynasty (and probably his uncle), who ruled in a 'diarchy' in the eastern part of the Hunnic confederation until 430 (the western part of the confederation was ruled by his brother *Octar*). After this date, *Ruga* governed alone, and until 435 (see Maenchen-Helfen (1973: 86-88) and Cossuto (2012: 45)).

5. Not only 'pure steppe people' amongst the last Huns

Those Huns that survived the massacre probably had amongst themselves (and, surely, in their Black Sea background) a sizeable percentage of Akatzirs (see above), a numerous warrior population (still existing during the 6th century, when Jordanes wrote his *Getica*), who were not Bulgars at that time, as shown by the following quote from Jordanes (*Getica* V)⁹:

Quidus in austros adsedit Agazzirorum, fortissima, frugum ignara, quae pecoribus et venationibus victitat. Ultra quos distendundur supra mare Ponticum Bulgarorum sedes, quos notissimos peccatorum nostrorum mala fecere

These Akatzirs had given hard time to the Hun rulers before Attila, but had been definitively subdued by Attila after a fierce war and had been given as *apanage* to Attila's son Ellak (Priscus, fragment 8). The Akatzirs did not know agriculture, but were hunters and breeders. Busagli (1986: 108-111) links these people to a Turkic language speaking groups that inhabited the Altai forest and were experts in metallurgy,

⁸ See Carolla (2008), who claims that this land was the first land legally granted on 433 AD by the Roman Consul *Aetius* (with imperial mandate) to *Ruga*.

⁹ "On the Southern part of these [the *Itemest-s*] live the Akatzirs, a very warmongering people, that doesn't know agriculture, and are shepherds and hunters. And beyond, until the Black Sea, there are the Bulgars, very famous for their terrible actions, caused by our sins".

the *Ağaçeri*, namely the 'people of the trees', or the 'people of the forests' (see Sinor 1990: 191). There is here no room to explore Bussagli's thesis of the assumed identification between the Akatzirs of Priscus and Jordanes and the Ağaçeri; however, it is important to remark that these Ağaçeri people actually existed, and spoke a Turkic language. After the death of Ellak, 'his' Akatzirs ended up under the rule of Ernak. They were neither 'original' Attila's Huns nor Bulgars, at the time when Jordanes wrote his work; on the contrary, they were shepherds (and *not* horse breeders) and hunters, they did not know agriculture and only during Attila's ruling time they became part of the Hun confederation.

This brief example clearly demonstrates that the adherence to a socio-economical system – such as, in our case, a 'forest oriented' or 'steppe oriented' system – does not totally or necessarily impact on the language of a given community, at least not amongst the steppe people of Eurasia. In other words, the important thing for the Arpadians (and their followers), like for any other steppe dynasty, was that the dynasty itself was *not* to be identified with the language(s) they spoke, but, rather, with a very complex socio-economical / cultural / value system, mainly based on the legitimacy of the dynastic right to lead the various tribes of a confederation and to occupy and use specific areas of land – this is indeed the tradition that the Magyars elaborated for themselves during the centuries, tradition that has, nevertheless, true historical basis.

If Ellak and Ernak actually ruled also over a powerful group of shepherds and hunters, the Akatzirs, it is also amongst these people (later on absorbed by other groups) that the Hunnic tradition has been preserved. Similarly, the legitimacy of these steppe aristocracy to stay in the lands belonging to the Attila's dynasty – centuries before the 'return' (*reditus*) of the Magyars to the Danubian area as 'steppe people' – has been preserved with a wealth of mythological elements, until historical times (all these cultural elements being collected and harmonized in the Hungarian *Gesta*, redacted centuries after the adoption of Catholicism).

6. Conclusion

In the complex steppe world there were, basically, two main groups of peoples: the 'Forest People' (Akatzirs), that preserved the Hun 'legitimacy tradition', and the Hunnic people, that changed their name (into Fossatisii and Sacromontisi) when they became a sedentary people – in addition to a few other ethnic and linguistic mixtures. Thus, in

order to solve the (apparent?) contradictions embedded in the conventional model of the ‘origin’ of the Hungarian language and people, we need to involve other disciplines in our research (history, archaeology, ethnography, etc.), rather than limiting ourselves to the sphere of linguistics. In other words, we need to understand the socio-economical system, the *Weltanschauung* and the rules of the steppe people and aristocracy (in turn based on the practice of equestrian nomadism), rather than try to explain their intricate ethno-linguistic relations by investigating their language(s) in isolation, cut off from their unalienable historical and social context.

References

- BARTOLINI, E. (ed.). 1991. *Jordanes. Storia dei Goti*. Firenze: TEA Editore.
- BURY, J.B. (ed.). Priscus (fragment 8): <http://sourcebooks.fordham.edu/halsall/source/priscus1.asp>
- Bussagli, M. 1986. *Attila*. Milano: Rusconi.
- CAROLLA, P. 2008. Priscus Panita, *Excerpta et fragmenta*. Berlin /New York: De Gruyter.
- COSSUTO, G. 2009. *Tracce “turche” in Europa Medievale. I popoli delle steppe in Europa dalla comparsa degli Unni alla nascita della Turchia*, Roma: Aracne.
- COSSUTO, G. 2012. *Europa e nomadi guerrieri. Gli sciti, gli unni e altri popoli delle steppe alle origini dell’Europa*, Roma: Chillemi.
- COSSUTO, G. 2014. *I Signori del Danubio. I nomadi guerrieri e l’Europa alto-medievale*, Roma: Chillemi.
- GOLDEN, P.B. 1992. *An introduction to the History of the Turkic peoples: ethnogenesis and state formation in medieval and early modern Eurasia and the Middle East*, Wiesbaden: Harrassowitz.
- KRISTÓ, Gy. 1996. *Hungarian history in the Ninth Century*, Szeged: Szegedi Középkorász Műhely.
- Macbain, B. 1983. Odovacer the Hun? *Classical Philology*, 78 /4: 323-327.
- MAENCHEN-HELFEN, O.J. 1973. *The World of the Huns: Studies in Their History and Culture*, Oakland: University of California Press.
- MARCANTONIO A., NUMMENAHO P., SALVAGNI M., 2001. The “Ugric-Turkic battle”: a critical review. *Linguistica Uralica*, 2: 81-102.
- MARCANTONIO, A., 2002 *The Uralic Language Family: Facts, Myths, and Statistics*, Cambridge University Press.
- MARCANTONIO, A. 2004. The role of János Sajnovics in comparative linguistics: a critical review. In P. Klesment (ed.), *Myths and Facts in Uralistics. Fenno-Ugristica* 26:151-169. [Hungarian version in A. Marcantonio 2006. *A történelmi nyelvészet és a magyar nyelv eredete. Angela Marcantonio válogatott tanulmányai*. Budapest: Hun-idea. 55-87].

- MARTINDALE, J. et al. 1980. *The Prosopography of Later Roman Empire, II: AD 395-527*. Cambridge University Press.
- PRITSAK, O. 1956. Der Titel Attila. In M. Woltner & H. Bräuer (eds.), *Festschrift für Max Vasmer zum 70-Geburtstag: am 28. Februar 1956*. Berlin: Harrassowitz. 404-419.
- PRITSAK, O. 1965. Yowar und Kawar. *Ural Altaische Jahrbücher*, 36: 390-93.
- PRITSAK, O. 1982. The Hunnic Language of Attila clan. *Harvard Ukrainian Studies*, VI/4: 428-476.
- RADY, M. & VESZPRÉMY, L. 2010. *Anonymus, Notary of King Béla: The Deeds of the Hungarians*, Budapest: Central Europe University Press.
- RAWLINSON, G. 1909. *Herodotus: the Histories*, <http://classics.mit.edu/Herodotus/history.4.iv.html>
- REYNOLDS, R.L. & LOPEZ, R.S. 1946 'Odoacer: German or Hun? *The American Historical Review*, 52/ 1: 36-53.
- RÓNA-TAS, A. 1988 "Ethnogenese und Staatsgründung. Die türkische Komponente in der Ethnogenese des Ungartums", in: *Studien zur Ethnogenese* Bd. 2, Rheinisch-Westfälische Akademie der Wissenschaften, Abhandlung 78, Opladen. 107-142.
- RUSSU, J.I.I. 1981. *Etnogeneza Românilor*, Bucharest: Editura Științifică și Enciclopedică.
- SAJNOVICS, J. 1770. *Demonstratio idioma Ungarorum et Lapponum idem esse*. Tyrnavia. http://reader.digitale-sammlungen.de/en/fs1/object/display/bsb10524231_00022.html
- SCROFANI, G. 2013. *sub voce 'Odoacre'*. *Dizionario Biografico degli Italiani*, 79. [http://www.treccani.it/enciclopedia/odoacre_\(Dizionario-Biografico\)](http://www.treccani.it/enciclopedia/odoacre_(Dizionario-Biografico))
- SINOR, D. 1990. *The Cambridge History of Early Inner Asia*. Cambridge University Press.

Information Structuring and typology: Finnic and Samic word order revisited through the prism of orality

M.M. Jocelyne Fernandez-Vest

1. Introduction

In this article, the presentation of the Information Grammar of Orality (IGOR) theory, a theory originally inspired by several decades of field linguistics among North-Western Uralic communities, and later developed within the ISTY-project at CNRS (Information Structuring and Typology, 2008-2015) will concentrate on 3 arguments:

- 1/ One should take into account 2 kinds of typologies:
 - i) *discourse typology*, based on criteria of genre / context / relations between interlocutors;
 - ii) *language typology*, apt to explain what kind of resources are mobilized in individual linguistic systems for building meaningful constructions.

2/ In order to increase their relevance regarding language use, word order studies should replace the analysis of 'neutral' sentences with that of 'dialogic' utterances taken from various discourse genres.

3/ The methodology lays stress on Answers (A) as primary elements of Information Structuring, the pivot of 2 binary information strategies involving Initial Detachments (IDs) and Final Detachments (FDs), Theme-Rheme and Rheme-Mneme respectively.

2. Information Structuring of oral discourse

Enunciation theories (Benveniste 1966-74, Kerbrat-Orecchioni 2009, among others) have shown that the main characteristic of oral exchanges is the anchoring of the speech. Orality is characterized by a situational

dependency resulting from the co-presence of interlocutors: defining is thus the unavoidable contextualization of oral and spoken languages (Fernandez-Vest 1987: 217-230, 1994: 118-119; Hagège 1986 1993: 3-4). Some recurrent language features can be attributed to this oral motivation: an oral exchange is marked by an improvised construction (lack of time for planning) and a strong redundancy necessary for decoding. Conversely, the typological evolution of an orally transmitted language that acquires the status of a written language proves its gradual 'oral demotivation' (see 3.2.).

2.1. Discourse Particles (DIPs)

In oral and spoken languages, Discourse Particles are generally numerous. This has been proved to be the case of Northern Sami, whose purely oral tradition could still be observed in the 1970s. See for instance this short excerpt of a dialogue between two elderly Sami informants:

- (1) (A) a. [How far is it exactly from here to there / your home in Báđoš / from here / the market place?]
 (B) b. *Goal mo bat dal dat lea? Galhan dat lea vissa ... beannot miilla vai ... gal dat guokte miilla lea gal.*
 c. *Eambo* dat gal lea.
 d. *Gal* dat liikká lea eambo gal.
 e. *Ammal* ... ammal jo VIHTTA miilla gal lea dákko Deatnorái. [...]
 (B) b. How much could it actually be? Yes indeed it is surely ... one and a half miles or what ... yes two miles there is yes.
 c. More it is for sure.
 d. Yes it is though more yes.
 e. Maybe ... maybe even FIVE miles yes from here along the Deatnu. [...]

(Fernandez-Vest 1987: 585-589)

A symptomatic profile of DIPs can be pictured as follows:

- they have no propositional meaning
- they qualify the discourse process rather than the structure
- they anchor the message to attitudes / feelings in an implicit way.

One can also add sub-categorizations, e.g. nuclear vs. peripheral, structural vs. interpersonal DIPs (Fernandez-Vest 1994: 21-34).

2.2. Detachment Constructions (DECs)

Another characteristic of spoken language is its unintegrated syntax, that must be recognized in its own right, and not considered as the result of an erroneous construction (see for instance Miller & Fernandez-Vest 2006; Miller & Weinert 2009: 360-372). The thorough analysis of diverse corpora has brought me to identify an underlying system for the use of what I call 'detachments' (rather than 'dislocations', referring to a generative perspective). But interpreting this system also implies to take into account longer utterances and their co-text, as we shall see below.

A simple definition of Information Structuring (IS) is the following:

Information Structuring studies the impact on their discourse construction of the speaker's and addressee's need to express their assumptions (private and shared knowledge, status of referents, personal appreciation of the situation) through dividing the information into more or less salient and relevant. The chosen strategy, i.e. the ordering and highlighting of constituents that reflect the linguistic encoding of information structuring categories is also determined by the cotextual relations of reference. Consequently, IS should be studied in longer texts rather than in single utterances: thematic chaining and thematic progression are essential components of the IS strategy.

(Fernandez-Vest 1987: 612; 2015: 14)

Our IGOR methodology relies upon a triple organization of the utterance – as recognized by several language theoreticians (Pierce 1978; Daneš 1964; Hagège 1993, among others) – with the main difference that I insist on making a resolutely textual *and* interactional definition of the Theme/Topic ('what is spoken about') and of the Rheme/Focus ('what is said about it'), and on emphasizing the importance in spoken language of the least investigated element, the third one, Mneme in my terminology, a Post-Rheme characterized by formal properties (often a flat intonation) and semantic values (e.g. reference to some supposedly shared knowledge – Fernandez-Vest 1994:197-200, 2004). Mneme differs from Lambrecht's Antitopic (1994:184-191, 2001) in two essential points: IS is relevant at the enunciative level, and should be analyzed also beyond the sentence.

Summing up, this methodology is doubly tripartite: 3 levels – enunciative, morphosyntactic, distributional semantic – and 3 constituents at the enunciative level; it has served to bring out two basic strategies

- the *Binary strategy 1* (Theme–Rheme), with a 1st element frequently detached (Initial Detachment, ID)
- and the *Binary strategy 2* (Rheme–Mneme), where the 2nd constituent is detached (Final Detachment, FD). It is a typical construction for impromptu speech, mostly absent in written style.¹

Combining the two strategies, *circular cohesion* is a recurrent device available for insuring the internal and textual coherence of utterances (Fernandez-Vest 1995, 2009, 2015: 169-188).

This methodology also reverses the Question-Answer test used by many linguists as the focus-triggering context (see Matic & Wedgwood 2013 for a critical evaluation): starting from the Answer as Minimal Communicative Utterance (MCU), we shall investigate its extensions forward and backward, IDs and FDs. Remarkable in my corpora of old Sami was the high proportion of FDs, in simple Answers (Rheme–Mneme) as well as Multiple Answers², e.g.

- (2) [And your parents' house was made of...?]
 – *Hirsa... hirsavisti* =>> *Guda dumá aso* [Rh] *dat hirssat* [FD-Mn].
 'Log... a log-hut =>> Six thumbs thick [Rh] *the logs* [FD-Mn].'

(Fernandez-Vest 1987: 552)

3. DECs from a comparative perspective

3.1. DECs and internal contrastivity

In view of establishing the relation oral vs. written style in a given language, it is extremely difficult to gather dialogic corpora collected in equivalent conditions of enunciation. I conducted between 1982 and 2002 with my main research groups at CNRS (LACITO, Langues et Civilisa-

¹ See some French examples of these 2 binary strategies, borrowed from Hergé's dialogues (*Tintin et les Picaros*, 1976), and their translation in several languages, e.g. *Les fenêtres* [ID], *elles ne s'ouvrent pas* 'The windows [ID], they do not open' / *Ça part vite, ces machins-là* [FD] 'it goes off quickly, these things' (Fernandez-Vest 2015: 198-211).

² Part of the typology of Qs and As originally built upon a Sami corpus (Fernandez-Vest 1987: 443-460), Multiple Answers appear as sequences of utterances linked by a quick tempo (double arrow =>>).

tions à Tradition Orale, then OSTERLITS, Oral/Oural, Structures en Transit et Recherches Linguistiques sur la Traduction du Sens) several experiments about ‘internal’ language registers, including for my part oral vs. written versions of *scientific texts* in French, English and Scandinavian languages. They revealed that the oral presentations were about 30 % longer (DIPs, DECs, repetitions, reformulations) than the corresponding published articles (Fernandez-Vest 1994: 143-158; 2015: 108-123).

Another approach of this relation was centered on the comparison of two versions of a collection of Finnish artists’ private interviews. In spite of the close faithfulness proclaimed by the editor of the collection (published by the Finnish Literature Society, SKS), the impromptu style of the recorded conversations were systematically changed into shorter, logico-syntactic sentences, e.g.:

(3) [– Doesn’t the creativity process evolve with time?]

[ORAL] – *Kyl siin varmasti vähän eri eri eri mekanismi / mekanismi on hiukan ehkä muuttunut / tän* [GEN.SG] *luovan* [GEN.SG] / *prosessin* [GEN.SG] / *mekanismi* [NOM.SG].

‘Yes there surely a little diff different mechanism / the mechanism has slightly maybe changed / *of this creative / process / the mechanism.*’

[WRITTEN] – *Luulisin myös luovan* [GEN.SG] *prosessin* [GEN.SG] *mekanismin* [ACC.SG] *iän mukana muuttuneen* [PAST.PART ACC.SG]

‘I would also believe *that the mechanism of the creative process* along with age *has changed.*’

In the oral version (my transcription), the proportion of IDs and FDs, inversely distributed in monological vs. dialogical parts, is globally balanced; in the edited version, they are reduced to respectively 0,5% and 0%. (Fernandez-Vest 2006: 185-191, 2015: 123-127).

3.2. Languages in transition

In this North-European area, Sami has been an ideal laboratory for following its typological evolution after it was pushed into (/ promoted to) a written culture. Let us remind that in traditional Sami, information strategies shaped by orality were prominent: word or-

der served mainly pragmatic aims, IS and emphasis; the utterance internal articulation and the intrasentential connection were mostly insured by DIPs.

The recent development of the written mode has implied a gradual decline of certain grammatical and semantic categories: spatio-temporal deictics and DIPs. A periodical study of articles collected in the new Sami press has shown the influence of Scandinavian neighboring languages, even in usual reports that are still grounded in dialogic exchanges: adpositions often involve reverse order, and cleft constructions are substituted for thematizing IDs (Fernandez-Vest 2011, 2012a: 83-91). FDs can still be found in the dialogues of fiction works, but their survival is probably a question of relatively short time (Fernandez-Vest 2005, 2009, 2016).

A new collaboration with the Giellagas Institute at Oulu University³ has allowed me to feature for the last four years the typological profile of modern Sami dialogues. In Questions as well as in Answers, the number of DIPs has been steadily reduced, and their semantics restricted: of the most usual dialogic DIPs, *na*, *gal*, *dat* (Fernandez-Vest 1987: 390-407, 443-460; 1994: 57-65), the omnipresent thematizing DIP *dat* has become very rare, the other two have survived with some changes: *na* is generalized as an opening for adjacent pairs (– *Na man...?* – *Na diedus...* ‘Well how...?’ – ‘Well of course...’), *gal* has gained a (confirming, invalidating or strengthening) dominant position.

What about IS? Apart from highly emotional utterances, fronting is blurred by subordinate clauses introduced by the conjunction *ahte* ‘that’ (influenced by its Finnish and Scandinavian equivalents – *Fi. että*, *Sw. att* –?).

Detachment Constructions are relatively rare. IDs can still be found in simple Answers of average length (less than 20 words), e.g.:

- (4) [How was Sami music handled at school in former times?]
 – *Ovdal áigge sámemusihkka skuollas [ID] / dat lei olu uhcanut.*
 ‘In former times Sami music at school [ID] / it was really minimized.’

³ Research stipend no 33-3562 (2013-2016) from the KONE Foundation, Helsinki, «Northern Sami, from oral discourse to written text: Looking for an authentic ethnolinguistic identity». Examples (4)–(6) (rewritten with my own transcription) are borrowed from Giellagas Archives; many thanks to Marko Jouste and Minna Rasmus (Elle Ovlá Elle Máreha Minna) for giving me access to those archives.

FDs occur only twice in a corpus of 4 hours recording, in its most basic form, that is, a mere repetition of the Question-Theme:

- (5) [Is that music an essential part?]
 – *Galhan dat – hui stuorra – oassi – lea dat musihkka* [FD].
 ‘Yes indeed it – a very big – part – is that music [FD].’

Consequently, the most prominent thematizing devices are the cleft constructions, directly inspired by their use in neighboring languages, e.g.:

- (6) – *Dat lea várra dat mii lea buot eanamusat váikkuhan.*
 ‘It is certainly that which has mostly influenced.’

pro *Dat dat* [th.DIP] *lea várra buot eanamusat váikkuhan.*
 in traditional Sami.

(Fernandez-Vest 2017)⁴

Let us propose some explanation for the frequency vs. rarity of detached and other segmented structures in oral vs. written styles.

4. Information Structuring (IS) and oral syntax: the case of Non Finite Constructions (NFCs)

This choice to situate IS at the enunciative level, distinct from the morphosyntactic one (2.2. above), should naturally be more convincing if we can draw some correlations between information units and morphosyntactic constituents, and, still better, if we can show that these correlations involve different constructions in oral vs. written styles. A good example of this is the case of NFCs, also referred to in the literature as “conjunctive converb constructions” (Nedjalkov 1995) – or even gerund, adverbial participle, clause equivalent (Haspelmath & König (eds.) 1995). Exclusive adverbial clauses in Proto-Uralic (supposedly), many of these constructions have been sup-

⁴ For a typological study of Sami Questions and Answers from a cross-linguistic perspective, see Fernandez-Vest *et al.* (2017).

planted in several Uralic languages by conjunctive clauses. This evolution has been generally explained by the pressure of surrounding languages: Uralic languages, that form a continuum extending from fusion to agglutination, have traditionally been divided into an eastern group, which (supposedly) remains close to Proto-Uralic typology, reinforced by contact with Turkic languages, and a western group strongly influenced by Indo-European languages (see Comrie 1988 for the differences of word order). Nenets and Komi-Zyrian are good examples of the changes nowadays ongoing even in the eastern group: the huge differences observed in the speakers' discourse between the use of gerundials vs. subordinate clauses reflect the degree of bilingualism and exposure to Russian syntax (Tereščenko 1965; Leinonen 2002, 2005; Fernandez-Vest 2012b). In the western group, the case of Finnish will provide us with a strong argument: the two types of constructions officially cohabit in the language, but NFCs are generally associated with written style.

4.1. Non Finite Constructions (NFCs) in Finnish

Although nowadays rare in oral style, NFCs are an important chapter of Finnish grammars: related to verbo-nominal polarity (a typological characteristic of the Uralic language family; Setälä & Sadeniemi 1966, Ikola 1978, among others), NFCs are one of the five grades that illustrate the notion of "sententiality", early put forward by Finnish grammarians (Hakulinen & Karlsson 1979), and developed by the *Comprehensive Grammar of Finnish*. NFCs, divided into 5 infinitives and 2 participles, that build original and morphologically complex constructions, lack several of the basic features (modality, verb of negation, thematic word order, personal suffixes...) used for defining the Finnish sentence (ISK 2004: 834, Fernandez-Vest 2008).

Two types of NFCs, relatively frequent in modern Finnish, will help us to explain their preferential connection with written style: the *referative constructions* and the *temporal constructions*.

Regarding the referative construction, example (3) above offered an illustration of the duality available: while the artist's oral Answer was elaborated with uncertainty – repetitions, hesitations, adverbial modulations – and ended with a FD chunked into three segments intended to disambiguate (literal repetition of the Ques-

tion-Theme), the edited version introduces, with admittedly complex morphological devices, a clear Answer (an opinion verb followed by an object clause). This written strategy implies a cognitive reductionism of the speaker's discourse⁵. The global sense of the message was maintained, but the rhythm and successive stages of the speech following a cognitive process of slow, laboured reflection upon an abstract topic disappeared. This strategy, favoring clarity in the Answer, obliterated the speaker's efforts to ponder her judgment and to finally insure the cohesion of her speech which the FD stood for.

Our paraliterary corpus of recorded and transcribed vs. edited interviews also offered several typical examples of the other NFC frequent in written style: the temporal construction, e.g.:

- (7) [Can one distinguish in the creative work exact phases (inspiration, putting into words...) that would take place successively?]
 [ORAL] – *ð Silloin^a kun^b mä^c / korjaan^d jotain kohtaa niin^e [DIP] mä^f tiedän että se^g vaikuttaa kaikkiin muihin ... kohtiin eli / eli^h seⁱ oivallus jonka mä^j saan tätä kohtaa korjatessa^k niin^l [DIP] se^m säteilee niinkunⁿ [DIP] / KOKONAISSUUTEEN niin^o [DIP] [...]*
 'ð Then^a when^b me^c / I correct^d a certain point well [DIP] me^f I know that that^g influences all the other ... points in other words / in other words^h thatⁱ inspiration that me^j I receive this-point-while-I-correct^k well^l [DIP] it^m extends somehowⁿ [DIP] / ON THE TOTALITY well^o [DIP] (...)'

[WRITTEN] *Korjatessani^b työssä jonkin kohdan tiedän^f, että korjaus^g vaikuttaa kokonaisuuteen eli oivallusⁱ säteilee kokonaisuuteen (...)*

'While I am correcting^b in the work a certain point^f, I know that the correction^g influences the whole in other words inspiration extends over the totality (...)'

The impromptu oral version of the speech turn is, as in the previous examples, characterized by hesitations, repetitions, prosodic rather than syntactic devices (e.g. the segmentation placed after the con-

⁵ For a psycholinguistic measuring of the influence of this complexity on syntactic parsing, see Hyönä & Vainio 2001, Fernandez-Vest 2015 (239-241).

nectors), numerous actualizers, numerous DIPs of thematic junction, informative word order.

The written version replaces DEC's by direct constructions or by verbo-nominal NFCs. Since the action of the subordinate clause is simultaneous with the main clause, the temporal construction is here built up with the inessive case of the infinitive II, to which a possessive suffix is added (co-referentiality of the subjects of the two clauses):

- | | | | | | |
|-------|---|------------|-----------|----------------|-----------|
| (a) | <i>silloin</i> | <i>kun</i> | <i>mä</i> | <i>korjaan</i> | (oral) |
| | 'then | when | I | correct' | |
| (b) > | <i>korjate-ssa-ni</i> (b') | | | | (written) |
| | INFII-INESS.SG-POSS1SG | | | | |
| | 'in my correcting', ie. 'while I am correcting' | | | | |

One can see in this example that the NFC is part and parcel of the total strategy intended to eliminate the uncertain phases of the thought elaboration improvised in real time, and thus 'clean out' and synthesize the Answer.

Will a comparison with our still orally shaped Sami language bring some further argument toward an orality-bound interpretation of IS and its syntactic correlates?

4.2. A corpus-based comparison of Finnish and Sami Non Finite Constructions (NFCs)

Northern Sami also possesses NFCs, as complex as their Finnish counterparts, although built upon different morphological bases. The Sami temporal NFC, for instance, expresses the simultaneity of the action between the subclause and the main clause with a special form called Gerund I: *-diin*, *-dettiin* (+ a possessive suffix for co-referentiality of the two subjects). If the hypothesis of NFCs as 'genuine ancient Uralic constructions' held, one should expect that NFCs were frequent in traditional Sami. The study of one of the few bilingual corpora available for comparison, before a unified written standard was gradually imposed upon Sami speakers, contradicts in fact that hypothesis. We shall complete it with later translations.

Let us take a few examples of temporal NFCs and their translated equivalents, that permit to single out a significant number of regular differences between the two languages.

• From Finnish to Sami

Using translated texts again for a relevant comparison, one can notice immediately that, in the recent translation into Sami of a Finnish novel (Mukka 1966, 2008), the Finnish temporal NFC is systematically rendered by a conjunctive clause:

(8a) Fi. *Tull-e-ssa-an minua vastaan portailla tai tiellä, Ulla käyttäytyi nyt niin kuin ennenkin*
 ‘When coming in front of me in the stair case or on the road, Ulla behaved now as before’

(8b) Sa. *Go Ulla bodii mu ovddal ráhpain dahje geainnu alde, de [DIP] láhttii dego tábálaččat*
 ‘When Ulla came in front of me in the stair case or on the road, (then [DIP]) she behaved as usually.’

The IS of the Sami sentence is clearly binary – a thematic clause followed by a rhematic one⁶ – whereas the IS of the Finnish sentence is less transparent: the NFC functions as an adverbial clause rather than a separate Theme-clause.

This systematicity suggests an interpretation in terms of orality – which had to be verified through looking at the reverse translations.

• From Sami to Finnish

This corpus is taken from one of the first text anthologies selected and translated by a team of native Sami speakers (*Skabmatolak*, 1974). Several examples of translated excerpts showed how the systematic binary articulation of the Sami sentence allowed potentially two different interpretations: (i) a Theme-clause + a Rheme-clause, or (ii) a Rheme-clause + a Mneme-clause, whereas the Finnish translation did not leave any room for interpretation, with no possibility of word order variation (within an adverbial clause firmly attached to the subject of the main verb).

A typical example of this, taken from a writer’s short story (describing first experience of high mountain skiing) is the following:

⁶ For the extension of our tripartite IS methodology to complex utterances and their integration into whole texts, that cannot be developed here, see Fernandez-Vest 2015: 34-40, 129-147.

(9a) Sa. *Danin son luoitáge erenoamáš váraid, go dat gobádat lahkana.*
 ‘That’s why he glides down very cautiously, when that depression approaches.’

(9b) Fi. *Siksi hän laskettaakin hyvin varovasti painanteen lähetessä.*
 ‘That’s why he glides down very cautiously at the approaching of the depression.’

(S. Aikio 1974 /1968)

The Sami sentence (9a) allows a variation in the enunciative interpretation of the sub-clause: the temporal clause could be either (i) a *rhetic clause*, that adds a precise information on the time (and place) when the skiing becomes particularly cautious, as announced in the thematic clause (*Danin --- váraid*), or (ii) a *mnemonic clause*, recalling the approaching of the hollow (already mentioned in a preceding sentence) after a rhematic clause telling that the boy becomes more cautious in connection (*Danin*) with his already mentioned fear of steep descents and knowledge of a depression in the vicinity). The Finnish sentence, entirely compact, permits only one reading: the Rheme-sentence one, with a final adverbial NFC. But the main difference, from the point of view of Information Structuring, is that the Sami utterance could also allow an internal segmentation in the subclause, since the depression has already been activated, manifested by the presence of a demonstrative adjective (*dat*):

(9a’)Sa. *Danin son luoitáge erenoamáš váraid,*
 ‘That’s why he glides down very cautiously,
go dat lahkana, (dat) gobádat.
 [Rh-clause] [Mn-FD]
when it approaches, the depression.’

This preference for flexibility, typical of oral enunciation, seems to be an explanation for the rarity of NFCs in Sami, in official comparison with another configurational language, Finnish – whose written tradition has already been installed for... little over 100 years.

5. Conclusion

Appealing to the broader co-text and transposing to a longer complex utterance (/sentence) the criteria defined for the three enunciative constituents of a short basic sentence is not without its risks: this part of the theory will still need further investigation. Nevertheless, taking into account from the point of view of IS the respective typological specificity of oral and written styles of the same language has proved to be useful for explaining several fundamental syntactic differences between the two styles. DEC's offer a perfect illustration of the enunciative process' impact on the syntactic structure.

As for NFCs, things are admittedly more complex: a thorough observation shows that the nature of the NFCs is not identical in Eastern and Western Uralic languages (where NFCs are limited to a few syntactic functions of the landmark word; see Comrie 2005). But, at least in the present state of their evolution, the different degree of orality is also an essential evaluation criterion, which can partly contradict the explanation of a purely Uralic construction substituted with an Indo-European structure.

Besides, I dare claim that the comparison of oral and written corpora in lesser-known languages acquiring a written status could give these languages a chance, not only to 'be saved', as the objective of many international projects is nowadays formulated, but to resist the constant pressure from neighboring languages with already well-established written norms, and thereby to maintain some of their typological characteristics. This short presentation of an orality-centered hypothesis as part of the IGOR theory is a step in that direction.

References

- BENVENISTE, E. 1966-74. *Problèmes de linguistique générale*, 1-2. Paris: Gallimard.
- COMRIE, B. 1988. General features of the Uralic languages. In D. Sinor (ed.), *Handbuch der Orientalistik 8: Handbook of Uralic Studies, 1. Uralic Languages*. Leiden: E.J. Brill. 451-477.
- COMRIE, B. 2005. La typologie des langues ouraliennes, comparées avec les autres grandes familles d'Eurasie septentrionale. In M.M.J. Fernandez-Vest (ed.), *Les langues ouraliennes aujourd'hui. Approches linguistiques et cognitives – The Uralic languages today. A linguistic and cognitive approach* (Bibliothèque de l'École des Hautes Etudes 340). Paris: Editions Honoré Champion. 75-86.

- DANEŠ, F. 1964. A three-level approach to syntax. *Travaux linguistiques de Prague* 1. Prague: Academia. 225-240.
- FERNANDEZ-VEST, M.M.J. 1982 (/1977). *Le finnois parlé par les Sames bilingues d'Utsjoki-Ohcejohka (Laponie finlandaise) – Structures contrastives, syntaxiques, discursives*. Paris: SELAF (L'Europe de Tradition Orale, 1).
- FERNANDEZ-VEST, M.M.J. 1987 (/1984). *La Finlande trilingue. 1. Le discours des Sames. Oralité, contrastes, énonciation*. Paris: Didier Erudition.
- FERNANDEZ-VEST, M.M.J. 1994. *Les Particules Énonciatives dans la construction du discours*. Paris: PUF (Linguistique Nouvelle).
- FERNANDEZ-VEST, M.M.J. 1995. Morphogenèse orale du sens: de l'espace des langues aux objets de discours. In M.M.J. Fernandez-Vest (ed.), *Oralité et cognition: invariants énonciatifs et diversité des langues*. *Intellectica* 20 (1). 7-53.
- FERNANDEZ-VEST, M.M.J. 2004. Mnémème, Antitopic – Le Post-Rhème, de l'énoncé au texte. In M.M.J. Fernandez-Vest & S. Carter-Thomas (éds.), *Structure Informationnelle et Particules Énonciatives – essai de typologie*. Paris: Editions L'Harmattan (*Grammaire & Cognition* 1-2). 65-104.
- FERNANDEZ-VEST, M.M.J. 2005. Information structure and typological change: Northern Sami challenged by Indo-European models. In M.M.J. Fernandez-Vest (dir.), *Les langues ouraliennes aujourd'hui: approche linguistique et cognitive – The Uralic Languages today: a linguistic and cognitive approach*. Paris: Editions Honoré Champion (Bibliothèque de l'École des Hautes Etudes 340). 563-576.
- FERNANDEZ-VEST, M.M.J. 2006. Vers une typologie linguistique du détachement à fondement ouralien d'Europe, *Bulletin de la Société de Linguistique de Paris (BSL)* C1, 1: 173-224.
- FERNANDEZ-VEST, M.M.J. 2008. Subordination et degrés de phrasticité dans quelques langues ouraliennes nord-occidentales: exemples samiques et fenniques. *La Linguistique* 44 (2). 99-116.
- FERNANDEZ-VEST, M.M.J. 2009a. Vers une typologie informationnelle des constructions à détachement. In D. Apothéloz, B. Combettes & F. Neveu (éds.), *Les linguistiques du détachement. Actes du Colloque International de Nancy (7-9 juin 2006)*. Bern / Berlin: Peter Lang (Sciences pour la communication 87). 251-262.
- FERNANDEZ-VEST, M.M.J. 2009b. Typological evolution of Northern Sami: spatial cognition and Information Structuring. In J. Ylikoski (ed.), *The Quasiqui-centennial of the Finno-Ugrian Society*, Helsinki: Mémoires de la Société Finno-Ougrienne (Suomalais-Ugrilaisen Seuran Toimituksia) 258. 33-55.
- FERNANDEZ-VEST, M.M.J. 2012a. *SAMI. An introduction to the language and culture, with a Sami-English-Sami lexicon*. Helsinki: Finn Lectura.
- FERNANDEZ-VEST, M.M.J. 2012b. Typological evolution of information grammar in Uralic. In S. Csúcs (ed.), *Papers from the XIth International Congress of Finno-Ugric Studies; 8-15 August 2010, Piliscsaba (Volume IV)*. Budapest: Reguly Társaság. 39-45.

- FERNANDEZ-VEST, M.M.J. 2015. *Detachments for Cohesion – Toward an Information Grammar of Oral Languages*. Berlin / Boston: De Gruyter Mouton (Empirical Approaches to Language Typology, EALT 56).
- FERNANDEZ-VEST, M.M.J. 2016. Detachment Linguistics and Information Grammar of Oral Languages. In M.M.J. Fernandez-Vest & R.D. Jr. Van Valin (eds.), *Information Structuring of Spoken Language from a Cross-Linguistic Perspective*. Berlin/Munich/Boston: De Gruyter Mouton (Trends in Linguistics. Studies and Monographs, TILSM 283). 25-32.
- FERNANDEZ-VEST, M.M.J., AMON, M., GADELIH, K.E., PAN, V.J., ACHARIYAYOS, J., DO-HURINVILLE, D.T. 2017. Information Structuring of Dialogic pairs from a cross-linguistic perspective. Evidence from some European and Asian languages. In M.M.J. Fernandez-Vest & H. Metslang (eds.), *Perspectives on Questions, answers and information structure; Language Typology and Universals (STUF) – Sprachtypologie und universalienforschung, Special Issue, 70:3*. 391-454.
- FERNANDEZ-VEST, M.M.J. 2018. Ecology of Northern Sami: its typological evolution from oral to written language. In B. Biró, I. Piechnik & U. Tuomarla (eds.), *Finland Suomi 100: language, culture and history*. 27-44.
- HAGÈGE, C. 1986 (/1985). *L'Homme de paroles*. Paris: Fayard (Folio Essais).
- HAGÈGE, C. 1993. *The Language Builder. An Essay on the human signature in linguistic*. Amsterdam / Philadelphia: Benjamins (Amsterdam Studies in the Theory and History of Linguistic Science, ser. IV, Current issues in linguistic theory 94).
- HAKULINEN, A. & KARLSSON, F. 1979. *Nykysuomen lauseoppi*. [Syntax of modern Finnish] Helsinki: Suomalaisen Kirjallisuuden Seura (SKST 350).
- HASPELMATH, M. & KÖNIG, E. (eds.). 1995. *Converbs in cross-linguistic perspectives*. Berlin: Mouton de Gruyter.
- HYÖNÄ, J. & VAINIO S. 2001. Reading morphologically complex clause structures in Finnish. *European Journal of Cognitive Psychology* 13 (34). 451-474.
- IKOLA, O. 1978 (/1971). *Lauseenvastikeoppi*. *Nykysuomen lauseenvastikkeiden ja niihin verrattavien rakenteiden selvittelyä* [Grammar of non finite constructions. The non finite constructions of modern Finnish and other similar constructions]. Helsinki: SKS (Tietolipas 76).
- ISK = *Iso suomen kielioppi* [A Comprehensive Grammar of Finnish]. 2004 HAKULINEN, A., VILKUNA, M., KORHONEN, R., KOIVISTO, V., HEINONEN, T-R. & ALHO, I. Helsinki: Suomalaisen Kirjallisuuden Seura (SKST 950).
- KERBRAT-ORECCHIONI, C. 1984, *L'énonciation. De la subjectivité dans le langage*. Paris: Armand Colin (Collection Linguistique).
- LAMBRECHT, K. 1994. *Information structure and sentence form. Topic, focus and the mental representations of discourse referents*. Cambridge: Cambridge University Press.
- LAMBRECHT, K. 2001. Chap. 80. Dislocation. In M. Haspelmath, E. König, W. Oesterreicher & W. Raible (eds.), *Language typology and language universals*.

- Sprachtypologie und Sprachliche Universalien. Typologie des langues et universaux linguistiques. An International Handbook*, II. Berlin /New York: Walter de Gruyter. vol. 2. 1050-1078.
- MATIĆ, D. & WEDGWOOD, D. 2013. The meaning of focus: The significance of an interpretation-based category in cross-linguistic analysis. *Journal of Linguistics* 49, 127-163.
- MILLER, J. & FERNANDEZ-VEST, M.M.J. 2006. Spoken and written language. In G. Bernini and M.L. Schwartz (eds.), *Pragmatic organization of discourse*. Berlin / New York: Mouton de Gruyter (Empirical Approaches to Language Typology, Eurotype 20-8). 9-64.
- MILLER, J. & WEINERT, R. 2009 (/1998). *Spontaneous spoken language*. Oxford: Oxford University Press (Oxford Linguistics).
- MUKKA, T.K. 1966. *Laulu Sipirjan lapsista*. Jyväskylä: Gummerus.
- MUKKA, T.K. 2008. *Sipirjá*. Sami translation of *Laulu Sipirjan lapsista* (by Jovna-Ánde Vest). Kárášjohka: Davvi Girji.
- NEDJALKOV, V.P. 1995. Some typological parameters of converbs. In M. Haspelmath & E. König (eds.), *Converbs in Cross-Linguistic Perspectives*. Berlin: Mouton de Gruyter. 97-136.
- PEIRCE, C.S. 1978. *Écrits sur le signe, rassemblés, traduits et commentés par Gérard Deledalle*. Paris: Seuil (L'ordre philosophique).
- SETÄLÄ, E.N. & SADENIEMI, M. 1966 (/1919 15th edition). *Suomen kielen lauseoppi*. Helsinki: Otava.
- Skabmatolak. Sabmelaš kirjjalašvuoda antologiija – Tulia kaamoksessa. Saamelaisen kirjallisuuden antologia*. [Fires in the polar night. An anthology of Sami literature]. 1974. Samuli Aikio, E. Itkonen & P. Sammallahti (eds.), (Sulo Aikio, 1974 /1968: Čierastallam – Mäenlaskua [Skiing down]. 295-300). Helsinki: Otava.
- TEREŠČENKO, N.M. 1965. *Nenecko-russkij slovar'*. Moscow: Sovjetskaja Ėnciklopedija.

Issues of comparative Uralic and Altaic Studies (4): On the origin of the Uralic comparative marker

Juha Janhunen

1. Introduction

This paper forms a part in a series dealing with selected issues of comparative Uralic and Altaic Studies. As far as Altaic is concerned, my general premises correspond to the anti-Altaicist line of argumentation, according to which the so-called Altaic languages, recently also renamed “Transeurasian”, are not mutually related, that is, they do not form a divergent language family with a common protolanguage. They do, however, share both material and structural properties, which are best explained as convergent developments due to a complex network of prolonged and recurrent areal contacts between the individual entities, which include not only the families traditionally termed Altaic, but also the Uralic languages.

The Uralic languages, by contrast, do form a coherent language family in the traditional sense, even though the depth of Proto-Uralic is considerable, for which reason the comparative evidence between the main branches of the family is often synchronically elusive and involves even many irregularities and unexplained details. At the same time, the Uralic languages have also interacted both with each other and with the so-called Altaic languages, which is why it is occasionally difficult to distinguish between primary inherited features and secondary traces of areal interaction. It is not rare to observe parallel structural or even material developments that have taken place independently in the different branches of Uralic, as well as in one or several families of the Altaic complex.

2. The Uralic comparative as a denominal form

An example of possible parallel developments is offered by the category of comparison. In general, morphologically expressed comparison of nominals is not a particularly characteristic feature of the Uralic languages, though occasional examples are found in several branches of the family (for a general survey of the situation, cf. Fuchs 1949). Instead, most Uralic languages, like also the languages of the Altaic complex, express comparison by syntactic and lexical means. The comparative is in this context often expressed by the so-called comparative construction involving the base of comparison in an ablative case form ('A is big from B' = 'A is bigger than B'). The superlative, in turn, can be expressed by independent lexicalized superlative markers ('most big' = 'biggest'), or also by constructions using a quantified base of comparison ('big of all' = 'biggest').

Some Uralic languages have, however, a morphological marker expressing the comparative. Interestingly, a formally identical comparative marker is attested in two geographical extremes of the Uralic family: in Finnic and Hungarian (Hungaric). The original shape of the marker may be reconstructed as **-mpA*, e.g. Finnish *suuri* 'big' : COMP *suure-mpi* : OBL *suure-mpa-* : *suure-mma-* 'bigger', Hungarian *nagy* 'big' : COM *nagy.o-bb* : OBL *nagy.o-bba-*. The phonetic differences between the synchronic data may be explained as regular developments: the denasalization and voicing of **mp* to *bb* in Hungarian, and the raising of the final vowel **A* to *i* (: *A*) in Finnic (with additional developments in the individual Finnic languages).

What is surprising is that traces of a similar comparative are very scarce in the Uralic languages located between Finnic and Hungarian. The only other branch that has this marker in the function of a regular comparative is Samic, located further northwest of Finnic, e.g. (Northern Sami) *nuorra* 'young' : COMP *nuora-t* : OBL *nuora-bu-*. It may be noted that the stem vowel *u* (< **o*) in the Sami oblique forms is not exactly regular. The "regular" shape **-mpA* is, however, attested in the pronominal form (Northern Sami) *nubbi* ~ (Southern Sami) *mubpie* 'other, second' < **mubbe* < **muu-mpA*, from the pronoun **muu* 'other' (Korhonen 1981: 246-247).

All of this is common knowledge in Uralic Studies, and there is general agreement that although the comparative marker is absent in the intermediate languages its presence in Finnic and Samic (in the west)

and Hungarian (originally in the east) allows us to reconstruct it as a feature dating back to Proto-Finno-Ugric. In view of the linguistic and geographical distance involved we might, however, also speculate that the Finno-Samic and Hungarian data represent parallel developments from a presupposition that was present in the protolanguage but that had perhaps not yet reached the status of a regular morphologically marked comparative. This idea has received some support in the past (Ravila 1937).

However this may be, there have been attempts at finding a further “origin” for the comparative marker. The received truth in Uralic Studies today is that we are dealing with a denominal form that originally expressed some kind of contrast. This is, in particular, suggested by pronominal examples like Finnish *jo-mpi ku-mpi* = Estonian *e-mb ku-mb* ‘either one’, based on the pronominal roots **yo-* (relative), **e-* (demonstrative), **ku-* (interrogative). In this context, local expressions such as Sami *dá-ppe* ‘here / from here’ vs. *do-ppe* ‘there / from there’: *do-ppil(d)* ‘in / from that direction’, Mordvin *to-mbale* ‘there’, and Mari *tu-mbal-* ‘there’ have also been mentioned (Collinder 1960: 260-261), but their status remains unclear, since they obviously contain also some other element, possibly an independent noun corresponding to Finnish *pieli* or *puoli* ‘side’.

Another line of explanation views the comparative originally as an evaluative (diminutive-augmentative) category (‘somewhat big’ > ‘bigger’). This explanation is supported by Samoyedic data like Tundra Nenets *ngarka* ‘big’: EVAL *ngarka-mpoy*^o ‘somewhat big’, in which nominals with an adjectival meaning take the evaluative marker *-mpoy*^o (Fuchs 1949: 151, also Collinder l.c.). This marker can also be attached to regular nouns, as in Tundra Nenets *yaxa* ‘river’: EVAL *yaxa-mpoy*^o ‘river of medium size’, which resembles the use of the comparative in Finnish examples like *ranta* ‘shore’: COMP-ESS *rann.e-mpa-na* ‘closer to the shore’.

The Samoyedic connection of the Finno-Ugric comparative forms remains, however, uncertain. For one thing, the evaluative marker *-mpoy*^o does not express comparison in the proper sense. Also, this marker seems to be present only in Nenets and is absent even in Enets, the closest relative of Nenets (cf. Siegl 2013: 180). Moreover, the form of the marker *-mpoy*^o is problematic, though it is clearly composed of at least two elements, *-mpo-* and *-y*^o, of which the latter is a common adjectival suffix (< **-yə*) with cognates in all Samoyedic languages. It is

difficult to take a definitive stand on the matter, but it appears more likely that the comparatives in Finnic-Samic and Hungarian are of a different origin.

3. The Uralic comparative as a deverbal form

Although the conception that the Uralic comparative marker is originally a denominal suffix is today prevalent, it is interesting to view the issue in the light of a different framework proposed by Ramstedt already a century ago (1917). In this framework (mentioned by Colinder 1960: 260, 273, but either forgotten or ignored by later scholars), the comparative marker **-mpA* is analysed as being composed of two elements **-m-* and **-pA*, of which the former is a denominal verbalizing suffix forming translative verbs, while the latter is a deverbal nominalizing suffix forming participles or “verbal nouns”. If correct, this analysis means that comparatives are originally nominalized translative verbs (‘growing big’ > ‘bigger’).

The suffixes **-m-* and **-pA* are well attested throughout the Uralic family (Lehtisalo 1936: 110-113, 249-261). Translative verbs in **-m(V)*- are synchronically transparent in, for instance, Samoyedic, as in Tundra Nenets *ngar(-)ka* ‘big’ : TRANSL *ngar.ə-m-* ‘to become big(ger)’. In Finnic, the labial nasal has been dentalized to *-n-*, possibly in analogy to the causative forms, in which the cluster **-m-t-* was regularly homorganized to *-n-t-*, as in Finnish *suuri* ‘big’ : TRANSL *suure-ne-* ‘to become big(ger)’ : TRANSL-CAUS *suure-n-ta-* ‘to make big(ger)’ : TRANSL-CAUS-REFL *suure-n-t-u-* ‘to become big(ger)’. The causative complex **-m-t-* survives also in Hungarian in the shape *-d-*, as in *sötét* ‘dark’ : TRANSL-SG3 *sötét.e-d-ik* ‘to become dark(er)’, in which the transitive stem is detransitivized by the use of the medial conjugation (*ikes igeragozás*).

Similarly, the element **-pA* is widely used in the Uralic languages to form participles, as in Finnish *anta-* = Hungarian *ad-* ‘to give’ : PPLE PRS Finnish *anta-va* = Hungarian *ad-ó* < **amta-* : **amta-pa*. It is, consequently, entirely plausible to view a form like Finnish COMP *suure-m(-)pi* ‘bigger’ as being derived from the corresponding verbal base TRANSL *suure-n(e)-* ‘to become big(ger)’. (It may be noted that the synchronic quality of the nasal in the cluster *mp*, as also in the cluster *nt*, is irrelevant, since all nasal+stop clusters in Finnic are homorganic. The nasal in these clusters could therefore also be analysed as

an archiphoneme, i.e. *Np Nt*. The intervocalic nasal *n* is, however, a distinctive segment and can be identified as original **m* only in the light of comparative data.)

This analysis of the comparative marker **-mpA* as TRANSL-PPLE **-m-pA* takes us further to many related questions. One of them concerns the status of adjectives in Proto-Uralic. There are indications that at least some adjectival words in Proto-Uralic were verbals. This is still synchronically the situation in the Samoyedic languages, where verbals are attested in adjectival functions in items like Tundra Nenets *nyar^oya-* ‘to be red’, *yepə-* ‘to be hot’, *mebyes-* ‘to be strong’. When used adnominally, such items have to be nominalized: PPLE PRS *nyar^oya-na* ‘red’, *yep^o-dya* ‘hot’, *mebye-ta* ‘strong’. Adjectival verbals are the rule in many languages further east in Eurasia, as in Koreanic and Japonic. Assuming that this was the original situation also in (Pre-Proto-)Uralic, the denominal translative verbs in **-m-* would have to be analysed as primary deverbal inchoatives.

A derivational set like Tundra Nenets *paewə-* ‘to be dark’: PPLE *paew^o-dya* ‘dark’: INCH *paewə-m-* ‘to become dark(er)’ suggests that the corresponding Finnish items *pime-ä* (< **pime-tä*) ‘dark’: *pime-ne-* ‘to become dark(er)’ are also ultimately deverbal forms (from a root that may be reconstructed as Proto-Uralic **pilmi-*). The same is true of Tundra Nenets *pyir-cya* (< **pir-kä*) ‘high’: *pyirə-m-* ‘to become high(er)’, corresponding to Finnish *pit(-)kä* ‘long’: *pite-ne-* ‘to become long(er)’ (from **pidi-*). In both cases, the adnominal forms seem to be nominalizations of original verbal roots. There are, however, problems to be solved before the details of the derivational relationships can be explained in full. Comparative information from Samoyedic suggests that there may have been additional elements following the verbal root in some of the relevant forms (Tapani Salminen, personal communication). It remains to be clarified whether these elements were present already at the Proto-Uralic level.

Another issue rising from the analysis of the comparative marker in **-m(-)pA* as a nominalization of translative (or inchoative) verbs concerns the use of the so-called consonant stem. It is well known that the Proto-Uralic high vowel **i* (or **I*), as reconstructed for non-initial syllables, can be absent before certain suffixes in those languages that otherwise preserve the vowels of non-initial syllables (Finnic, Samic). Due to its positional alternation with zero, the high vowel may also have been realized as a quantitatively or qualitatively reduced vowel,

as opposed to the full low vowel **A* (= *a/ä*), of which traces have been preserved more widely in the Uralic languages. However this may be, the translative marker also occurs both with and without a vowel, which means that it has both a vowel and a consonant stem, as in Finnish TRANSL (vowel stem) *pite-ne-* 'become long(er)' : TRANSL-CAUS (consonant stem) *pide-n-tä-*.

Consonant stems are typically attested before the suffix-initial consonants **t *n*, as exemplified by Finnish *vuosi* (< **wooti*) : OBL *vuote-* : PART *vuot-ta* : ESS *vuon-na* (< **woot-na*), in some fully lexicalized cases also before **k *m*, as in *pite-* : *pit-kä* 'long', *kuole-* 'to die' : *kal-ma* 'death'. The comparative marker **-m(-)pA*, when analysed as the nominalization of derived verbal stems in *-m(V)-*, suggests that a consonant stem could also be triggered by a suffix-initial **p*. Although this seems to be the only example of a consonant stem before **p*, the sequence **-m-p-* is in no contradiction with the general framework of Uralic morphophonology.

4. The analogy of the Turkic comparatives

When Ramstedt proposed his deverbal analysis of the Uralic comparative form he had in mind a possible Turkic parallel. Like Uralic, the Turkic languages often express comparison by syntactic or lexical means. They do, however, possess an element, **(I)rAk*, which expresses an increased degree of properties and is generally analysed as a comparative marker, as in *yakshī* 'good' : COMP *yakshī-rak* 'better', *köp* 'much' : COMP *köp-i.räk* 'more', as attested in many Common Turkic languages. This marker has no primary cognate in Bulghar Turkic, suggesting that it may be a Common Turkic innovation, but it has been secondarily borrowed into Chuvash, as well as into neighbouring Finno-Ugric Mari, where it likewise forms comparatives of adjectival nominals (Fuchs 1949: 172).

According to Ramstedt (1952: 199-200) **(I)rAk* is a composite suffix, consisting of the denominal verbalizing suffix **(V)r-*, which typically forms translative/essive verbs, as in Old Turkic *ak* 'white' : TRANSL/ESS *ak-ar-* 'to be/become white', and the deverbal nominalizing suffix **(V)k*, as observed in various types of lexicalized nouns, as in *tanu-* 'to report' : NMLZ *tanu-k* 'witness'. Both suffixes are well attested in Common Turkic since the Old Turkic period (Erdal 1991: 224-261, 499-507), and they also occur in the oldest layer of Bulghar Turkic borrowings in Mongolic, allowing them to be reconstructed as **-rA-* and **-kA*, respectively, as in

Turkic **köök* ‘blue’ : TRANSL **köök-er-* < < **kö(ö)ke-re-* ‘to become blue’ → Mongolic (**köke-re-* id., Turkic (**id-* ‘to send’ : NMLZ (**id-uk* < **idu-ka* ‘(the) sent (one)’ > ‘holy’ → Mongolic **iduka-n* > (**iduga/n* ‘shamaness’.

Ramstedt assumed, consequently, a structural parallel between the Uralic comparative in **-mpA* = TRANSL-NMLZ **-m-pA* and the Turkic comparative in **-(I.)rAk* = TRANSL-NMLZ **-(I.)r-Ak*. His explanation on the Turkic side has, however, been challenged by a hypothesis according to which **-(I.)rAk* is actually a grammaticalized trace of the independent word +*(*y*)*irak* ‘far’ (Räsänen 1957: 74-75, 108-109, following Bang). This hypothesis could possibly be supported by the fact that **-(I.)rAk* normally occurs in word-final position and does not take further suffixes, making it similar to a clitic or “particle” (Erdal 2004: 150-151), rather than a base for a nominal paradigm. By contrast, the Uralic comparative in **-mpA* can take both inflectional and derivational suffixes. This argument loses its power, however, when we realize that *(*y*)*irak* = *(*y*)*ira-k* is itself is a nominalization in **-(V)k* of the verb **yira-* ‘to be distant’.

It has to be concluded that Ramstedt’s explanation may well be correct for both Turkic and Uralic. However, even if the Turkic data were to be explained differently, there is another Turkic parallel which was not yet noticed by Ramstedt. In modern Kazakh comparatives can be formed in two alternative ways: with the Common Turkic suffix **-(I.)rAk* and with the innovative suffix *-LAw*, with regular alternants in *-LAw* ~ *-dAw* ~ *-tAw* depending on the stem-final segment. The two suffixes are basically synonymous, though the forms in *-LAw* have also been identified as “diminutives”, since they can have the connotation of moderation, e.g. *ülken* ‘big’ : COMP *ülken-i.rek* ‘bigger’ : DIM *ülken-dew* ‘(a little) bigger’ (Muhamedowa 2015: 237-238).

Now, the Kazakh “diminutive” suffix *-LAw* is nothing else but a regular Kipchak Turkic nominalization in *-w* (< **-g*) of verbs containing the denominal translative/essive suffix *-LA-*, as in *zhili* ‘warm’ : TRANSL/ESS *zhili-la-* ‘to be/become warm’ : NMLZ *zhili-la-w* ‘(a little) warmer’, *biyik* ‘high’ : TRANSL/ESS *biyik-te-* ‘to be/become higher’ : NMLZ *biyik-te-w* ‘(a little) higher’. There are indications that the “diminutive” suffix is gaining ground in Kazakh and may gradually be replacing the original comparative suffix. It remains to be investigated whether innovative “diminutives” or comparatives of the same type are emerging in other Kipchak languages. In the Turkic context, the suffixes **-LA-* and **-g* are well attested since the Old Turkic period (Erdal 1991: 172-223, 429-455).

5. Conclusion

From the above we may conclude that the Uralic comparative forms, as attested in three branches of the family – Finnic, Samic and Hungarian – are likely to reflect a common source based on the nominalization of Proto-Uralic translative (or inchoative) verbs. Although it is difficult to see any direct areal or chronological connection between the Uralic and Turkic comparatives, the Turkic parallels suggest that this is a strategy that is typologically “natural” for languages of the Ural-Altai type.

Typological parallels are an important dimension that link the different Ural-Altai languages and language families in spite of the fact that these languages are not united by any primary genetic affinity. Most recently, Ylikoski (forthcoming) has noticed that there is an interesting structural parallel between the South Sami relational use of the Uralic comparative in **-mpA* and the Ewenic (Ewenki-Ewen) contrastive/selective use of the Tungusic comparative in **-dlmA-r*. In both cases, the comparative marker can be added to kinship terms. Although this need not mean that the two markers have an analogous origin, a functional parallel like this can serve as a source for conclusions concerning diachronic typology. This an important future field for *real* “Transeurasian” language studies.

Acknowledgements

My warm thanks are due to Professors Angela Marcantonio and Cinzia Franchi, as well as all the other organizers of the Padua workshop on *The State of the Art of Uralic Studies* (November 2016), in which a first version of this paper was presented. I am also grateful to Dr. Tapani Salminen for valuable consultation concerning the Samoyedic data.

Grammatical abbreviations

3 = third person	EVAL = evaluative	PPLE = participle
CAUS = causative	INCH = inchoative	PRS = present
COMP = comparative	NMLZ = nominalization	REFL = reflexive
DIM = diminutive	OBL = oblique	SG = singular
ESS = essive	PART = partitive	TRANSL = translative

References

- COLLINDER, B. (1960). *Comparative Grammar of the Uralic Languages*. A Handbook of Uralic Languages, Part 3. Stockholm: Almqvist & Wiksell.
- ERDAL, M. (1991). *Old Turkic Word Formation: A Functional Approach to the Lexicon*, vols. 1-2. *Turcologica* 7. Wiesbaden: Otto Harrassowitz.
- ERDAL, M. (2004). *A Grammar of Old Turkic*. Handbuch der Orientalistik VIII, 3. Leiden: Brill.
- FUCHS, D.R. (1949). Der Komparativ und Superlativ in den finnisch-ugrischen Sprachen. *Finnisch-Ugrische Forschungen* 30, pp. 147-230. Helsinki.
- KORHONEN, M. (1981). *Johdatus lapin kielen historiaan*. Suomalaisen Kirjallisuuden Seuran Toimituksia 370. Helsinki.
- LEHTISALO, T. (1936). Über die primären ururalischen Ableitungssuffixe. *Mémoires de la Société Finno-Ougrienne* 72. Helsinki.
- MUHAMEDOWA, R. (2015). *Kazakh: A Comprehensive Grammar*. Routledge Comprehensive Grammars. London: Routledge.
- RAMSTEDT, G.J. (1917). Suomalais-ugrilaisen komparatiivin syntyperä. *Virittäjä* 21, pp. 37-39. Helsinki.
- RAMSTEDT, G.J. (1952). *Einführung in die altaische Sprachwissenschaft*. II. Formenlehre. *Mémoires de la Société Finno-Ougrienne* 104: 2. Helsinki.
- RÄSÄNEN, M. (1957). *Materialien zur Morphologie der türkischen Sprachen*. *Studia Orientalia* 21. Helsinki.
- RAVILA, P. (1937). Über das finnisch-ugrische Komparativsuffix. *Finnisch-Ugrische Forschungen* 24, pp. 29-58. Helsinki.
- SIEGL, F. (2013). *Materials on Forest Enets, an Indigenous Language of Northern Siberia*. *Mémoires de la Société Finno-Ougrienne* 267. Helsinki.
- YLIKOSKI, J. (forthcoming). The so-called relation forms of nouns in South Saami. *Finnisch-Ugrische Forschungen* 64. Helsinki.

Previous papers in the series

- JANHUNEN, J. Issues of comparative Uralic and Altaic studies (1): The case of Proto-Mongolic *x. *Hoppou Jinbun Kenkyuu* 10. Tsumagari Toshirou Sensei Taishoku Kinengou, pp. 97-104. Sapporo, 2017.
- JANHUNEN, J. Issues of comparative Uralic and Altaic studies (2): Medial *p in Pre-Proto-Mongolic. In: M. Németh & B. Podolak & M. Urban (eds.), *Essays in the History of Languages and Linguistics*. Dedicated to Marek Stachowski on the occasion of his 60th birthday. Kraków, 2017, pp. 297-305.
- JANHUNEN, J. Issues of comparative Uralic and Altaic studies (3): The Turkic plural in *s. *Altai Hakpo* 27, pp. 155-168. Seoul, 2017.

Revisiting the theory of the Hungarian vs Chuvash lexical parallels

László Marác

1. Introduction

In this article I shall discuss the lexical parallels occurring between Hungarian and Chuvash (a Turkic language) and, occasionally, Mongolian. These parallels have been studied for long time, since the work by József Budenz (1871). The *Hungarian Etymological Dictionary (A magyar nyelv történeti-etimológiai szótára (TESz))*, for which see Benkő 1967, 1970 & 1976) identifies 33 undisputed Hungarian vs Chuvash lexical correspondences¹, that line up together, against ‘Common Turkic’, in displaying specific sound alternations. These alternations are generally referred to as ‘rhotacism’ and ‘lambdacism’, that is: Hungarian vs Chuvash (as well as Mongolian) on the one hand, and Common Turkic on the other hand display /r/ vs /z/ and /l/ vs /š/, respectively. The conventional explanation for these alternations, and related Hungarian vs Chuvash parallels, relies on the traditional classification that Hungarian is a Uralic language, whilst Chuvash is a Turkic language, more precisely, the only representative of the so-called ‘West Old Turkic’ branch (WOT; see below). It is argued that these Hungarian lexical items have been borrowed from either Chuvash, or one of its hypothesized predecessors, such as the Volga Bulgar languages. The possibility of a different direction of borrowing, that is, Chuvash adopting the Hungarian counterparts (or even a totally different explanation alto-

¹ I am indebted to Angela Marcantonio for sharing with me her insights on the issue of the Hungarian vs Turkic lexical correspondences, and for many useful remarks to improve the present text. The research leading to these results has received funding from the European Union’s Seventh Framework Program (FP7/2007-2013) under grant agreement no. 613344.

gether; see footnote (12)), is simply not even taken into consideration within the framework of the conventional model. However, this model makes controversial claims; for example, it distinguishes two branches of Turkic, the so-called 'Lir' versus 'Shaz' branches, also referred to as WOT (as mentioned) and East Old Turkic (EOT) / Common Turkic, respectively. Here I shall put forward an alternative explanation for the rhotacism and lambdacism alternations in Chuvash, arguing that these alternations are the result of word transfer *from* Hungarian *into* Chuvash, and not from Chuvash into Hungarian – as widely claimed. As a consequence, the *lingua genesis* of Chuvash can be considered to be the outcome of intensive and layered processes of contact and substrate interference between different languages / language families co-existing in the linguistic area of the Volga Bends. If the rhotacism and lambdacism (and other) isoglosses in Chuvash are actually originating from Hungarian – rather than the other way round – it follows that WOT is a 'phantom-category'. This conclusion is also supported by the fact that all the reconstructed forms of WOT listed in the two volumes by Róna-Tas & Berta² (2011; hence RT&B) – the dictionary that lists all the Hungarian words classified as loan words (mainly) from Turkic – are very similar or identical to the listed, attested forms of EOT. This suggests that the alternations between Hungarian, Chuvash and Mongolian on the one hand *vs* Common Turkic on the other hand, should be explained in different terms. In the last section of this essay

² In this article I will heavily rely on the work by Ligeti (1986), instead of the more recent study on Turkic *vs* Hungarian lexical parallels put forward in the etymological dictionary by Róna-Tas & Berta (2011). Although Róna-Tas & Berta's (RT&B) corpus is up-to-date, more extended and richer with attested forms from the various Turkic languages with respect to Ligeti's corpus, the authors' analysis, classifications and conclusions do not differ from those put forward by Ligeti. Quite the contrary, RT&B accept Ligeti's basic, theoretical framework and solutions – this being in turn in line with conventional wisdom. Moreover, Ligeti's analysis has the advantage that the lexical items under investigation are less burdened with the assumed sound structure of the proto-Uralic and / or proto-Hungarian forms, as is the case in RT&B corpus – although Ligeti too accepts the Uralic origin of Hungarian and the conventional 'borrowing model'. Ligeti's theoretical analysis too is driven by the reconstruction of clear patterns, but the lexical parallels are not presented in the form of an etymological dictionary. Last, but not least, in this study I am not so much interested in the derivation of the individual lexical items (that is, the derivation of the Hungarian borrowed words from their Turkic parallel), but, rather, in the analysis and interpretation of the (reconstructed) sound patterns, much in the style of Ligeti – although, as it will turn out clear, I hold a fundamentally different view on these issues. Hence, I shall take the patterns discussed in Ligeti (1986) as the basis of my research.

I shall put forward an alternative line of research in order to account for these issues.

The presentation in this article runs as follows. In section 2., the standard classification of the Turkic languages will be discussed. In section 3., the lexical correlations between Hungarian and the Turkic languages will be spelled out. In section 4., the Hungarian vs Chuvash lexical parallels will be systematically illustrated and accounted for. In section 5., a research agenda will be proposed, according to which the phonological alternations / isoglosses under discussion are to be interpreted in terms of ‘Central Asian Sprachbund isoglosses’ – this in turn being in accordance with the conclusions (independently) attained by Marcantonio (2014).

2. The classification of the Turkic languages

2.1. Following Johanson (1998: 81-126), the Turkic languages may be divided into six branches. This classification takes into account geographical and typological properties as well³:

Diagram 1. Six branches of Turkic according to geographical and typological criteria

1. Southwestern (SW) branch: Oghuz Turkic;
2. Northwestern (NW) branch: Kipchak Turkic;
3. Southeastern (SE) branch: Uyghur Turkic;
4. Northeastern (NE) branch: Siberian Turkic;
5. Chuvash: representing Oghur or Bulgar Turkic;
6. Khalaj: representing Arghu Turkic

In this classification, Oghur Turkic (also spelled *Ogur*, *Oguric*; *Bulgar*, *Bolgar*, and its variants) is also referred to as ‘Lir Turkic’, while the other branches are referred to as ‘Common Turkic’ or ‘Shaz Turkic’ – as anticipated above. At the present stage of research, it is unclear when these two major branches of Turkic are supposed to have diverged. Some

³ In Tekin (1990: 5-18) an overview is presented of the different proposals for the classification of the Turkic languages, including proposals by Arat, Benzing, Menges, Poppe, Doerfer, as well as his own classification. Whatever classification is chosen, Chuvash (or its assumed predecessors: Bulgar, Volga Bulgar, or Proto-Bulgar) is always considered as a separate group of the Turkic language family.

scholars hypothesize that the Oghur branch broke off from Common Turkic perhaps as early as 500 BC. Note that all the Turkic languages that are spoken today are of 'Eastern Turkic' type, that is, of Common Turkic origin. There is however one exception: the Chuvash language, spoken on the western periphery of the Turkic language family. These two major groups of Turkic are spoken by completely different numbers of peoples: the eastern branch is spoken by more than hundred million speakers, while the western branch is spoken by hardly one million people. This asymmetry is not accounted for in the literature.

Let us now have a look at the hypothesized genealogical tree of Turkic in accordance with the classification of Johanson (1998: 81-126)

Diagram 2. Hypothetical genealogical tree of Turkic

1. Proto-Turkic: Proto-Bulgar; Common Turkic
2. Proto-Bulgar: *Danube Bulgar; Volga Bulgar
3. Volga Bulgar: Chuvash
4. Common Turkic: Oghuz, Arghu, Kipchak, Karluk, Siberian ...
5. Oghuz: Turkish, Azeri, Turkmen ...
6. Arghu: Khalaj
7. Kipchak: Cuman, Tatar, Kazakh, Kyrgyz.... .
8. Karluk: Uzbek, Uyghur, Salar ...
9. Siberian: Tuvan, Tofa, Khakas, Sarygh Yughur, ...

If we compare the geographic-typological classification (diagram 1) with the hypothesized genealogical classification (diagram 2), we observe that the Oghur or Bulgar Turkic branch matches with the branch starting with Proto-Bulgar in the genealogical tree. This implies that all other branches of the genealogical tree cover the Shaz Turkic or Common Turkic languages. Thus, both approaches assume that the Oghur, or Bulgar languages (the Lir Turkic languages) are a separate branch of the Turkic family. Mainstream Turcology assumes also that Oghur Turkic or Proto-Bulgar was historically spoken in the Hunnic Empire, Old Great Bulgaria (*Magna Bulgaria / Onoguria*), and later on in Volga Bulgaria and the Danube Bulgar Khanate (Danube Bulgaria) – this group would have included the languages of the Huns, the Bulgars, the Khazars and the Eurasian Avars. Note that these languages are only very sparsely documented and, in fact, they are practically unknown. Nowadays its only extant member is the Chuvash language

(as mentioned), that cannot be understood by speakers of the Shaz Turkic languages, mutually intelligible. This being the case, it is uncertain whether Chuvash is directly descended from any of the members of the Oghur group, or a separate branch within this dialect group (see Johanson 1998).

2.2. The Oghur branch is characterized by some (relatively⁴) regular and systematic alternations, as against Common Turkic. Two of the most striking correspondences are the following:

Diagram 3. Oghur ~ Common Turkic sound correspondences

Oghur	/l/ ~ Common Turkic /š/
Oghur	/r/ ~ Common Turkic /z/

As mentioned above, the first alternation in diagram 3 has been referred to as lambdacism, and the second one as rhotacism (the sounds involved in these correspondences being at the origin of the labelling: *Lir* vs *Shaz* Turkic). This formulation implies that the source language is a reconstructed variant of Common Turkic and the recipient language is a reconstructed variant of Oghur. The final sound of the Oghur name is /r/, instead of /z/, which is itself an instance of the ‘*r* ~ *z* alternation’, being cognate with the name of the Oghuz group within Common Turkic (see diagram 2.). Consider the *r* ~ *z* alternation in the following diagram, referring to the equivalents of the word ‘nine’ in different Turkic languages:

Diagram 4. *r* ~ *z* alternation in Turkic languages

languages	‘nine’
Old Turkic	<i>toquz</i>
Turkish	<i>dokuz</i>
Azeri	<i>doqquz</i>
Uzbek	<i>toqqiz</i>

⁴ Marcantonio (2014, footnote 34) points out that the rhotacism vs lambdacism isoglosses do not always materialize within the appropriate context – as is ‘normal’, given the fundamental variable nature of languages. Thus, the correlation is not perfect, although several options come to mind to account for these ‘irregular’ cases. However, for the sake of the argument, I will concentrate on the existing, traditional ‘regular’ pattern (as pursued by Budenz, Gombocz, Ligeti and others), setting the ‘irregular’ cases aside for the moment. See the last section for discussion of these cases.

Uyghur	<i>toqquz</i>
Kazakh	<i>toǵız</i>
Kyrgyz	<i>toǵuz</i>
Altay	<i>toǵus</i>
Khakas	<i>toǵıs</i>
Tuvan	<i>tos</i>
Sakha/Yakut	<i>toǵus</i>
Khalaj	<i>toqquz</i>
Chuvash	<i>tăχăr</i>

This diagram shows that only in the Chuvash cognates we find rhotacism for the final consonant of the word for ‘nine’, *tăχăr*, whereas in all the other instances of the Common Turkic branch we find zetacism, a final /z/, which has become voiceless in some younger branches of Turkic, including Altay, Khakas, and Sakha/Yakut. Hence, rhotacism *vs* zetacism is a (relatively) clear dividing criterium between Oghur and Common Turkic. The same is true for the ‘*l* ~ *š* alternation’, as in Chuvash *xěl* and Common Turkic *qış* ‘winter’ (see Tekin 1990: 11). In the next section I shall illustrate that Hungarian lines up with the Oghur branch, when there are lexical correspondences among Common Turkic, Chuvash and Hungarian.

3. The lexical correlations between Hungarian and the Turkic languages

3.1. Hungarian vs Chuvash parallels

Ligeti (1986: 9) discusses the history of the ‘Old Turkic loan words of Chuvash type in Hungarian’ in depth⁵. According to the author, this concept was an idea of the German scholar József Budenz, and was first published in the academic journal *Nyelvtudományi Közlemények* (1871: 67-135), in the heat of the ‘Ugric-Turkish Battle’⁶. Budenz argued that

⁵ This article will closely follow the classification and the illustration of Chuvash and Turkic material as presented and discussed in Ligeti (1986), unless indicated otherwise.

⁶ The ‘Ugric-Turkish Battle’ was an academic debate revolving around the issue of the primacy of the ‘Ugric’ (read ‘Finno-Ugric’) as against the Turkic genetic relation of Hungarian. The debate between Budenz, who argued for a Hungarian-Ugric genetic relation, and Ármin Vámbéry, who considered Hungarian a ‘mixed’ language of Ugric and Turkic elements, started around 1870 (see Marcantonio *et al.* 2001; Marác

the Hungarian loan words of Chuvash type (from Old Turkic) could be accounted for by postulating a ‘dummy’, that is, a predecessor language of Chuvash, like Proto-Chuvash or Old Chuvash, that would be the source of the sound matches in Hungarian – the common features being, mainly, rhotacism and lambdacism, as discussed,

Note that the starting point of rhotacism in Proto-Turkic could be, in principle, either /r/ or /z/. If we assume Proto-Turkic /r/ as the original sound, this sound would have then developed into Common Turkic /z/ (zetacism), but would have been preserved in Oghur /Chuvash as /r/. The second possibility is to assume an original Proto-Turkic /z/, that developed into Oghur /Chuvash /r/ (rhotacism), and was preserved in Common Turkic as /z/. The latter hypothesis is widely accepted in Turcology, whilst the inclusion of Mongolian lexical affinities into this picture has raised support for the former hypothesis, especially among the proponents of the Altaic theory, that is: Mongolian would be an Altaic language, related with Turkic, that patterns with Oghur /Chuvash within the domain of the isoglosses under discussion. If this were the case, one may well conclude that Chuvash must have been the oldest Turkic variant.

Consider the following examples where Hungarian, Chuvash and Mongolian phonological patterns line up together, as opposed to Common Turkic:

Diagram 5a. Hungarian vs Chuvash vs Mongolian as against Common Turkic (Ligeti 1986: 14)

meaning	Hungarian	Chuvash	Common Turkic	Mongolian
‘write’	<i>ír</i>	<i>šir</i>	<i>yaz-</i>	<i>ĵiru-</i>

2012 for further details). Supporters of the Uralic theory consider the Finno-Ugric /Uralic origin of Hungarian as settled since this Ugric-Turkish Battle, although, in reality, the data reported in this regard from both camps are not decisive, as admitted by Budenz himself (see again Marcantonio *et al.* 2001 and Marcantonio 2002: 42). Although Hungarian is classified nowadays as an Uralic language, it contains a statistically significant number of lexical as well as morphological correspondences with Turkic, and, to a lesser extent, Mongolian and Tungusic. There is an impressive academic literature documenting this ‘state of the art’, including the quoted RT&B dictionary; see also Marcantonio (2016, 2017a & 2017c). The position generally defended in textbooks is that these correspondences between Hungarian and Turkic (/Altaic) cannot be taken as evidence in favour of a Hungarian *vs* Turkic (/Altaic), or ‘Ural-Altaic’ genetic relation, although no convincing explanation is provided for this – but see again Marcantonio (2014, 2016 & 2017a) for further discussion and a different point of view.

Diagram 5b. Hungarian vs Chuvash vs Mongolian as against Common Turkic (Ligeti 1986: 14)

meaning	Hungarian	Chuvash	Common Turkic	Mongolian
'calf'	<i>borjú</i>	<i>pāru</i>	<i>buzāyu</i>	<i>birayu</i>

Ligeti (1986: 48-52) argues that if the Altaic language family is rejected altogether, then the correspondences between Turkic and Mongolian must be explained otherwise, that is: the Mongolian borrowing should be assumed to derive from Common Turkic, having nothing to do with Chuvash. In any case, the Mongolian counterparts in the $r \sim z$ correspondences are comparable to the Hungarian ones in the relevant correspondences. Although the issue of the Turkic borrowing into Mongolian is still unsettled, what seems to be important for Hungarian scholars, like Ligeti, is that the phenomenon of the Mongolian borrowing had nothing to do with that of the Hungarian borrowing: "The period of the Turkish-Mongolian borrowings are not even settled approximately, but they are so old that from the perspective of the Hungarian borrowings it does not matter" (Ligeti 1986: 52; my translation).

Budenz' ideas were taken up as the starting point for further research. Gombocz refined the list of Hungarian *vs* Chuvash phonological and lexical correspondences in his monograph: *Die bulgarisch-türkischen Lehnwörter in der ungarische Sprache* (published in 1912), where he first reported also instances of lambdacism. Once again, much of the research concentrated on the question of which Turkic people the Hungarians had borrowed their loan words from, particularly the 'cultural and environmental' loans words. Gombocz (1912) pointed at the Volga Bulgars, a position also defended by the Hungarian Turcologist Gyula Németh (1921: 205-207), who linked the $r \sim z$ correspondence to the morphological structure of the names of the peoples themselves (*Oghur*, *Bulgar*, and so on). The time of borrowing would have been the period stretching from the 5th to the 9th century AD, according to both authors. These claims have since been adopted by mainstream linguists working in this field. Thus, basically, the Hungarian loan words from Turkic are assumed to be relatively 'young' within this paradigm. However, due to the existence of complicated phonological patterns and deviations, several older unknown variants of Chuvash must be postulated in order to account for all the relevant correlations.

In the following paragraph, I shall illustrate the phonological and lexical correspondences among Chuvash, Hungarian and Mongolian on the one hand, and Common Turkic on the other hand, in more detail.

3.2. Grouping together the Hungarian vs Chuvash parallels

The Hungarian etymological dictionary TESz identifies 33 safe Old Turkic loan words of the Chuvash type in Hungarian, together with other 9 borrowed elements considered to have, however, an uncertain status. Another 43 cases might also belong to this group, according to Ligeti (1986: 11). Ligeti distinguishes two groups with respect to the Chuvash type of borrowing in Hungarian: in the first group, Chuvash (and its assumed earlier stages) would have preserved older sounds, whereas the second group would reflect innovations on the Chuvash side. The first group of phenomena is presented under (A); and the second group of phenomena under (B), following Ligeti (1986: 14-52).

A. The phenomena illustrated in this group are rhotacism (A1), lambdacism (A2) and the isogloss: Hungarian *vs* Chuvash *vs* Mongolian word initial /š-/ *vs* Common Turkic /s-/, that is, ‘sigmacism’ (A3). In all these instances, the Chuvash, Hungarian and Mongolian correspondences pattern together, against Common Turkic:

A1. Rhotacism. Ligeti (1986: 14) lists 16 cases of rhotacism; five of these appear both in Chuvash and Mongolian: *borjú* ‘calf’, *iker* ‘twin’, *ír* ‘write’, *ökör* ‘ox’, and *sár* ‘mud’⁷:

Diagram 6a. Hungarian /r/ ~ Chuvash /r/ ~ Common Turkic /z/ ~ Mongolian /r/

meaning	Hungarian	Chuvash	Common Turkic	Mongolian
‘calf’	<i>borjú</i>	<i>pāru</i>	<i>buzāyu</i>	<i>birayu</i>
‘twin’	<i>iker</i>	<i>yēker</i>	<i>ekiz</i>	<i>ikire</i>
‘ox’	<i>ökör</i>	<i>vākār</i>	<i>öküz</i>	<i>üker</i>
‘mud’	<i>sár</i>	<i>šur</i>	<i>sāz</i>	<i>siroi</i>
‘filter’	<i>szűr</i>	<i>sěr-, sör-</i>	<i>süz</i> ‘clean’	X

⁷ In all the diagrams ‘X’ indicates lexical gaps.

In the following diagram the Chuvash counterparts are missing. According to Ligeti (1986:15) Chuvash *tüs* is a loan word from Common Turkic:

Diagram 6b. Hungarian /r/ ~ Common Turkic /z/ ~ Mongolian /r/

meaning	Hungarian	Chuvash	Common Turkic	Mongolian
'polecat'	<i>görény</i>	X	<i>küzän</i>	<i>kürene</i>
'bald (headed), bare'	<i>tar</i>	X	<i>taz</i>	<i>taraqai</i>
'have patience, endure'	<i>túr</i>	X (<i>tüs</i>)	<i>töz</i>	<i>türe-</i>

In the following cases, the Mongolian counterparts are missing.

Diagram 6c. Hungarian /r/ ~ Chuvash /r/ ~ Common Turkic /z/

meaning	Hungarian	Chuvash	Common Turkic	Mongolian
'ring'	<i>gyűrű</i>	<i>šěrě</i>	<i>yüzük</i>	X
'knee'	<i>térd</i>	<i>čěr</i>	<i>tiz /tizik / tiz</i>	X

The Hungarian body part term *térd* 'knee' has no Uralic counterpart, but has equivalents in Turkic, including Old Turkic *tiz*, Kyrgyz *tizä*, Osman *diz* and Chuvash *čěr* 'knee'. Despite this being a basic item of the Hungarian vocabulary, TESz (III, 1976: 895) assumes that Hungarian *térd* is a loan word from Old Turkic, more precisely from the Chuvash type of Old Turkic. TESz further assumes that the root *tér-* of *térd* originates from the Chuvash form **tír* or **tēr*, root to which the Hungarian diminutive suffix *-d* or *-gy /d'* would have been attached, once borrowed, yielding *térd*. However, the problem with this analysis is that the assumption of the attachment of a diminutive suffix has no independent evidence, neither are the forms in question documented in older stages of Chuvash; the explanation is therefore *ad-hoc*.

The following diagram displays cases in which both the Chuvash and the Mongolian counterparts are missing. Ligeti (1986:15) claims that Chuvash *tiněs* and Mongolian *tenggis* are borrowings from Common Turkic:

Diagram 6d. Hungarian /r/ ~ Common Turkic /z/

meaning	Hungarian	Chuvash	Common Turkic	Mongolian
'pole'	<i>karó</i>	X	<i>qazuk</i>	X
'dragon'	<i>sárkány</i>	X	<i>sazayan</i>	X
'sea'	<i>tenger</i>	X	<i>tāñiz</i>	X
'buttermilk'	<i>iró</i>	X	<i>yáz</i> (Tatar)	X

A2. Lambdacism. Lambdacism appears in four cases only, two of which have an uncertain status. There are no Chuvash counterparts for the two safe etymologies, that are based exclusively on Common Turkic; nevertheless, they are referred to as 'Old Turkic loan words of the Chuvash type' (Ligeti 1986: 17). Notice also that only one Mongolian counterpart is attested:

Diagram 7. Hungarian /l/ ~ Common Turkic /š/ ~ Mongolian /l/

meaning	Hungarian	Chuvash	Common Turkic	Mongolian
'noon, south'	<i>dél</i>	X	<i>tüš</i> (Kyrgyz)	<i>düli</i> 'middle', half'
'lay down'	<i>dől, düll</i>	X	<i>tüš-</i> (Uyghur) 'fall'	X

A3. Sigmaticism. Common Turkic initial /s-/ corresponds to Hungarian, Chuvash and Mongolian /š-/ (written *s* in Hungarian), as in *sár*, *sárga* 'white, pale, yellow', *seper*, *söpör* 'sweep, broom', and *sarló* 'sickle'. These words are again referred to as 'Old Turkic loan word of the Chuvash type' (Ligeti 1986: 18). Compare:

Diagram 8. Hungarian /š/ ~ Chuvash /š ~ ś/ ~ Common Turkic /s/ ~ Mongolian /š/

meaning	Hungarian	Chuvash	Common Turkic	Mongolian
'yellow, light'	<i>sárga, sár</i>	<i>šurě</i> 'white'	<i>sariγ</i> 'yellow, pale'	<i>šira</i> 'yellow'
'brush, sweep'	<i>seper, söpör</i>	<i>šăpăr</i> 'broom'	<i>sipir-</i>	<i>sigür-, ši'ür-</i>
'sickle'	<i>sarló</i>	<i>šurla</i>	X	X

Let us now present and discuss the phenomena pertaining to the B group, as anticipated above:

B. The Chuvash variants in the following correspondences are considered as innovation by Ligeti (1986: 13):

1. Proto-Turkic initial /č/ ~ Chuvash /š/ ~ Hungarian /š/ (written *s*), as in *saru* 'shoe /sandal';
2. Proto-Turkic final /k/: in multi-syllable Chuvash words final /-k/ is deleted, whilst in Hungarian it *may* spirantize into /-γ/, as attested in Old Hungarian documents, such as *A tihanyi apátság alapítólevele* 'the foundation charter of the nunnery of Tihany' (TA; issued in 1055). In this document we find the sound development indicated above, as seen, for example, in the velar fricative of *mene-h* [mëne-γ] 'go-ing', from *men-ni* 'to go'. This velar fricative has subsequently disappeared, causing the lengthening of the preceding vowel, as we can see in the development from Old Hungarian *mene-h* > Modern Hungarian *men-ő* 'going' (ő = /ö:/), the long vowel being indeed the Modern Hungarian form of the present participle suffix (for more detail on these data see Marcantonio⁸ (2017c)). The same phenomenon is to be found in many other, similar instances, such as *apró* 'small', *seprő* 'yeast of wine'.
3. Proto-Turkic initial /y-/ ~ Chuvash /ś/ ~ Hungarian three variants: /d'/ (written *gy*), /s/ (written *sz*), and /Ø/, as seen in: *gyertya* 'candle', *szél* 'wind', *ír* 'write', respectively;
4. Proto-Turkic initial /ya-/ ~ Chuvash /i/ ~ Hungarian /i:/ (written *í*), as in *ír* 'write';
5. Proto-Turkic initial /y-/ ~ Chuvash /ś-/ ~ Hungarian /ń-/ (written as *ny*), as in *nyár* 'summer' and *nyak* 'neck';
6. Proto-Turkic medial and final /d/ ~ Chuvash /r/ ~ Hungarian /d/ and /z/, as in *búza* 'wheat';
7. Proto-Turkic final /-ŋ/ ~ Chuvash /-m/ ~ Hungarian /-m/, as in *gyom* 'weed'.
8. Proto-Turkic first syllable /a/ ~ Chuvash /i/ ~ Hungarian /i/, as in *tinó* 'young bullock, ox'.

These eight types are represented in the following diagrams, respectively:

⁸ I am indebted to Angela Marcantonio for providing me with these data and related comments from her (2017c) article, before its publication.

B1.**Diagram 9.** Proto-Turkic initial /č/ ~ Chuvash /š/ ~ Hungarian /š/ (written s)

meaning	Hungarian	Chuvash	Common Turkic	Mongolian
'shoe'	<i>saru</i>	X	<i>čaruq</i>	<i>čaruχ</i>

According to TESz (1976, III: 497), Hungarian *saru* is an Old Turkic loan word. The reconstructed Turkic form could be *čaruγ, *čariγ, or could be of the Chuvash-type *šaruγ, *šariγ. The borrowing could have taken place also when there was /-k/ in final position, instead of final /-γ/.

B2.**Diagram 10.** Hungarian final /o:, ö:/ (written ó, ő respectively) ~ Common Turkic /k/_

meaning	Hungarian	Chuvash	Common Turkic	Mongolian
'small'	<i>apró</i>	X	<i>oprak</i>	X
'yeast of wine'	<i>seprő</i>	<i>šapre</i>	<i>šeprek</i>	X

TESz (1967, I: 167) considers *apró* as deriving from Old Turkic, whereby the source could have been *opray. The final spirant has been lost mostly in Hungarian, as mentioned above, yielding a diphthong that then developed into a long vowel. According to TESz (1976, III: 519), *seprő* originates from Old Turkic, for which consider the Kyrgyz cognate *söbrö*. Vámbéry, quoted in Gombocz (1912: 116), refers to the Common Turkic variant with final /-k/, variant that is not given in TESz.

B3.**Diagram 11a.** Proto-Turkic initial /y-/ ~ Chuvash /š-/ ~ Hungarian /d'-/ (written gy)

Meaning	Hungarian	Chuvash	Common Turkic	Mongolian
'shredder, slicer; slice'	<i>gyalu</i> <i>gyalul</i>	X	<i>yiš-, yiši-</i> 'slice'	X

⁹ Here the capital D refers to phenomenon of the voiced and voiceless alternation (according to context) of the dental plosive in Chuvash (and in Turkic in general).

'ring'	<i>gyúrú</i>	<i>šěřě</i>	<i>yüzük</i> 'ring' (Turkish)	X
'knead'	<i>gyúr</i>	<i>šar</i>	<i>yuyur-</i>	<i>ǰiyura</i>
'candle'	<i>gyertya</i>	<i>šurDa</i> ⁹	<i>yarta</i> 'candle'	X
'bulrush'	<i>gyékény</i>	X	<i>yikän, yekän</i>	<i>ǰigesün, ǰegesün</i>
'belly, stomach'	<i>gyomor</i>	X	<i>yumur</i> (Osman)	X

Diagram 11b. Proto-Turkic /y-/ ~ Chuvash /š-/ ~ Hungarian /s-/ (written as *sz*), or Hungarian /š/ (written as *s*)

meaning	Hungarian	Chuvash	Common Turkic	Mongolian
'wind'	<i>szél</i>	<i>šil</i>	<i>yēl</i>	X
'sickle'	<i>sarló</i>	<i>šurla</i>	X	X

Notice that *sarló* 'sickle' has no counterpart in Common Turkic. Ligeti (1986: 24) has to assume that a Proto-Turkic /y-/ had been changed into Ancient or Old Chuvash /š-/.

Diagram 11c. Proto-Turkic /y-/ ~ Chuvash /š-/ ~ Hungarian /Ø-/

meaning	Hungarian	Chuvash	Common Turkic	Mongolian
'write'	<i>ír</i>	<i>šir</i>	<i>yaz-</i>	<i>ǰiru-</i>

This diagram equals the one in (5a) above; however, here the focus is on the initial sound correspondence, whilst (5a) demonstrates the Hungarian, Chuvash and Mongolian correspondence (rhotacism), as against Common Turkic.

B4.

Diagram 12. Proto-Turkic /ya/ ~ Chuvash /i-/ ~ initial Hungarian /i:/ (written as *í*)

Semantic Meaning	Hungarian	Chuvash	Common Turkic	Mongolian
'write'	<i>ír</i>	<i>šir</i>	<i>yaz-</i>	<i>ǰiru-</i>

B5.

Diagram 13. Proto-Turkic initial /y-/ ~ Chuvash /ś-/ ~ Hungarian /ń-/ (written as *ny*)

Meaning	Hungarian	Chuvash	Common Turkic	Mongolian
'neck'	<i>nyak</i>	X	<i>yaqa</i> 'collar' (Old Turkic)	X
'summer'	<i>nyár</i>	<i>śur</i> 'spring'	<i>jaka</i> 'collar' (Osman) <i>jaz</i> 'spring' (Old Turkic) <i>jaz</i> 'summer' (Turkmen)	X

According to TESz (1970, II: 1031), *nyak* (that has no Chuvash counterpart) is of unknown origin. According to the etymological dictionary by Lakó & Rédei (1978, III: 476-477), the origin of Hungarian *nyár* is unclear, whilst, according to TESz (1970, II: 1036), *nyár* is probably of Old Turkic origin, having being borrowed into Hungarian in Uralic times.

B6.

Proto-Turkic medial and word final /-d/ often corresponds in Hungarian and Chuvash to /-d/ or /-t/, and sometimes in Hungarian to /-z/ – note that the stem of all the nouns in diagram (14a) co-occur with suffixes. In diagram (14b), Ligeti argues (1986: 30) that Proto-Turkic /d/ corresponds to Hungarian /z/ in the case of *búza*.

Diagram 14a. Proto-Turkic medial and final /-d/ ~ Chuvash /-t/ ~ Hungarian /-d/

Meaning	Hungarian	Chuvash	Common Turkic	Mongolian
'navel'	<i>köld-ök</i>	<i>kănt-ăr</i>	<i>kind-ik</i>	X
'wise'	<i>ild-om</i>	<i>yăłtt-am</i>	<i>ıld-am</i>	<i>ıld-am</i>
Turkic 'smart'			<i>yıl-dam</i>	<i>ĵıld-am</i>

Diagram 14b. Proto-Turkic /-d/ ~ Chuvash /-t/ ~ Hungarian /-z/

meaning	Hungarian	Chuvash	Common Turkic	Mongolian
'wheat'	<i>búza</i>	<i>pări</i> 'wheat' ¹⁰	<i>buydai</i>	X

B7.**Diagram 15.** Proto-Turkic final /-ŋ/ ~ Chuvash /-m/ ~ Hungarian /-m/

meaning	Hungarian	Chuvash	Common Turkic	Mongolian
'weed'	<i>gyom</i>	<i>śum, śăm</i>	<i>yoŋ</i>	X

Proto-Turkic final /-ŋ/ and /-n/ have only a few correspondences in Hungarian and Chuvash; moreover, they are not clear, since Hungarian *gyom* and Chuvash *śum, śăm* have no convincing equivalent in Common Turkic. Nevertheless Ligeti (1986: 35) maintains that the Hungarian root word is a borrowing from Old Turkic of the Chuvash type.

B8.**Diagram 16.** Proto-Turkic first syllable /a/ ~ Chuvash /i/ ~ Hungarian /i/

meaning	Hungarian	Chuvash	Common Turkic	Mongolian
'young bullock, ox'	<i>tinó</i>	<i>tīna</i>	<i>tāna</i>	X

According to Ligeti (1986: 36-39), Chuvash has no long vowels, in general. Hence, if there is a correspondence with a long vowel in Hungarian, like in *sár* 'mud', the Chuvash counterpart typically displays a short vowel, as in *śur*. This being the case, the question arises: where does the long vowel of the Hungarian counterparts originate from? In order to answer to this question, Ligeti proposes a highly speculative derivation: in earlier stages of Chuvash (including Proto-Turkic) there must have been a long vowel that has been diphthongized in Old Chuvash; these diphthongs would have then been borrowed into Hungarian either as a diphthong or as a long vowel (Ligeti 1986: 37-38). This path of development would be illustrated by the Chuvash and

¹⁰ According to Ligeti (1986: 30) Proto-Turkic /-d/ only remains as /-d/ after /l/, /n/, /r/; otherwise neologisms pop up in Chuvash, such as /r/.

Common Turkic equivalents of Hungarian *kék*: medial /v/ in Chuvash *kǎvak* must have been changed first into a glider /y/, and then into a diphthong (Ligeti 1986: 38), before appearing as a long vowel in Hungarian, as shown below:

Diagram 17. ‘blue’ in Hungarian vs Chuvash vs Common Turkic vs Mongolian

meaning	Hungarian	Chuvash	Common Turkic	Mongolian
‘blue’	<i>kék</i>	<i>kǎvak</i>	<i>kök</i>	<i>küke</i> ‘blue, green’

4. Accounting for the Hungarian vs Chuvash parallels

4.1. Analysing the data

Let us now analyse the set of data presented in the previous section. The first relevant observation to be made is that in the above reported diagrams the Hungarian counterparts are always present, whereas the Turkic (/Altaic) parallels can be missing. The Common Turkic counterparts are only missing in two cases (in A3/8 and in B3/11b), whilst the Chuvash and Mongolian parallels are missing more frequently. In Chuvash the parallels are absent in ten of the cases presented; compare below the number (included within parenthesis) of the missing counterparts, including: in A1/6b (3); in A2 (2); in B1 (1); in B3/11a (3); and in B5 (1). The Mongolian parallel is even more frequently missing, that is, in 23 of the cases presented; compare: A1/6a (1); A1/6c (2); A1/6d (4); A2 (1); A3 (1); B2 (2); B3/11a (4); B3/11b (2); B5 (2); B6 (2); B7 (1); and B8 (1).

Thus, a substantial percentage of the assumed Chuvash parallels are missing. This leads to the anomaly that the forms of Chuvash (or its earlier variants), from which the Hungarian items would have been borrowed, are, in fact, unknown – or, perhaps, they never existed. Hence, an earlier, hypothetical starred form of Chuvash has to be postulated, a form that, in turn, is reconstructed on the basis of the Common Turkic equivalent as well as the Hungarian counterpart – in a clear circular way. This reconstruction is even more problematic when both the Chuvash and Mongolian parallels are missing, like in A1/6d, A2, B2, B3/11a, and B5. Thus, the reconstructed Chuvash forms from which the Hungarian words are supposed to have been borrowed, are highly speculative, as in the following examples: *karó* and *író* (compare A1/6d), *dól*,

dűl (compare A2); *apró* (compare B2); *gyalu*, *gyékény*, *gyomor* (compare B3/11a) and *nyak* and *nyár* (compare B5). The same applies to instances of lambdacism, as presented in A2. When the Chuvash and Mongolian counterparts are missing, obviously, we are left with the Hungarian and Common Turkic correspondences only, so that it cannot be excluded that lambdacism could be accounted for as originating from an old borrowing from Proto-Turkic into Hungarian, or vice versa (still assuming, for the sake of the argument, that 'borrowing' is the correct process here).

Another instance where the (assumed) direction of borrowing from Chuvash runs into problems is B8, that is, Proto-Turkic first syllable /a/ vs Chuvash /i/ vs Hungarian /i/. In this case it is assumed that the (Old) Chuvash (velar) vowel /i/ originates from Proto-Turkic /*ā ~ a/, another velar vowel, whilst the Hungarian front counterpart /i/ would be derived from (Old) Chuvash velar /i/. However, the Hungarian front vowel /e/ or /i/ also appear in those cases when there is no velar /i/ in (Old) Chuvash, as shown in Hungarian *gyertya* ~ *gyirtya* ~ *gyortya* vs Common Turkic *yarta*, but Chuvash *śurDa* ('candle'). Thus, the Hungarian front vowel /e/ or /i/ might derive from more than one source – as is often the case in the process of borrowing.

The hypothesized loan words *karó* and *író* (compare A1/6d), that are missing from Chuvash and Mongolian, support the thesis that rhotacism developed independently in these languages.

It is also interesting to observe that the Hungarian items sometimes have a counterpart only in Chuvash, without any equivalent in Common Turkic, such as Hungarian *sarló* (B3/11b). Variants of this word also appear in some Uralic languages: Votyak (/Udmurt) *śurlo*, Komi-Zyrian *t'śarla* and Cheremis (/Mari) *sarla*, but according to TESz (III: 495), these are loan words originating from Chuvash *śurla*.

Another set of parallels where there is only a Hungarian vs Chuvash lexical correspondence is Hungarian *disznó* 'pig, swine', and its Chuvash equivalent *sisna*. In all these cases it is completely unclear what the 'Chuvash type of Old Turkic loan word' in Hungarian might mean, since there is no 'type' available other than the Chuvash one. The Chuvash counterparts of *sarló* and *disznó* could be very well loan words from (Old) Hungarian into (Old) Chuvash.

In addition to the anomalies and asymmetries that show up if we assume a borrowing model in which Hungarian is always the recipient and Chuvash (or one of its earlier variants) always the donor language, there is another pressing reason why this conventional model is highly

questionable. There is no solid explanation able to account for the quick transformations of Chuvash in the light of the slow pace of the dissolution and development of the other Turkic languages. It has to be assumed that, for some reason, the earlier stages of Chuvash (Ancient, Old, Volga Bulgar and Proto-Bulgar) have been changing very quickly between the 5th and the 9th century AD. The dissolution of the Turkic languages is however a relatively recent phenomenon, because these languages, apart from Chuvash, are mutually intelligible. This clearly shows that the whole ‘Chuvash-complex’ model embraced in mainstream linguistics is conceptually misguided.

4.2. Mainstream vs alternative interpretations

4.2.1. According to traditional, mainstream interpretation, the Hungarian *vs* Turkic (/Altaic) phonological / lexical correspondences, and related isoglosses, as discussed above (rhotacism in Chuvash, Hungarian and Mongolian, *contra* zeticism in Common Turkic, etc.), are *always* the result of borrowing, whereby Hungarian is *always* the recipient language and Old Turkic the donor language, rather than vice versa (this interpretation being re-proposed recently, for example, in Bence (2014)). However, on the basis of the data and arguments presented above, this explanation does not appear to stand up to scrutiny. As a matter of fact, the most pressing argument in support of the conventional interpretation is that Hungarian has been classified as a Uralic language, whose core, ancient vocabulary ‘must’ be, therefore, of Uralic origin. Consequently, any other non-Uralic correspondence within the basic lexicon ‘must’ necessarily be an instance of borrowing into Hungarian. The other option, namely that it is Hungarian – or its earlier variants – that is functioning as the donor language in these cases, is not even taken into consideration. However, there is strong, empirical evidence that rhotacism in Hungarian cannot originate from Chuvash (or its older variants), because rhotacism in Chuvash is a late development, going back only to the 11th century AD. This has been convincingly argued for by the German Turcologist Johannes Benzing in his article: *Die angeblichen bolgartürkischen Lehnwörter im Ungarischen* (1942), as discussed below.

4.2.2. The first established records of the Turkic languages are the Orkhon inscriptions, by the Kök Türks, dating from the 8th century. These runic inscriptions were discovered in 1889 in the Orkhon Valley,

in Mongolia. Another important source of the Turkic languages is *The Compendium of the Turkic Dialects*, originally written in Arabic (*Divânü Lügati't-Türk*), in the second half of the 11th century (c.1072-74 AD), by the Uyghur scholar Mahmud al-Kashgari. *The Compendium* is the first, early account of the Turkic language family in the form of a comprehensive dictionary, pertaining mainly to the south western branches of Turkic. *The Compendium* was indeed meant for use by the Caliphs of Bagdad, the new Arabic allies of the Turks. The Al-Kashgari's dictionary also includes the first known map of the areas inhabited by Turkic speaking peoples. Benzing (1942: 25) considers Mahmud al-Kashgari's dictionary as the 'witness' of the fact that rhotacism in Chuvash cannot have occurred before the 11th century. His argument goes as follows. First, Chuvash /z/ and its Turkic correspondent /z/ have both been derived simultaneously from Proto-Turkic /ð/. Mahmud al-Kashgari demonstrates that Bulgar, the hypothesized ancient form of Chuvash, still possesses the /z/ sound. There is no record or report about any sound change from /z/ to /r/, this being documented only since 1230, in the Russian chronicle *Trunowe*. Next to the Hungarian loan words supposedly demonstrating the sound change /z/ > /r/, like in Hungarian *ökör* 'ox', we also find the hypothesized sound change Proto-Turkic */ð/ > /z/, like in Hungarian *búza* 'wheat'. Benzing (1942: 26) proposes the following derivation: Proto-Turkic **būđaj* or **bud[q]aj* > Early Bulgar **buzaj* > Late Bulgar **buraj* > Chuvash *pāri* 'dinkel wheat'. If this is the derivation, then Hungarian /z/ could only originate from Early Bulgar, whilst a borrowing from Chuvash itself must be excluded, because this would yield the wrong result. Second, another important item of evidence for the dating of the assumed borrowing is the following sound change: Common Turkic /(y)a/ > Chuvash /i ~ i/, as in Common Turkic *yaz* 'write', corresponding to Chuvash *šir*, in turn derived from reconstructed proto-Bulgar **jir*. This lexical item could have been the form borrowed by Hungarian, resulting in *ír* 'write'. Certain Arabic loan words in Chuvash also participate in the same sound change /(y)a/ > Chuvash /i ~ i/. On the basis of this, Benzing claims that the sound change in question cannot have taken place before the 10th century, the period of Islamization of the Volga Bulgar area. Moreover Mahmud al-Kashgari's examples relating to this change still contain the /a/. This is the case in *Suvar*, one of the Oghur languages closely related to Chuvash, where we find *bal* 'honey', and not its nowadays Chuvash variant *pyl* 'honey'. Benzing (1942: 27) concludes his argumentation as follows:

Da wir nach den obigen Darlegungen allen Grund zur Annahme haben, daß der Wandel $z > r$ und $a > y(i)$ im Bolgarischen nicht vor dem XI. Jahrhundert erfolgt ist, können die Ungarn die bisher für bolgarischen gehaltenen Lehnwörter keinesfalls von den Bulgaren übernommen haben, denn um diese Zeit hatten sie keine Berührung mehr mit den Bulgaren. Von *búza* 'Weizent' wird man annehmen können, daß es ein altes Lehnwort ist; die Wörter mit dem Lautwandel $z > r$ müssen jünger sein und sind wohl erst nach der Landnahme zu den Ungarn gekommen

Thus, Benzing's research (*ibidem*), together with the work by Mahmud al-Kashgari, convincingly demonstrates that the Hungarian vs Chuvash lexical affinities cannot be the result of Bulgar Turkic loan words into Hungarian, borrowed before 1073 AD. A similar conclusion has been reached also by Ramstedt (1957), who argues that the initial Proto-Turkic $/*y-/$ corresponds to Common Turkic $/y-/$ and Chuvash $/ś- \sim s-/$. However, Ramstedt (*ibidem*) concludes that the sound change $/y-/ > /ś-/$ is not older than the 13th century. This in turn calls into question the claim put forward by Gombocz (1912: 121) that cases such as Hungarian *szél* (see B3/11b) are derived from Chuvash, or its earlier variants. If Hungarian had borrowed directly from Proto-Turkic words with $/*y-/$, we would run into a problem of predictability, because initial Proto-Turkic $/*y-/$ may correspond to various sounds in Hungarian: $/d'/$ (compare B3/11a), $/s/ \sim /ś/$ (compare B3/11b), $/Ø-/$ (compare B3/11c) and $/ń-/$ (see B5/13). In sum, apart from the fact that the data, as documented in (Old) Chuvash, are not able to yield the expected form in Hungarian (according to the conventional model), the conventional system of sound correspondences runs into a number of serious anomalies and asymmetries. This shows that the whole chain of hypotheses associated with the concept of the presence of Old Turkic loan words of the Chuvash type in Hungarian is wanting.

Hungarian could have not borrowed the words of the Chuvash type after 1073 AD either, because at that time they would have left the Volga Region since about two centuries. It is also extremely unlikely that the Pechenegs or the Cumans would have brought the Old Turkic loan words of Chuvash type after 1073 AD to the Carpathian Basin – where the Hungarians had settled in the 9th century – as suggested by Benzing. This is because both the Pecheneg and Cuman are Shaz Turkic languages, and by no means Lir Turkic, Oghur languages. This being the case, the only solid, alternative explanation for the Hungarian *vs*

Chuvash phonological and lexical affinities is that it is the Hungarian substrate which is responsible for the (A) group of correspondences between Hungarian and Chuvash, at least for the cases in A1 and A3 (A2 is unclear, because the Chuvash equivalents are unknown). Menges (1968: 99) draws a similar conclusion, with the difference that, in his opinion, the Hunnic language is the major player in this context:

...we may well say that Chuvash with its *r* and *l* in place of Turkic *z* and *š* – the only consistent and truly regular feature in the phonology of Chuvash – stands completely apart from the Turkic group and in line with Mongolian and Tungus. Thus, this particular feature is best to be supposed as one of the more significant criteria for a strong, basically non-Turkic stratum in Chuvash and Proto-Bulgarian, which is to be identified with the Hunnic ethnic and linguistic element in the Proto-Bulgarian complex. The conclusion might be drawn that Hunnic was not Turkic, but another Altajic language which was, in this respect at least, most closely related with Mongolian and Tungus

5. An agenda for further research

5.1. In the previous paragraphs I have presented (relatively) systematic and regular lexical parallels between Turkic and Hungarian – parallels mainly of the Lir Turkic, and not of the Shaz Turkic type. As discussed, these parallels are conventionally accounted for by claiming that they are (all) the effect of borrowing into Hungarian, and from Lir Turkic only. In turn, this claim relies on the postulation of the existence of a WOT branch within the Turkic family. It is sometimes also assumed that West Turkic represents the oldest Turkic variant, due to its correspondences with Mongolian¹¹, if the latter is considered to be related to Turkic within the framework of the Altaic language family.

However, I have argued that these claims are untenable. There are, in fact, clear cut data on the basis of which it can be argued that it is Hungarian the donor language (assuming, for the sake of the argument, that

¹¹ As shown in the text, Mongolian patterns with the Lir-variant too. Ligeti does not offer a convincing explanation for this, but rules out the possibility that these variants had anything to do with Hungarian (Ligeti 1986: 52). Hungarian is sometimes missing within this pattern, as in Old Turkic *toquz* vs Chuvash *tāχār* 'nine'. It is reasonable to assume that these cases are instances of borrowings between Common Turkic and Chuvash, probably due to the 'Kipchakization' of Chuvash.

we are indeed dealing with borrowing¹²), triggering the Lir-isogloss in Chuvash. First, the time frame of the borrowing of the Lir-variant suggests that Chuvash cannot be the donor, as clearly argued for in Benzing (1942). Second, if we look at the Lir correspondences among Chuvash, Mongolian and Hungarian, we observe that the Hungarian counterparts are always available, whilst the Chuvash and Mongolian patterns are often incomplete. As a matter of fact, the conventional model has to rely on the establishment of 'reconstructed' forms for the missing parallels, and not on 'attested' forms, of course. In contrast, the thesis that it is Hungarian the language that functions as a substrate and donor language, does not encounter these difficulties. This being the case, there is no evidence for assuming that the direction of borrowing is from a Chuvash type of source into Hungarian.

At this point it is worth observing that the alternative explanation put forward here appears to be supported also by a socio-linguistic considerations. Németh (1991) argues for the existence of a *Sprachbund* in the area of the Volga Bends, pointing out that the exceptional position of Chuvash within the Turkic family is due to the mixing of Hungarian, Cheremis (/Mari), a Uralic language, and Kipchak Turkic, whereby the Hungarian substrate would have been particularly prominent. This being quite a plausible scenario, there is no need either for a WOT branch (as postulated by RT&B), or 'a dummy Bulgar Turkic category', that would be responsible for the Lir-effect in Hungarian. This conclusion ties in with the conclusion reached by Menges (1968: 99), although Menges assumes that the substrate in question is mainly of Hunnic origin, as discussed above.

5.2. The model of explanation proposed here, despite accounting for the relevant data in a more coherent and systematic way with respect to the conventional model, runs nevertheless into some anomalies that have to be accounted for through further research. First, consider the following two quotes by Marcantonio (2014:16 & foot note (34), respectively) on the topic of the Hungarian *vs* Turkic correspondences:

¹² The possibility that we are not dealing here with the process of borrowing, or, perhaps, not with 'borrowing only', but, rather, with a mixture of processes – borrowing, inheritance, share drifts, etc., as is typically the case in languages sharing many correlations – cannot be excluded a priori. However, this line of research cannot be pursued in this short essay.

It is certainly true that several of them [Hungarian vs Chuvash parallels] do indeed present phonological features that are characteristic of the West Turkic, Bulgharic languages, the so-called ‘rhotacism’ and ‘lambdacism’ isogloss, that is, the occurrence in Hungarian and (essentially) Chuvash (as well as, at times, Mongolian) of *r* and *l* as against the occurrence of *z* and *š* in all the other Turkic (‘Common Turkic’) languages, respectively [...]. However, there are plenty of relevant words or suffixes that, obviously, do not contain these sounds, or any other ‘Chuvash criteria’, so to say, and therefore it is not always possible to apply them for assessing the presumed Bulgharic nature of the loan word in question. This is also at times recognized in the specialistic literature; see for example Zimonyi (2012: 84 ff.) and Ligeti (1986: 36–48). Even Róna-Tas & Berta (2011) – who, as we have seen, ‘always’ reconstruct a West Old Turkic /Bulgharic form for the listed Turkic loan words of their dictionary – report 268 loan words that display no relevant criteria, as against 199 loan words that display Chuvash criteria. Not to count that the rhotacism and lambdacism isogloss, like any other phonetic /phonological (or any other type of) process in languages, are not always implemented regularly.

Rhotacism and lambdacism, however, do not always materialize (as one would expect). For example, Chuvash *r* may also correspond to Common Turkic *r*, as in *yur* vs *qār* ‘snow’ respectively; Chuvash *l* may also correspond to Common Turkic *l*, as in *śul* vs *yōl* ‘way’, respectively¹³.

Thus, the Hungarian vs Common Turkic correspondences that do not display the rhotacism or lambdacism isoglosses need to be accounted for in any case, independently of the postulation of the WOT branch. As a matter of fact (as the careful reader will easily realize) RT&B, despite arguing strongly in favour of the existence of the WOT vs EOT branch division (in line with conventional interpretation), list data that are clearly in contradiction with their own classification, as it emerges from what follows:

- for each single Hungarian word of Turkic origin listed in their dictionary, RT&B propose a ‘WOT reconstruction’, from which the Hungarian term would have been borrowed; however, scrupulously, the authors also report (plenty of) the actual, attested parallel forms from EOT languages;

¹³ Note however that these cases are few in number, so that fairly recent borrowings is the first hypothesis that comes to mind. ‘Kipchakization’ of Chuvash in these cases is a reasonable starting point for further research.

- from these attested EOT forms it comes up evident that almost all (if not all) of the reconstructed WOT forms are (very) similar or identical to the attested EOT ones (as already pointed out in the first quote by Marcantonio reported above) – this being the case also when the Lir *vs* Shaz alternations do actually materialize (as in *gyűrű vs yüzük, ökör vs öküz*, etc.).

This being the state of the art, the postulation of the WOT branch – the branch that would have been the ‘only source’ of the borrowing – however much drawing from a well established, high-level tradition of Hungaro-Turkic research, is, actually, superfluous (as already pointed out by Marcantonio (2014)). Even if the conventional empirical model did account for the data satisfactorily, one must bear in mind that the model itself is fraught with a major shortcoming at the theoretical level: the process of borrowing is, typically, the end result of an often long and messy process of contact and interference among real peoples, speaking real languages, and not the result of contact and interference among theoretical, abstract forms created by linguists – forms that are then supposed to have been borrowed in an orderly and systematic manner (see Marcantonio 2017b).

6. Conclusion

The data and arguments presented in this article strongly suggest the postulation of a Central Asian *Sprachbund*, that includes Old Hungarian, Old Turkic, Mongolian, etc. (in Marcantonio (2014) and Marác (2015) plausible variants of this basic scenario have been put forward; see also Marác 2017). This model¹⁴ appears to offer adequate explanation for the following phenomena, as observed above: a) the existence of minimal, Hungarian *vs* Common Turkic pairs only, the Chuvash and /or Mongolian counterparts being often missing; b) the close similarity of the Hungarian words of Turkic origin to their Common Turkic parallel – even when a Chuvash corresponding form is actually attested;

¹⁴ Notice that the claim made here regarding the existence of a Central Asian *Sprachbund* that would include also the early Magyars are at odds with the conclusion reached by B. Obrusánszky in this volume, according to whom “there are no ancient historical sources, archaeological findings or ethnographic traces that prove, or even just refer to, the existence of (part of) the Hungarians in Western Siberia, or the Middle Volga regions”.

c) the fact that Hungarian lines up with Mongolian when the rhotacism and lambdacism alternations occur. Last, but not least, *the claim put forward here that the r ~ z isogloss is a wide spread, Central Asian Sprachbund isogloss*, can also account for the final /-r/ in the root of the numeral ‘hundred’, as attested in Tundra Nenets, a Samoyedic language. The reconstructed proto-Samoyedic form is *yür (see Róna-Tas & Berta 2011, II: 1113). Within the framework proposed here, there is no need to assume that the Samoyedic form with final /-r/ has been borrowed from Old Chuvash, assumption that, in turn, presupposes an undocumented migration of the ancestors of the Chuvash people westwards, as Erdal (1993: 162) claims:

Es scheint alles für die Hypothese zu sprechen, daß Sprecher des Urwolgalbolgarischen in Südsibirien weilten, bevor sie in das Kaukasusgebiet gelangten. Dort dürften sie sowohl die Samojuden getroffen als auch den Mongolen zahlreiche lexikalische und grammatikalische Elemente übermittelte haben

Whatever the case, I shall leave an in-depth elaboration of this ‘Central Asian Sprachbund puzzle’ for future research.

References

- AYDEMİR, H. 2005. The main pillars of the Turkic rhotacism-zetacism, I. *sämiz, sämir-, sämri-, semre*. *Studia Etymologica Cracoviensia* 10: 15-34.
- BENCE, G.I. 2014. Legújabb viták az altaji nyelvek rokonságáról. *Nyelvtudományi Közlemények* 112: 1-5.
- BENKŐ, L. (ed.) 1967-1976. *A magyar nyelv történeti-etimológiai szótára I-III*. Budapest: Akadémiai Kiadó
- BENZING, J. 1942. Die angeblichen bolgarturkischen Lehnwörter im Ungarischen. *Zeitschrift der Deutschen Morgenlandischen Gesellschaft* XCVIII: 24-27.
- BERTA, Á. & RÓNA-TAS, A. 2002. Old Turkic Loan Words in Hungarian: Overview and Samples. *Acta Orientalia Academiae Scientiarum Hungaricae* 55 (1-3): 43-67.
- BUDENZ, J. 1871-1873. Jelentés Vámbéry Á. Magyar-Török szóegyezéseiről. *Nyelvtudományi Közlemények* 10: 67-135.
- CZUCZOR, G. & FOGARASI, J. 1862-1874. *A magyar nyelv szótára I-VI*. Pest: A Magyar Tudományos Akadémia.
- ERDAL, M. 1993. *Die Sprache der wolgalbolgarischen Inschriften*. Wiesbaden: Harrassowitz.

- GOMBOCZ, Z. 1912. *Die bulgarisch-türkischen Lehnwörter in der ungarische Sprache, Mémoires de la Société Finno-Ougrienne XXX*. Helsinki.
- IBRAYEV, S. (ed.) 2012. *Turkic World Almanac*. Astana: Turkic Academy.
- IBRAYEV, S. (ed.) 2013. *Түркі дүниесі: ғылыми-көпшілік альманахы: 2013*. Astana: Foliant.
- JOHANSON, L. 1998. The History of Turkic. In L. Johanson & É.Á. Csató (eds.), 81-125.
- JOHANSON, L. & CSATÓ, É.Á. (eds.) 1998. *The Turkic Languages*. London-New York: Routledge.
- KAŞGARLI, M. 1998. *Divanü Lûgat-it-Türk tercemesi*. Edited by Besim Atalay. Ankara: Türk Dil Kurumu.
- LAKÓ, Gy. & RÉDEI, K. 1972-1978. *A magyar szókészlet finnugor elemei I-III*. Budapest: Akadémiai Kiadó.
- LIGETI, L. 1986. *A magyar nyelv török kapcsolatai a honfoglalás előtt és az Árpád-korban*. Budapest: Akadémiai Kiadó.
- MARÁ CZ, L. 2012. The "Ugric-Turkic" War and the Origin of the Hungarian Language. *International Review of Turkish Studies* 2/4: 8-23.
- MARÁ CZ, L. 2012. Eurocentrism and the "Ugric-Turkic" Question. In S. Ibrayev (ed.), 360-369.
- MARÁ CZ, L. 2013. Towards a Theory of "Ural-Altai" Isoglosses: a Research Agenda for an Eurasian Linguistic Area. In S. Ibrayev (ed.), 92-103.
- MARÁ CZ, L. 2015. *Towards Eurasian Linguistic Isoglosses: the Case of Turkic and Hungarian*. Astana: International Turkic Academy.
- MARÁ CZ, L. 2017. *Elhallgatott közép-ázsiai nyelvrokonaink*. Budapest: Frig Kiadó.
- MARCANTONIO, A. 2002. *The Uralic Language Family. Facts, Myths and Statistics. Publications of the Philological Society. Vol. 35*. Oxford/Boston: Blackwell.
- MARCANTONIO, A. 2010. *On the Origin of Hungarian: Hungarian vs Turkic vs Finnish parallels*. Paper presented at the: 17th Conference of the Finno-Ugric Studies Association of Canada (FUSAC). Montreal 29-31, May 2010. Concordia University.
- MARCANTONIO, A. 2014. *The Linguistic Influence of Turkic on Hungarian*. Turkic World Almanac. Astana: International Turkic Academy. 1-22.
- MARCANTONIO, A. 2016. Magyar-türk nyelvi korrelációk: kölcsönzés, genetikai örökség vagy valami más? Paper presented at the conference: *A nyelvtörténeti kutatások újabb eredményei IX, 2016. április 27-28. University of Szeged, Hungary* [Abstract & power point presentation available in: <http://sites.google.com/site/angelamarcantonio/Home>].
- MARCANTONIO, A. 2017a. Magyar-uráli és magyar-török morfológiai megfelelések a tudomány jelenlegi állása szerint. In Ferenczné Szócs, É., Gazda, J. & Szabó, E. (eds.), *Kőrösi Csoma Sándor Gyökereink*. Kovászna: Kőrösi Csoma Sándor Közművelődési Egyesület. 85-102.
- MARCANTONIO, A. 2017b. Finnish and Uralic studies in Finland and elsewhere: Their contribution to linguistic knowledge. In B. Bíró, I. Piechnik & U. Tuomarla (eds.), *Finland Suomi 100: Language, Culture, History*. Kraków / Helsinki: Jagiellonian Library & University of Helsinki. 70-82.

- MARCANTONIO, A. 2017c. Uralic vs Turkic Morphology in Hungarian: The State of the Art. In A. Marcantonio 2017. *The Nature of the Hungarian vs Turkic Linguistic Correlations: Is Hungarian really a 'Proto-typical' Uralic Language?* Monograph supplement of *Rivista di Study Ungheresi* 2017, n. 16. Rome: 'La Sapienza' University Press. 45-100.
- MARCANTONIO, A., NUMMENNAHO, P. & SALVAGNI, M., 2001. The "Ugric-Turkic Battle". A Critical Review. *Linguistica Uralica* 37/2: 81-102.
- MENGES, K. 1968. *The Turkic Languages and Peoples: an Introduction to Turkic Studies*. Wiesbaden: Otto Harrassowitz.
- NÉMETH, Gy. 1921. On ogur, hét ogur és Dentümogyer. *Magyar Nyelv* XVII: 205-207.
- NÉMETH, Gy. 1991. *A honfoglaló magyarság kialakulása*. Budapest: Akadémia Kiadó.
- RAMSTEDT, G.J. 1957. *Einführung in die altaische Sprachwissenschaft I. Lautlehre*. Helsinki: Suomalais-Ugrilainen Seura.
- RÓNA-TAS, A. & BERTA, Á. 2011. *West Old Turkic: Turkic Loanwords in Hungarian, Part I: Introduction, Lexicon A-K*. Wiesbaden: Harrassowitz.
- RÓNA-TAS, A. & BERTA, Á. 2011. *West Old Turkic: Turkic Loanwords in Hungarian, Part II: Lexicon L-Z, Conclusions, Apparatus*. Wiesbaden: Harrassowitz.
- TEKIN, T. 1990. A new Classification of the Turkic Languages. *Türk Dilleri Araştırmaları*. 5-18.
- VÁMBÉRY, Á. 1870. *Magyar és török-tatár szóegyezések*. Pest: Nyelvtudományi Közlemények különnyomat.
- ZIMONYI, I. 1012. *A magyarság korai történetének sarokpontjai. Elméletek az újabb irodalom tükrében*. Szeged: http://reald.mtak.hu/597/7/dc_500_12_doktori_mu.pdf.

Are the Hungarians Ugric?

Borbála Obrusánszky

1. Introduction

1.1. Were the Hungarians originally Ugric, therefore, ultimately, Uralic (U) peoples? The answer to the above raised question is controversial among Hungarian scholars, because the evidence provided by the ancient historical sources and that (supposedly) provided by the linguistic model are in contradiction with one another. Árpád, the first royal chief of the Hungarian Kingdom, in the Carpathian Basin, claimed that Attila and the Huns were his ancestor, and that he came from Scythia. The Scythian and Hunnic genealogical tradition had been well known among the Hungarians of the Carpathian Basin for centuries (see also Cossuto in this volume). Indeed, both the Kingdom official documents and folk poetry clearly preserve this tradition (see below), and nobody had had any doubt about it for centuries (see Domanovszky 1933, Hóman 1925, Hölbling 1999, etc.). From folk music¹ (for which see Szabolcsi 1934), to ethnographical studies (Orbán 1868, Szendrey 1923/24:143-149), that have gathered items of Hunnic tradition among the Hungarians, to natural science research (archaeology, palaeo-anthropology and genetics (see Henkey (2002) and Fóthi (1997))), the evidence strongly suggests that the early Magyars have their origin in the Eurasian steppe, being directly connected with the Scythians and the Asian Huns (see Neperáczi 2017). Despite this long and coherent tradition, supported by a variety of research and records,

¹ Music is said to be the second language. There is no similarity between Hungarian and Uralic musicology; however almost the same melodies can be found from Central and Inner Asia, from Turkic and Mongolian tribes, as well as the Caucasus (see Szabolcsi 1934).

it is the linguistic U model that still prevails: the Hungarians were /are U people, whose ancestors came from Western Siberia or the Middle Volga regions (Fodor 2012: 119-129, 125-146).

1.2. In contrast to the testimony provided by the historical sources and the results of natural science research, the U linguistic theory does not provide any concrete, extra-linguistic evidence (historical sources and/or archaeological, anthropological findings, etc.), as to how and why the *Hungarian* peoples would fit into the U language family – neither could the theory provide this, having been founded ‘exclusively’ on linguistic data. As to the linguistic evidence itself, the reality is that it is rather wanting, for the following reasons. First of all, the conventionally identified lexical and morphological correlations are not statistically significant, as is now pointed out by several scholars (see for details Marcantonio 2002, chapter 4., and the other authors quoted there). Second, a good percentage of the (supposedly) U cognates and grammatical elements are to be found also in other, non-U languages, such as Turkic and Mongolian, this fact too being well known within specialist literature (see again Marcantonio 2002). Third, Hungarian does not share either its functional morphology (verbal and nominal case ending paradigm), or its derivational morphology, with any of the U languages, as shown again in Marcantonio (2002, chapter 8, and 2017) – shared morphology being widely considered to be a *contitio sine qua non* to establish language families. Last, but not least, the general typological features shared by Hungarian and the other U languages are also shared (all or in part) by the other Eurasiatic languages. Thus, the following questions should be asked:

- What could be the truth?
- How can this contradiction between the linguistic model – supporting the U origin of the Hungarians – and the evidence provided by the other, relevant disciplines – supporting their steppe / nomadic origin – be resolved?

2. The thesis of the Uralic origin of the Hungarians: Historico-political influences

2.1. The debate revolving around the issue of the origin of the Hungarians begun at the time of the Habsburg Empire, when the Habsburg House encouraged and supported some foreign scholars to question the long standing historical, cultural tradition according to which the

early Magyars were of Scythian /Hunnic origin. In the meantime, the language family-tree model was set up, and language families such as the Indo-European family, the Semitic family, the Ural-Altai and then Finno-Ugric (later on Uralic) family, etc. were established. However, the language family model itself is fraught with several shortcomings, as pointed out in numerous works addressing this issue (for a recent assessment see François (2014)). In addition, it is generally accepted (wrongly, in my opinion) that the results of the application of the comparative method do not need to be supported by extra-linguistic evidence, such as the historical and social background of the languages under investigation. This in turn means that peoples / languages, who lived far apart from one another in space and time may be – wrongly – included within the same language family. This is indeed the case of the Hungarians, this situation having been triggered, as mentioned, by the Habsburg House. More in details, after Hungary lost its independence war of 1848/49, old members of the Hungarian Academy of Science were removed (in 1858) and young scholars, loyal to the Habsburg administration, were appointed (see Kónya & Kosáry 1975). Among these newly appointed scholars there was the German József Budenz, who was invited to Hungary with the task of setting up the Hungarian language system into the Ugric (read ‘Finno-Ugric’) model, although he did not speak either Hungarian, or Finnish and related languages. He did not know any eastern languages either, such as Turkic and Mongolian. Excellent Hungarian linguists and ‘orientalists’ – such as the jurist and philologist János Fogarasi, the linguists and turcologists József Thury and *Ármin Vámbéry*, the linguist Gábor Bálint², etc. – vainly protested against what they considered to be the wrong course of action, the wrong linguistic method of analysis (as adopted by Budenz) and, consequently, the wrong results. In the meantime, Pál Hunfalvy (whose original German name was Paul Hunsdorfer), the librarian in chief of the Academy, wrote an essay titled the ‘Ethnography of Hungary’ (Hunfalvy 1876), where he provided the framework for a new historical concept, a new historical background regarding the Hungarians, despite the fact that he could not support his claims through the necessary records and historical sources. By doing so, he bypassed the whole of the Hungarian historical / cultural tradition, questioned the authen-

² For example, Bálint had a special talent for languages and linguistics: he spoke around 30 languages and was able to compare the various systems of the Eurasian languages.

ticity of the Hungarian *Royal Chronicles* and the Hungarian *Diplomas*, and created a fictitious homeland, in line with the newly established linguistic model. Scholars such as Thury (1897b & 1890), Nagy (1909) and others since then pointed out these mistakes – once again, in vain.

2.2. Despite the pressure exercised by the Habsburgs, during the 20th century it became clear that the tenets and predictions of the U theory could not be supported by evidence drawn from other fields of study (not to count the scantiness of the linguistic evidence itself, as discussed above). Certainly, genetic, cultural and linguistic kinship may be disconnected from one another; the similarities in languages do not necessarily mean genetic relationship among the peoples that speak them, and, *vice versa*. This being the case, some Hungarian linguists (such as Honti 2010), claim that only the ‘linguistic kinship’ is of relevance for the U language family, and, therefore, Hungarian. However, I personally disagree with this statement, that, in fact, appears to be an *ad hoc* justification for the lack of any (sort of) extra-linguistic evidence in support of the existence of the U family. Other kinds of ‘kinship’ are relevant too, and need to be taken into account (as pointed out by ethnographic studies; see Bodrogi 1961), including, for example, ‘artificial kinship’, encompassing sworn brotherhood or alliances signed by blood – a common custom among steppe peoples, from Scythians, to Huns, Hungarians, Mongols, etc. Thus, it is vital in this context to take into account evidence drawn possibly from all the humanistic and scientific disciplines that might have a bearing on linguistic issues.

3. Proto-Uralic speech community: Are there historical records?

Let us now investigate whether there are any historical records pointing to the existence of a unified, ethno-linguistic proto-U community, to be located somewhere between Western Siberia and the East European plains. Within this context, one should also examine whether *Magna Hungaria*, a major site of the ‘ethnogenesis’ of the Hungarians (according to official Hungarian historiography), really existed (topic that will be dealt with in the next paragraph, par. 4.).

3.1. Based on a previous study by Tomaschek (1889: 7-16), Klima (1999, 2002 & 2005) lists all the available historical sources that *have*

been interpreted as referring to one or the other of the U people. Klima (1999), quoting the Greek historian Herodotus (*The Histories*, including its account of the *Greek-Persian Wars*; see Muraközi (2002) & Herodotus (1909)), claims that Herodotus makes a reference to some U peoples in his Book IV. In reality, Herodotus refers to the *Scythians* and their migrations: they were expelled from their old residence by the *Issedon-s*, who occupied the home of the *Cimmerian-s* – this took place in the 7th century BC. According to other historical sources, the nomadic *Scythians* in Asia were chased out by the *Massageta-s* (Herodotus IV:11; see also Télyf 1863:16)³. Thus, according to Herodotus, only in the 7th century BC a great migratory movement of peoples took place, when part of the *Scythians* moved from Western Siberia to the Caucasus and the Easter-European Plains. This in turn means that, instead of the fictive U community, traces of the *Scythian* peoples can actually be retrieved, on the ground of historical, as well as archaeological research (according to Koryakova & Epimakhov 2002). Indeed, the denomination of that long historical period extending from the 7th to the 3rd century BC – in Siberia, Eastern Europe, Middle Volga regions, the Caucasus and even in the Carpathian Basin – is the ‘Scythian Age’. Then the ‘Sarmatian Age’ takes over, from about the 3rd century BC up to the 3rd century AD. This means that the U linguistic theory has actually located the home land of the Hungarians in what was, in fact, the land of the *Scythians*, both in terms of space and time. As a matter of fact, linguists claim to have found many Iranian loan words in present day U languages (Fodor 1973, Vékony: 1997: 265-266), and, surely, the Persians did not live in the far territory of Siberia, neither did they establish a state before the 6th century BC, in the southern part of modern Iran.

3.2. The steppe culture and civilisation must have had an impact on Siberia – mainly the southern and central territories – since the Early Bronze Age up to the Late Middle Ages and the times of the Russian occupation (Naumov 2006). As a matter of fact, the Scythian, Hunnic, Turkic and Mongolian Empires covered most of the Eurasian steppe, including the Ural and Volga regions, for about 3,000 /3,500 years. In contrast, the people who are classified as U were discovered only from

³ The *Massageta-s* lived in the Trans Caucasian and Central Asia region, at the time of Herodotus.

the Late Middle Ages, and do not feature in any ancient historical sources. The *Scythians* ruled both Siberia and those territories where the ancestors of the Hungarians lived – according to the Hungarian royal historical sources (Pauler & Szilágyi 1900)⁴. The same situation can be observed for the Volga River region: it has been populated by Scythian peoples since ancient times. It is therefore likely that Siberian peoples such as the Khanty and Mansi (whose original land was located around the Northern Sea, according to Klima⁵ 1999) had been influenced by the highly developed steppe societies for about 3,000-4,000 years. Unfortunately, we do not really know the ancient Scythian and Hunnic languages, because the relevant ancient Greek, Chinese, etc. sources preserved only words, and some expressions, but we can assume that their mutual, long lasting cultural and commercial connection proceeded hand in hand with a mutual influence on the respective languages. As we have already mentioned, and shall see also below, the archaeological findings suggest that the *Scythians* and the Huns lived in vast territories of Siberia for thousands of years, so that they could have represented the bridge among the far regions of Inner Asia, Siberia and the Carpathian basin. In contrast, there is no *direct* connection between the early Hungarians and the people of Siberia.

3.3. Let's now quote some more studies, according to which there would be ancient historical sources that refer to some of the peoples that are now classified as U, in particular the Hungarians. The Greek historians wrote about the territories of the Northern Black Sea and the Caucasian region, but had no knowledge about further away territories, like the territory of the Ural Mountains. Nevertheless, some scholars within U studies tried to identify in these sources a more or less direct reference to the Hungarians. For example, Harmatta (1997: 120-123) thought that the *Yurka* peoples (for which see also Cossuto in this volume) featured in Herodotus, could be the ancestors of the Hungarians, but this claim lacks evidence, since the Greek historian wrote clearly that the *Yurka* (~ *Jürka*) people were a 'Scythian nation'⁶.

⁴ Both the *Gesta Hungarorum* and the *Chronicum Pictum* state that the Hungarians lived in Scythia, their centre being the Maeotic Swamp; see below.

⁵ According to Klima (1999: 47-8), Abu-Hámid al-Garnati mentioned the *Jugra peoples* or the ancestors of the Ob-Ugric peoples.

⁶ Compare the following quote from Herodotus (IV: 22): "North of the *Budini* the land is uninhabited for seven days' journey; after this desolation, and somewhat

Some scholars (for ex. Zsirai 1930), identify the name *Yurka* with the place name *Yugra* ~ *Yugri*, although other scholars, such as Munkácsi (1894 & 1895) and Vásáry (2008) deny this. The Slavic chronicles of the 11th /12th centuries mention the *Yugra* as people living in the Northern, Arkhangelsk region, around the Pechora and Dvina River, in the western side of Ural Mountains (Klima 2002:108)⁷. Old maps do not show the name *Yugria* in the present-day Khanty regions⁸. As to the Khanty peoples (one of the languages/peoples classified as Ugric within the traditional U family tree, together with the Hungarians), according to Reguly (1844), they lived near present day Tatarstan, inside the former Golden Horde, which is situated in Eastern-Europe, not in Siberia⁹. Records from the *Novgorod First Chronicle*, in the year 1265, show that *Yugra* was a tributary to Novgorod. Military expeditions by Novgorod (for example, those mentioned in the *Novgorod Chronicles* of the year 1323, 1329 and 1364), and later on by the Muscovites, are also reported. Nowadays the *Yugria* ~ *Yugra*, etc. geographical names can be found in the Vologda region or the Russian city of *Perm*, north of Tatarstan, westward to the Ural Mountains. All this confirms the accounts of the historical sources, according to which the *Yugria* region was located nearby Novgorod. The *Yugor vs Ugric* identification came up much later than the Middle Ages, being in fact created by an envoy of the Habsburg House, Sigismund Herberstein, who visited Moscow (but not the far regions of Russia), in the 16th century (Herberstein 1549). Thus the name *Yurka* and *Yugria* (with its variants) have no connection with the Hungarians, the similarity

more toward the east wind, live the *Thyssagetæ*, a numerous and a separate nation, who live by hunting. Adjoining these and in the same country live the people called *Iyrkæ*; these also live by hunting [...]. Beyond these and somewhat to the east live the *Scythians* again, who revolted from the Royal *Scythians* and came to this country”.

- ⁷ Klima (2002: 108) states that, since 1445, the Ustug City waged war against Novgorod (today: Vologda), a war that also reached *Yugra*. Ustug has the modern name ‘Velikiy Ustug’, and is a city in Russia, in the Vologda Region, along the Rivers Yug and Suhona.
- ⁸ Solinus’s map of 1538 does not mention *either Khanty or Yugria, as being near the Ural Mountains*; see: https://www.raremaps.com/gallery/browse/creator_id/2773. Ortelius’s map of 1562 describes *Ioghoria* as being located in the far northern part of Russia; see: <https://www.raremaps.com/gallery/enlarge/50396>.
- ⁹ Reguly (1844: 316-317) sent reports from his own journey as follows: “Among the Cheremis peoples there is a nation, whose name is Ostyak. People from Rostov gave name for them, who escaped from the converting to Russian belief, or Christianity, and settled down in the Land of the Bolgars and Golden Horde, in order to live under the Kazan Khan”.

in names being only a coincidence. There are also linguists who have observed this. For example, Kálmán (1998: 395) states the following:

This [...] belief in the Hungarians' close relationship with the Ob-Ugrians is based not so much on scientific arguments as on the fact of the Latin names Hungarus and Hungaria [...] the land of the Voguls [...] and Ostyaks [...] was referred to in Russian and West European sources as Ugra, Yugra, or Yugria [...] it was taken almost for granted that Yugria in the Ural region was the Hungarians' ancestral home

4. Historical sources in support of Magna Hungaria?

4.1. Turning now to the issue of *Magna Hungaria*, there are two questions to ask. First, did a region bearing such a name actually exist in ancient times – a region supposedly located between the Volga and Kama Rivers, from about 500 BC until 550 AD? Second, if it did exist, was it really one of the ancient homelands of the Hungarians, at some point during their westward migrations, as claimed by Hungarian historiography? The reality is that we have no historical sources that could confirm either the existence of such a region, or the dwelling of a Hungarian community in this area (this thesis is not supported by local Tatar or Russian scholars either, any more, according to Wheatcroft (2002)). Muslim geographers from the 10th/12th century wrote extensively about that territory, because the *Bulgar-s* converted to Islam in the year 922, and maintained constant contact with Muslim communities; however, no one of them knew about *Magna Hungaria* (see Kmoskó 1997, 2000 & 2004). This denomination appeared relatively late, in the course of the 13th century, when Christian monks began to use it to refer to the ancient lands of the Hungarians. As a matter of fact, neither the famous Monk Julianus, a Hungarian Dominican missionary who, supposedly, found a group of Hungarians in *Magna Hungaria*, nor the later missionaries, gave exact information about that land, but only general, vague descriptions.

4.2. The historian Tacitus¹⁰ mentions the name *Fenni* (Tacitus, *Historiae* I: 45-46). The geographer Claudius Ptolemaeus (*in his geographical work*), also mentions a population he calls *Phinnoi*¹¹, as well as *Sarmatia*,

¹⁰ Publius Cornelius Tacitus, one of the greatest Roman historians, was born in 55 AD and died in 117.

¹¹ http://penelope.uchicago.edu/Thayer/E/Gazetteer/Periods/Roman/_Texts/

with this referring to the Eastern European region, from the Black Sea to the Vistula River – the land of steppe horsemen¹². However, once again, neither of them makes any mention of *Magna Hungaria*. The next, significant turn took place in the course of 3rd century AD, when the Goths moved southward, from Scandinavia, toward the Eastern European Plain, and reached the Black Sea region. On their way to their destination they must have crossed the Volga region, which was an excellent water route for trading. Their historian, Jordanes (in his *Getica*; second half of the 6th century), did not mention *Magna Hungaria* either, although the Goths must have crossed the region that bears this name, because it is located on their way to the Caucasus and Black Sea. In contrast, Jordanes recorded some names of people who lived nearby (now classified as U), namely *Meren-s*, *Morden-s* and others, as well as the Finns and Estonians¹³. In addition, Jordanes (who knew Priscus' 5th century account of the Huns¹⁴) recorded the exact place of the peoples he calls *Hunuguri*, that is, the Hungarians (for the identification of the *Hunuguri* with the 'Hungarians', see discussion below). This place, in his opinion, was located next to the Black Sea¹⁵, where they would have lived at least since the 5th century. This being the case, the linguistic model according to which the ancestors of the Hungarians moved away from *Magna Hungaria*, right in Jordanes' times, cannot be true.

Ptolemy/3/5*.html

¹² Compare the Ancient Map of Russia by Abraham Ortelius: *Russiae, Moscoviae et Tartariae Descriptio Auctore Antonio Lenkensono Anglo edita Londini, 1562 & dedicata illustriss D. Henrico Sydneo Wallie presidi*.

¹³ Compare Jordanes (III: 21): "There also are other peoples. There are the *Scripi-Fennæ* (= 'Schreit-Finnen', 'Walking Finns', i.e., 'Skiing Finns', 'Lapps', that is, 'Ski-users'), who do not seek grain for food but live on the flesh of wild beasts and birds' eggs; for there are such multitudes of young game in the swamps as to provide for the natural increase of their kind and to afford satisfaction to the needs of the people". Compare Jordanes (V:36): "But on the shore of the Ocean, where the floods of the river Vistula empty from three mouths, the *Widi-warii* dwell, a people gathered out of various tribes. Beyond them the *Æsti* [i.e., the ancestors of the Estonians], a completely peaceful folk, likewise hold the shore of the Ocean". Compare also Jordanes (XXIII:116): "Among the tribes he conquered, were the *Gothi-Scytha* in Aunxeis, the *Wasinabroncæ*, *Merjans*, *Mordjans*, *Inmiscareis* (the *Cheremis* peoples), *Rogas*, *Tadzans*, *Aþaul*, *Nawego*, *Bubegenæ* and *Coldæ*. <http://www.harbornet.com/folks/theedrich/Goths/Goths1.htm>.

¹⁴ Priscus of Panium was a 5th century Roman diplomat and Greek historian and rhetorician (or sophist).

¹⁵ Regarding the Black Sea region Jordanes (V: 37) states: "Now the *Hunuguri* are known to us from the fact that they trade in ermine pelts. The audacity of the men mentioned above has intimidated them".

As a matter of fact, starting from the late 5th century, the *Hunugur*-s or *Onogur*-s are often included in the historical accounts by writers who lived in the Black Sea and Azov Sea regions. At this point, regarding the denomination *Hunuguri* ~ *Onogur*, the reader might ask how one can be sure that by such a denomination the historical sources referred to the Hungarians (as claimed here), and not to the Bulgars, as generally claimed elsewhere¹⁶ (see also Cossuto in this volume, who points out the “continuous tradition of the *Onugur-Bulgar* component of the would-be Hungarians”). The answer to this question will be put forward below, in paragraph 5.2.

4.3. In summary, on the basis of what has been argued above, it can be stated that there are not totally reliable items of information regarding the early Hungarians – not to count the fact that they lived and moved within steppe tribal confederations. This is why, in my opinion, one should take into account *also* the information provided by the official *Royal Chronicles*, that were written in the Hungarian royal court in the Middle Ages, representing the officially accepted documents¹⁷. They include: the *Gesta Hungarorum*, by the so-called *Anonymous Magister P.* (12th century), the *Royal Chronicles*, by the literate clerk Simon de Kéza (~ Simon Kezai; 1282-85), at the Hungaro-Cuman Court of the King Ladislaus, and the *Chronicum Pictum* (see *Chronicum Pictum* (1987)). *These Chronicles* link their contemporary *magyar peoples* not to the forest hunters *Budini*, *Thyssagetæ* or *Iyrcae*, but to the Scythian Steppe peoples, including the *Savards*, *Huns*, *Alans*, etc. – this discourse being in line with the sources and arguments put forward above. Referring to these *Chronicles*, some (last centuries as well as contemporary) Hungarian historians, such as Hunfalvy (1876), Kristó (1996), Szűcs (1993), argue – without the support of clear cut evidence – that the Hungarian Kings created a false historiography in order to trace back their own lineage to the Huns¹⁸ (even if some of the reported stories appear to be con-

¹⁶ The Bulgars (also referred to as: Bulghar, Bulgari, Bolgar, Bolghar, Bolgari), were Turkic semi-nomadic warrior tribes who flourished in the Pontic-Caspian steppe and the Volga region, during the 7th century. According to some researchers, for example Hyun (2013), their roots can be trace back to Central Asia.

¹⁷ Thury (1897b) argues that oral history had been very popular among the Hungarians for a long time; they recorded events and great victories in songs even for the Ottoman period.

¹⁸ Although there are certainly mistakes in the Hungarian Ancient *Chronicles*, including in the recording of geographical names (also because of mistakes in the copying

firmed by archaeological studies; see for ex. Botalov (2013), as well as, again, Györfy (1986: 205)). Finally, one may observe that the Maeotian Swamps appear to be the centre of the steppe peoples since the Bronze Age. Herodotus¹⁹ states that the Royal *Scythians* lived there, and archaeologists have found rich necropolises in the area. The Asiatic Huns too created their own centre in that area, as testified by the 6th century historian Procopius (see Procopius 1914: IX / 6; see also 1954): “From the Caspian Gate as far as Maeotia all nations of Huns lived” (this being the same territory mentioned by the Hungarian *Chronicles*). Similarly, all Hungarian *Chronicles* of the Middle Ages record that this place – a place serving as the centre of steppe people for centuries – was the dwelling place for their ancestors.

Whatever the case, and even if the Hungarian *Chronicles* were all wrong, in all or most aspects of their narrative, two relevant ‘facts’ emerge clearly, nevertheless: a) no ancient historical sources or legends are known that connect the Hungarians with the Siberian or Volga region; b) no mentions whatsoever are to be found in any way referring to a proto-U community, or languages, or ‘culture’ (as ‘assumed’ within the framework of the linguistic model), although mentions are to be found of single peoples that are now classified as belonging to the U family.

5. The names of the Hungarians in the historical sources

5.1. Savard, Sabir, Siyavurdi

The Byzantine Emperor Constantine Porphyrogenitus VII composed his work *De administrando imperio* (DAI) in 948-952. He mentions the *magyar-s*²⁰ among his list of ‘Turkic’ peoples / clans, giving some insight into their ancient history. According to the Emperor, the Turks had as their ancient name that of *Savardoi-Asphaloi* (DAI, chapter 38), and *this is* why scholars consider the ‘Savard-question’ as part of the ancient Hungarian history. Thus, the relevant question is: who are these

process), the main thrust of these texts is right, if we compare them with the information provided by other, foreign sources.

¹⁹ Herodotus (Book IV: 19) mentions that territory as the centre of the Royal Scythians.

²⁰ To be precise, the noun *magyar* *per se* does not feature in the DAI, that mentions instead the name *Megérη* within a list of Turkic clan /chief names – several of these names being Hungarian proper names still used nowadays; for details see Marcantonio (2002: 218 ff). See also Ligeti (1986: 400 ff.).

Sabirs and *Savards*? A nation that migrated from Western-Siberia, as Németh (1930), Czeglédy (1943) etc. claim, or autochthonous people of the Caucasus, as other scholars claim? Actually, most of the available sources, as well as scholars' research, indicate that they are, basically, autochthonous peoples of the (Trans) Caucasian area, that is, Modern Azerbaijan, Georgia and Armenia. These sources and scholars include: Thury (1897a:320-21, 398-403)²¹, Télfy (1863:7)²², Geybullayev (1986: 33), the Byzantine author Theophylaktos Simokattes (*History* VII: 8), who talks in terms of *Onogurs*, *Sabirs* and other 'Hunnic nations'. Similarly, Menandros (the Byzantine historian Menander Protector), in his *Historici Graeci Minores*, recorded the name *Unigur* (that can be identified with *Onogur* or *Hungarian*) in connection with a delegation of Turks in the year of 569/576 (for which see Györffy (1986: 80)), etc. Also the Muslim geographers of the 10th /12th centuries mention the *Siyavurdi*, that is, *Savards*, as peoples that live in the historical *Udi* province, eastward of Tbilisi and westward of the Tovuz and Semkir City, the latter located nowadays in Azerbaijan – as reported in Al-Hamadani's *Book of Countries* (see Kmoskó 2000: 147), and testified by the presence of many geographical names such as *Sabir* and *Madjar*. Thus, it seems clear that: a) the peoples named *Savard* ~ *Sabir* lived in the Caucasus, being probably autochthonous (according to most sources); b) there does not seem to be enough and /or clear cut historical sources to prove the Siberian origin of these peoples; c) the word *Sibir*, on the other hand, derives from *Buryat sheber* ~ *shiber* 'taiga, marsh', a word that spread over along with the Mongolian invasion of the 13th century (Alexandrovna, 2015: 26).

To conclude: the *Savards* are an important part of the Hungarian *Ethnogenesis*, and they are to be connected *not* with the Siberian region, but with the Caucasian region.

5.2. Hungaria ~ Onogoria

The name of the Hungarians all over the world is: *Hungary*, *Ungarn*, *Hongrie* etc. The first Hungarian royal dynasty also named themselves this way, and bore the title *rex Hungariae*, an expression that features

²¹ The author assumes that they were descendants of *Nimrod*, legendary King of the Hungarians and other nations of Caucasus.

²² Herodotus (I:104) states that: "From Lake Maeotis to the Phasis River and Colchis, the distance is 30 days for good pedestrians. From Colchis it is not far to reach Media; only one nation lives there, named *Saspir*. Those ones who pass their country, reach Media".

in the Caucasian region since the Early Middle Ages. Jordanes (V: 37) knew that a Hunnic tribe, the *Huniguri*, lived next to the Maeotic Swamp. The *Suda-Lexicon*²³ – drawing information from Priscus – recorded that the *Onogurs* sent envoys to the Byzantine empire in 463, and the peoples named *Onoguri* or *Hunnuguri* were described as ‘Huns’. Similarly, Theophylaktos Simokattes (*History* VII: 8), wrote that the *Barsil-s*, together with the *Onogurs*, *Sabirs* and other ‘Hunnic people’, established an alliance with the *Avar-s*, who came from the East in the course of the 6th century AD. It has been widely claimed that the *Onogurs* are to be connected with the *Bulgar* Turkic group, although there is no strong evidence for that; thus, it is really important to find out – on the basis of the historical sources – who these peoples really were. The *Bulgars* had this name as their own name since the 6th century AD, and never called themselves *Onogur*, as foreign sources (such as Priscus and Jordanes) did. Theophylaktos Simokattes (*History* VII: 8), who wrote his work in the first half of the 6th century, makes a difference between the peoples called *Onogur* and those called *Bulgar*: the *Bulgars* lived under the *Avars*, whilst the *Onogurs* and *Sabirs* are listed as ‘Hunnic people’. He also wrote that the *Avars* conquered the *Bulgars*, but forged an alliance with the (Hunnic) *Onogurs* and *Sabirs*. The *Onogundur vs Bolgar ~ Bulgar* equation appeared, only for a short period, during the 7th century, with the meaning: ‘the *Bulgars* who belong to the *Onogurs*’, expression that appears to indicate that the *Bulgars* were not independent people at that time, but lived under the *Onogurs*. In other words, the *Bulgars* were independent only for a short period, and never used either the *Onogur* or the *Onogundur* denomination for themselves, but only *Bulgar*. Moreover, a letter by Joseph, King of the Khazars, mentions the name *Bulgar* and *Ungri* separately, in this way suggesting that they were two different nations, with two different names (see Kohn 1881: 7-8). *Ravennatis Anonymi Cosmographia* (for which see Schnetz (ed.) 1990: 46), in the 9th century, mentions the *Onogoria* country as being located next to the *Pontus* (/ Black) Sea, near the Maeotic Swamp²⁴, or Azov Sea.

²³ The *Suda* or *Souda Lexicon* is an extensive, 10th century Byzantine encyclopedia of the ancient Mediterranean world, formerly attributed to an author called Suidas. It is an encyclopedic *Lexicon*, written in Greek, with 30.000 entries, that preserves (also) items of information from lost sources, including some fragments from Priscus’ lost work.

²⁴ Compare the following quote: “Item iuxta mare Ponticum poitur patria quae dicitur *Onogoria*, quam subtilius Livianus philosophas vicinam paludis Maeotidae”.

As to the Hungarians, they bore the name *Onogur* ~ *Hungarus*, and not only in the Caucasian region, but in the Carpathian Basin region too, as testified by several sources. For example, Merovingian sources since the 6th century record the name *Hungarus* next to *Pannonian Hun*; several Western European sources, as collected by Király (2006)²⁵, knew the name *Hungarian* even before Árpád's campaign of the 9th Century; *Nestor's Chronicles* mention that the Hungarians remained in the northern part of the Black Sea, between the Dniester and Dnieper River, where a 'Hungarian Mountain'²⁶ could be found (see also Ferincz (2015: 16-17, 34-35). The City of Kiev too has an 'Ugor hill', or 'Hungarian-hill'. The name *Hungarian* ~ *Hungarus* appears still in the 14th century: Pope Joannes XXII issued a *bull*a in 1329, where he mentions the Hungarians in the Caucasus as '*Ungari Asiatici*', pointing out that they lived next to the *Alans*, in the Northern Caucasus.

In summary: with regard to the names *Onogur* ~ *Hungar(us)*, it appears that only the Hungarians or *Magyars* bore them, both in the Black Sea region and the Carpathian Basin region. In contrast, the *Bolgars* did not use that ethnonym.

5.3. Magyar ~ Madjar ~ Maggariya

Most scholars of Hungarian or U studies claim that the Hungarians' self-denomination, *magyar*, is etymologically connected with the self-denomination of the Siberian Mansi peoples, that is *mańci*. Certainly, the two names display a superficial similarity; however, the established etymology is fraught with several shortcomings, as discussed in Marcantonio (2002: 212 ff.); see also Marcantonio (2014), in particular footnote (20), for a comprehensive discussion of the *mańci vs magyar (alleged) connection*; see also Czeglédi (2012: 28-48). As a matter of fact, as an alternative, a Turkic etymology has been proposed by two prominent Hungarian turkologists, Ligeti (1986: 400), and, more recently, Berta (2010), who both argue that *magyar* is a noun of Turkic origin – see again Marcantonio (2014) for details. Even if the conventional U etymology were correct, this similarity of nouns, on its own (without the support of

²⁵ Király (2006) collected the whole of the European sources that mention the Hungarians (Slavic, Merovingian, Latin, etc. accounts).

²⁶ According to *Nestor's Chronicle*, a 'Hungarian or Ugor' hill is to be found in Kiev. That hill still exists in Kiev today.

additional evidence, such as historical sources and extensive linguistic correlations), could not prove that the two groups of people are genetically related, of course. As to the relevant, early documents, the name *Madjar* is reported as being spread in the area of the eastern part of the Black Sea since the Early Middle Ages – right there where the previous settlements of the Hungarians are supposed to have been, according to the Hungarian *Chronicles*. Let us see some examples. The first document featuring the name *Madjar* is *Derbentname*²⁷, a source dating back to the 6th century, when Chosrou Anusirvan, the Persian Shah, built fortresses against the Huns. Two of these fortresses have a special name: *Ulu Madjar* or *Big Madjar* and *Kichi Madjar* or *Small Madjar* (see Helilov & Nyitray 2008: 54-61) – there are numerous such names that relate to the Hungarians, until the 14th century, spread over the Black Sea and Caucasian regions. Thury (1897a: 324) found the name *Madjar* also in the dwelling places of the *Savards* or *Sevordi-s*, and this was confirmed later on by Geybullayev (1986) and Helilov & Nyitray (2008). The Muslim geographers knew that the Hungarians lived in the Black Sea region in the 10th/12th centuries, already provided with a developed social organisation or kingdom: they locate the nation *Maggariya* near the Black Sea, near the Maeotic Swamp region. For example, Ibn Rusta (see Kmoskó 1997, I/1: 207-209), from the 10th century, wrote that the *Maggariya* nation lived along the eastern shore of the Black Sea. Al-Bakri (Kmoskó 1997, I/2: 256) gave us some special items of information: the *Maggariya* nation lived between the *Pecheneg-s* and the *Bulgars*, neighbouring with *Byzantine Rum* and *Alania* – although the ‘*al-ungalus*’ people, or Hungarians, lived in the neighbourhood of the Franks (Kmoskó 1997: 256-257). These sources wrote about the Hungarians or *Madjar-s* after the biggest part of them had moved into the Carpathian Basin, which in turn means that in the 11th/12th centuries the Hungarians lived in two major regions: the Caucasus and the Carpathian Basin. In addition to the Muslim sources, also the *papal* court gives us some items of information regarding a Caucasian Hungarian Kingdom, to whose king a papal *bull*a was sent (see Bendefy 1941: 58). Zichy’s expedition of 1895 found some place names relating to the Hungarians as *Madjar* in Northern Caucasus, around the Kuma River (see Szádeczky-Kardoss 2000: 91-95). To conclude this list,

²⁷ This is an ancient record, that summarises the history of *Derbent* City (nowadays in the Republic of Dagestan), located on the Caspian Sea, north of the Azerbaijan border, and surroundings, from the 5th to the 11th centuries.

one may observe that the Hungarian *Chronicles* refer to the Hungarians in two ways: as *Hungari* and as *Mogor* ~ *Moger* ~ *Magor*²⁸. In particular, Simon Kezai, in his *Prologus* (see Pauler & Szilágyi 1900: 480) describes the land of the ancient Hungarians as *Mogoria*.

Thus, according to the historical sources (since the 6th century up to the Early Middle Ages), the name *Madjar*, and its variants, appear to be spread all over the Caucasus. This contradicts the claim made within the U theory, according to which the name *Magyar* is to be connected with the self-denomination of the Siberian Mansi. Indeed, this etymological connection has been established by linguists in the 19th century, and it is not a correct etymology either, as we have discussed above. At this point it could be objected that, in fact, the name *Madjar*, and its many variants, are to be found as ethnonyms, toponyms and hydronyms also in the Volga / Kama and Bashkiria area, the present day Republic of Bashkortostan (see for details Marcantonio (2002: 219 ff.)). This is certainly true; however, the fact remains that there are no old historical sources in support of the existence of an 'ancient' area called *Magna Hungaria*. The steppe people migrated into the Kama /Volga and Bashkiria area from the Caucasus, and this is the reason why we find similar tribal and geographical name, such as *Bulgar*, *Mad(i)jar*, etc. Whatever the case, the geographic names and the medieval report from the 13th /14th centuries are not relevant sources for investigating the *ancient* history of the Hungarians.

6. Conclusion

It is hard to summarise in a coherent way the numerous – at times unclear or even discordant – items of information and historical data relating to the early appearance of the Hungarians on the scene of History, particularly within the framework of this short article. However, from the sources mentioned here (and there are many more), three 'facts' emerge clearly. First, both western and eastern sources, as well as the Hungarians' own *Chronicles*, make the very same basic claim: the earliest Hungarian settlements, since the 4th / 5th centuries AD, are to be found in the Caucasus / Black Sea region. Second, there are no ancient historical sources, archaeological findings or ethnographic traces that

²⁸ We have the name *Mogor* in the *Prologus* by Kezai (Pauler & Szilágyi 1900: 478), the name *Mogeri* in the *Prologus* by Anonymus (Pauler & Szilágyi 1900:329), and *Magor* in *Chronicum Pictum*, 2.

prove, or even just refer to, the dwelling of (part of) the Hungarian peoples in Western Siberia, or the Middle Volga regions. Last, but not least, there are no references in the historical sources to any pre-historical or historical U population / speech-community, or any (material and /or spiritual) U ‘culture’.

These findings clearly contradicts the tenet of the U origin of the early Magyars. This being the state of the art, I believe we should seriously rethink this – by now long standing – issue. Perhaps, some new scientific fields, such as genetics, will help to cast some light in this direction.

References

- ALEXANDROVNA, S.O. 2015. *Ethnic History of Russia*, I. Moscow Pedagogical Institute.
- ANONYMUS. GESTA HUNGARORUM. <http://mek.niif.hu/02200/02245/02245.htm>.
- BÁLINT, G. 1901. *A honfoglalás revíziója*. Kolozsvár/ Budapest: Magánkiadás.
- BENEFY, L. 1941. *Kunmagyaria. A kaukázusi magyarság története*. Budapest: Cserépfalvy.
- BERTA, Á. 2010. On the Turkic origin of the names of the Hungarian tribes. In Johanson, L. & Róna-Tas, A. (eds.), *Studies in Turkic Etymology*. Turcologica 85. Wiesbaden: Harrassowitz. 175-187.
- BODROGI, T. 1961. A magyar rokonsági terminológia vizsgálatának néhány kérdése (Műveltség és Hagyomány). *Magyar Néprajzi Lexikon*, III, *Műrokonság*.
- BOTALOV, S. 2013. About Hun European and Hun Asian. In Botalov, S. (ed.), *Hun forum. Problems, origin and identification culture of the Eurasian Huns*. Chelyabinsk: The Publishing center South Ural State University. 32-88.
- CHRONICUM PICTUM, 1987. = *Chronicum pictum Marci de Kalt, Chronica de gestis Hungarorum (Hungarian Chronicle with Pictures, Chronicles by Mark of Kalt, about the Deeds of the Hungarians)*. Translated by Bellus Ibolya. Budapest: Európa Könyvkiadó.
- CZEGLÉDY, K. 1943. *Keleten maradt magyar töredékek*. In Ligeti, L. (ed.), *A magyarság őstörténete*. Budapest: Franklin Társulat. 154-178.
- CZEGLÉDY, K. 1959. A szavárd-kérdés Thury József előtt és után. *Magyar Nyelv*, 5: 373-385.
- CZEGLÉDI, K. 2012. *A magyar nyelv szerepe őstörténetünk kutatásában*, VI. *Törzsneveink*. Pilisvörösvár: Magánkiadás.
- DOMANOVSKY, S. 1933. Kézai és a hunkrónika. In *Károlyi Árpád Emlékkönyv*. Budapest: Sárkány-Nyomda Részvénytársaság. 110-133.
- DOMONKOS, P & KLIMA, L. (eds.) 2002. *Finnugor történeti chrestomathia II*. Budapesti Finnugor Füzetek 18. Budapest: ELTE.
- FERINCZ, I. 2015. *Régmúlt idők elbeszélése. Magyar Őstörténeti Könyvtár*. Budapest: Balassi Kiadó.

- FODOR, I. 1973: *Vázlatok a finnugor őstörténet régészetéből*. Budapest: Régészeti Füzetek, II/15.
- FODOR, I. 2009. Rokonaink-e a szkíták? In Kovács Tibor (ed.), *Szkíta aranykincsek*. Budapest: Magyar Nemzeti Múzeum. 119-12.
- FODOR, I. 2012. *Őstörténeti viták és álviták*, IV. kötet. Budapest: Csodaszarvas. 125-146.
- FÓTHI, E. 1997. *Összehasonlító embertani vizsgálat a Kárpát-medence népeinek ethnogeneziséhez. Embertani kapcsolatok a 6-8. században az eurázsiai steppe és a Kárpát-medence között*. Móra Ferenc Múzeum Évkönyve, *Studia Archaeologica*, 4: 497-521.
- FRANÇOIS, A. 2014. Trees, waves and linkages: Models of language diversification. In C. Bower & B. Evans (eds.), *The Routledge Handbook of Historical Linguistics*. London: Routledge. 161-189.
- GEYBULLAYEV, G.A. 1986. *Toponimiya Azerbajdžana*. Baku: Elm.
- GYÖRFFY, Gy. 1986. *A magyarok elődeiről és a honfoglalásról*. Budapest: Gondolat.
- GYÖRFFY, Gy. 1993. *Krónikáink és magyar őstörténet*. Budapest: Balassi Kiadó.
- HARMATTA, J. 1997. *A magyarok nevei a görög nyelvű forrásokban. Honfoglalás és nyelvészet. A honfoglalásról sok szemmel III*. Edited by Kovács, L. & Veszprémy, L. Budapest: Magyar Tudományos Akadémia.
- HELILOV, M. & NYITRAY, Sz. 2008. *Ősmagyarok Azerbajdzsánban*. Budapest: Hun-Idea.
- HENKEY, Gy. 2002. *A magyarság és más kárpát-medencei népek etnikai embertani vizsgálata*. Turán könyvek II. Budapest: Magyar Őstörténeti Kutató és Kiadó Kft.
- HERBERSTEIN, S. 1549. *Rerum Moscoviticarum Commentarii [Notes on the Moscovites / Notes on Russia]*. Vienna.
- HERODOTUS, 1909 = Herodotus: *The Histories* [translated by G. Rawlinson]. <http://classics.mit.edu/Herodotus/history.4.iv.html>
- HONTI, L. (ed.) 2010. *A nyelvrokonságról Az török, sumer és egyéb áfium ellen való orvosság*. Budapest: Tinta Könyvkiadó.
- HÓMAN, B. 1925. *A magyar hun-hagyomány és hun-monda*. Budapest: Studium.
- HÖLBLING, T., 2009. *A honfoglalás forráskritikája I-II*. Pécs: Magánkiadás.
- HUNFALVY, P. 1876. *Magyarország ethnográfiája*. Budapest: Magyar Tudományos Akademia.
- HYUN JIN, K. 2013. *The Huns, Rome and the Birth of Europe*. Cambridge University Press.
- JORDANES, 2005. *A gótok eredete és tettei*. translated by Kiss Magdolna. Budapest: L'Harmattan Kiadó.
- KÁLMÁN, B. 1998. The history of the Ob-Ugric languages. In D. Sinor (ed.), *The Uralic Languages. Description, History and Foreign Influences*. Handbook of Uralic Studies I. Leiden: E.J. Brill. 394-412.
- KÉZAI, S. = *Kézai Simon mester krónikája*. <http://mek.oszk.hu/02200/02249/02249.htm>
- KIRÁLY, P. 2006. *A honalapítás vitás kérdései*. Nyíregyházi Főiskola: Ukrán és Ruszin Tanszéke.

- KLIMA, L. 1999. *Finnugor Történeti Chrestomathia*, I. Budapesti Finnugor Füzetek, 12. Budapest: ELTE Finnugor Tanszék.
- KLIMA, L. 2002. *Finnugor Történeti Chrestomathia*, II. Budapesti Finnugor Füzetek, 18. Budapest: ELTE Finnugor Tanszék.
- KLIMA, L. 2005. *Finnugor Történeti Chrestomathia*, III. Budapesti Finnugor Füzetek 19. Budapest: ELTE Finnugor Tanszék.
- KMOSKÓ, M. 1997. *Mohamedán írók a steppe népeiről*, I/1. Budapest: Balassi Kiadó.
- KMOSKÓ, M. 2000. *Mohamedán írók a steppe népeiről*, I/2. Budapest: Balassi Kiadó.
- KMOSKÓ, M. 2004. *Szír írók a steppe népeiről*. Budapest: Balassi Kiadó.
- KOHN, S. 1881. *Héber kútforrások és adatok Magyarország történetéhez*. Budapest: Zilahy Sándor Könyvkiadó.
- KÓNYA, S. & KOSÁRY, D. 1975. *A Magyar Tudományos Akadémia másfél évszázada 1825-1975*. Budapest: Akadémiai Kiadó.
- KORYAKOVA, L. & EPIMAKHOV, A.V. 2002. *The Urals and Western Siberia in the Bronze and Iron Ages*. Cambridge World Archaeology. Cambridge University Press.
- KRISTÓ, Gy. 1996. *Magyar honfoglalás – honfoglaló magyarok*. Budapest: Kossuth Kiadó.
- LIGETI, L. 1986. *A magyar nyelv török kapcsolatai a honfoglalás előtt és az Árpád-korban*. Budapest: Akadémiai Kiadó.
- MARCANTONIO, A. 2002 *The Uralic Language Family: Facts, Nyths, and Statistics*. Publications of the Philological Society, 35. Oxford /Boston: Blackwell.
- MARCANTONIO, A. 2014. The linguistic influence of Turkic on Hungarian. *Turkic World Almanac, Turkic Academy of Science*. 1-22. Reprinted in *Journal of Eurasian Studies*, 2015 / II:69-88 (online): www.federatio.org/joes.html
- MARCANTONIO, A. 2017. Magyar-uráli és magyar-török morfológiai megfelelések a tudomány jelenlegi állása szerint. In Ferenczné Szócs, É., Gazda, J. & Szabó, E. (eds.), *Kőrösi Csoma Sándor Gyökereink*. Kovászna: Kőrösi Csoma Sándor Közművelődési Egyesület. 85-102.
- MARQUART, J. 1903. *Osteuropäische und ostasiatische Streifzüge*. Leipzig: Dietrich'sche Verlagsbuchhandlung.
- MUNKÁCSI, B. 1894. Az ugorok legrégebb történeti emlékezete. *Ethnographia* V/1: 160-180.
- MUNKÁCSI, B. 1895. Az „ugor” népnevezet eredete. *Ethnographia* VI / 5-6: 349-387.
- MURAKÖZI, Gy. 2002. *Herodotos: Görög-perzsa háborúk*. Budapest: Osiris Kiadó.
- NAGY, G. 1909. A skythák. Székfoglaló értekezés. *Értekezések a történettudomány köréből*, 22: 1-96.
- NAUMOV, I.V. 2006. *The History of Siberia*. London / New York: Routledge.
- NÉMETH, Gy. 1930. *A honfoglaló magyarság kialakulása*. Budapest: Magyar Tudományos Akadémia.
- NEPERÁCZKI, E.D. 2017. A honfoglalók genetikai származásának és rokonsági viszonyainak vizsgálata archeogenetikai módszerekkel. PhD Dissertation. University of Szeged. <http://doktori.bibl.u-szeged.hu/3794/>.

- ORBÁN, B. 1868. *Székelyföld leírása*. Pest. Electronic version: <http://mek.oszk.hu/04800/04804/html/>
- PAULER, Gy. & SZILÁGYI, S. 1900. *A magyar honfoglalás kútffői*. Budapest: Magyar Tudományos Akadémia,
- PROCOPIUS, C. 1914. *History of Wars*, I-II. Loeb Classical Library 48. [Translated by Henry Bronson Dewing]. Cambridge (Mass.): Harvard University Press.
- PROCOPIUS, C. 1954. *The Gothic Wars, VII-VIII*. [Translated by H.B. Dewing]. Cambridge (Mass.): Harvard University Press.
- REGULY, A. 1844. *Beszámoló az oroszországi útvjáról*. Budapest. Akadémiai Értesítő.
- SCHNETZ, J.M. (ed.) 1990. *Ravennatis Anonymi cosmographia et Guidonis geographica*. Stuttgart: Taubner.
- SZABOLCSI, B. 1934. Népvándorláskori elemek a magyar népzeneben. Adatok a magyar néphagyományok keleti kapcsolataihoz. *Ethnographia*, XLV: 138-156.
- SZÁDECZKY-KARDOSS, L. 2000. *Zichy-expedíció Kaukázus, Közép-Ázsia, 1895*. Budapest: Magyar őstörténeti és Kiadó.
- SZENDREY, Zs. 1923/1924. Történeti népmondáink. *Ethnographia*, 34/35: 143-149.
- SZÜCS, J. 1993. *Az utolsó Árpádok*. Budapest: Történettudományi Intézet.
- TÉLFY, J. 1863. *Magyarok őstörténete. Görögforrások a scythák történetéhez*. Pest: Laufer.
- TOMASCHEK, W. 1889. Kritik der ältesten Nachrichten über den Skythischen Norden II. *Sitzungsberichte d. Phil.-Hist. Classe d. Kais. Akademie d. Wissenschaften*. 117. Wien. 7-16.
- THURY, J. 1897a. A magyaroknak szavarti aszfali neve. *Századok*, 31: 317-403.
- THURY, J. 1897b. Krónikáink és a nemzeti hagyomány. *Irodalomtörténeti Közlemények*, VII/ 3: 275-296.
- THURY, J. 1890. Ósköltészetünk. *Halas*. 1-34.
- TÓTH, T. 1965. A honfoglaló magyarság ethnogenezisének problémája. *Anthropológiai Közlemények*, 9:139-49.
- VÁSÁRY, I. 2008. *Magyar őshazák és magyar őstörténeszek*. Budapest: Balassi Kiadó.
- VÉKONY, G. 1997. A magyar ethnogenezis szakaszai, I-II. *Életünk*, I: 260-276, II: 384-408.
- WHEATCROFT, S.G. (ed.) 2002. *Challenging Traditional Views of Russian History*. New York: Palgrave Macmillan.
- ZSIRAI, M. 1928-30. Finnugor népnevek, I. Jugria. *Nyelvtudományi Közlemények*, XLVII: 252-295, 399-452.

The impossibility of the evolutionary metaphor: Neogrammarians, family trees and linguistic affinity

Péter Pomozi

1. The coming into being of the evolutionary metaphor

There are not many books that have had a greater impact on the development of science than Charles Darwin's *The Origin of Species*, published in 1859. In this publication, Darwin made also some observations on the topic of the development of languages. Actually, at some point, the author expresses his views on the issue of languages 'lineage' in a rather detailed manner (these observations were, originally, almost an entire page long); compare Darwin (1859/1872: 423):

If we possessed a perfect pedigree of mankind, *a genealogical arrangement of the races of man would afford the best classification of the various languages now spoken throughout the world; and if all extinct languages, and all intermediate and slowly changing dialects, were to be included, such an arrangement would be the only possible one.* [...] The various degrees of difference between the languages of the same stock would have to be expressed by groups subordinate to groups; but the proper or even *the only possible arrangement would still be genealogical*; and this would be strictly natural, as it would connect together all languages, extinct and recent, by the closest affinities, *and would give the filiation and origin of each tongue* [italics is mine, P. P.]

Darwin had already expressed some thoughts anticipating his views on language evolution (as stated in his 1859 work), in previous essays. As one can see from the above reported citation, on several occasion, in his *The Origin of Species*, Darwin equates the family tree of the development of species to the language family tree. He did nothing but stating that the origin of every language (be it alive or dead language) can be described, actually, can *only* be described, in the same way as the

genealogy of the *homo sapiens sapiens*. In other words, Darwin considers the evolutionary, biological metaphor as the only, perfect method of describing the lineage, the descent of languages. This was not, however, the first time that linguistics used metaphors drawn from the realm of (live) nature. For example, Humboldt had already equated languages to live organisms, and the so-called ‘philosophical school of linguistics’ of the early 19th century, in general talked about roots as well as the direct and /or lateral descendants of the roots in question, well before the work by Darwin. In Hungary, Count Teleki József (Teleki 1821) adopted the biological metaphors already at the time of the planning of the first, academic, full-scale Hungarian dictionary (*A magyar nyelv szótára*). For example, phrases such as *leány-nyelvek* ‘sister languages’, *törzsökös nyelvek*, literally ‘root-related languages’ – in the sense of *rokon-nyelvek* ‘related languages’ – were already good, popular metaphors. Nowadays too various branches of linguistics successfully adopt the biological metaphors¹, since these clearly show parallelisms between various ecological phenomena and human society phenomena. However, in my opinion, the conception of the Darwin-Haeckel genealogy, coupled with the related concept of the Schleicherian language family tree – having become a founding metaphor, actually, a dogma, of historical linguistics – did not really belong in the least to the class of ‘good metaphors’. Quite the contrary: these concepts, together with the speculative tenet of the primary, exclusive importance of reconstruction, triggered a backward step for the historical comparative linguistics both with regard to the promising, 17th-18th century German and Hungarian starting point linguistics², and the Humboldtian views.

¹ Compare, for example, the concepts of: *biodiversity* and *linguistic (cultural) diversity*, *endangered species* and *endangered languages*, *revitalization* in the field of eco-linguistics, and, partially, of sociolinguistics.

² Already in the year 1769 Sajnovics János made some remarkable observations of comparative nature regarding the Hungarian and Sami (Lapp) language, in his famous book published in Copenhagen: *Demonstratio idioma Ungarorum et Lapponum idem esse*. The same is true of the volume by the (Göttingen) scholar Gyarmathi Sámuel, *Affinitas linguae Hungaricae cum linguis fennicae originis grammaticae demonstrata* (published in Göttingen, in 1799), where the author draws into the comparison also other languages that we now define as Finno-Ugric. In the work by Sajnovics and Gyarmathi, the idea of that single sided, exclusive supremacy of the process of sound change and the (later developed) sound laws, had no place. According to Gyarmathi, a comparative analysis of morphology and syntax also plays an important role – not to count that in the comparative linguistic works of that time the concept of the reconstruction of the protolanguage had not appeared yet, either. The Hamburg scholar Martin(us) Vogeli(us) wrote *De Finnicae Linguae indole observationes* well a

Thus, experiments were done to stick to and sketch the concept of protolanguage (*Ursprache*, later on *Grundsprache*), whilst considering the Darwinian metaphor as the only plausible one in this context. Moreover, the thought gained ground that the current ‘daughter’ languages derive from the mother / protolanguage through numerous phases of development, in the same way as the human being developed from the first, primitive, single-cell living being. The protolanguage was considered as representing the original, primordial *status* of language, and the above cited passage by Darwin must have certainly contributed to this idea. However, this way of thinking is even more problematic than the family tree tenet itself, when it comes to have to account for language contacts and language change. All in all, this metaphor was a dangerous mistake, from many points of view, including from the point of view of current ‘language-in-contact’ studies, language typology, dialectology, etc. Even worse than this – the Historical Comparative Method (HCM) started to apply this metaphor as if, in its consciousness, it were the exclusively correct one (as mentioned).

As to Uralic studies, they too ‘canonized’ the genealogical, language family tree, considering it as reflecting the actual, reliable history of the formation of the modern Uralic languages, and so it worked all the way through up to 1956, when Harri Moora (1956) made several basic, critical remarks regarding the applicability, or otherwise, of the language tree.

2. The major problems of the evolutionary metaphor

The antecedents of the currently known language families have absolutely no direct connection, in any respect, with the earliest systems of the verbal coding proper of the *homo sapiens sapiens*. In other words, the period of formation of the verbal system of the *homo sapiens sapiens* is not to be looked for in the dawn of the formation of the nowadays Eurasian language families, but in periods of time whose antiquity is of a far greater order of magnitude (see Korhonen 1993: 299; Müller 2009: 239). This being the case, the Darwinian equation: ‘species-development’ = ‘language-development’ becomes clearly meaningless for historical linguistics, because, in biological terms, the human language has not

century before Sajnovics and Gyarmathi (between 1668 and 1670); in this work he first explored the possibility to assess a genetic relation between Finnish and Hungarian through (supposedly) scientific methods.

changed at all in the last several thousands of years or so. Only the verbal coding systems that developed like our current ones underwent change and variation, often as a consequence of intricate language contacts, within a (rather) normal multilingual framework. If we were to push the Darwinian-Haeckelian metaphor to its ultimate consequences, then – if we argue in favour of the ‘one-single-cell (in German *Monoren*) starting point’ – we should glance into that period of human language history that occurred well before the ‘out of Africa’ moment. Even so, the evolutionary metaphor would only have some sort of meaning if we accepted the premise of the mono-genesis of the human language, for which, however, there is no clear cut proof.

2.1. ‘Root-less’ trees

Given that the *Grundsprache* was considered, since the very beginning, as *Ursprache*, it is not clear what is there supposed to have been before that. As a matter of fact, if the verbal coding proper of human kind was formed about 40.000-120.000 years ago (as claimed by Korhonen (1993) and Müller (2009)), then the modern Eurasian languages had their own history even before the formation of their respective (assumed) protolanguages – supposing that these protolanguages could actually be reconstructed through the HCM. The thick trunk of the tree is nothing but the visualization the ‘top-point’ convergence (so to say), within the continuous waves of the processes of language convergence and divergence. Dixon (1997: 67) calls this the ‘punctuation period’ – a brief period of time with respect to the ‘status of equilibrium’ (see below). Looking at the issue in this light, a language family tree without roots, in my opinion, is equally a ‘non-sense’, as if we could imagine a gigantic tree without roots but with an alive, widely branching, green foliage³.

³ This is exactly what Zongor Endre did, with regard to the otherwise wonderful family tree he proposed – a tree that resembles a wide-branching, verdant oak-tree. The colourful picture by Zongor (1940) is the Hungarian version of the sketch drawn by the renown Finnish scholar Lauri Kettunen (1938). Kettunen’s and Zongor’s oak-tree show remarkable similarities with the family tree by Ernst Haeckel, proposed in his volume: *Stammbaum der Menschen*, that appeared in 1866. Whilst in Haeckel the *Monoren* and *Amoeben* are at the bottom of the species, in Kettunen-Zongor’s tree the lowest visible part of the tree indicates the Uralic phase. Among the modern family trees the most spectacular example is certainly the combined Uralic-Indo-European tree proposed by Minna Sundberg (for Sundberg’s linguistic tree model see Taggert (2017)).

The reason of the fortunes of the root-less tree is to be looked for within the context of the history of ideas connected to the formation and development of the HCM. On the basis of the *gentem lingua facit* principle, the HCM, for long time, considered as a 'fact' the interdependence, the equation 'history of language' = 'history of peoples'⁴, so that there was no need to raise questions regarding the phases preceding the protolanguage phase itself. In fact, what else could have been there before the 'linguistic amoeba'? Looking at the issue from this perspective, it is easy to understand why any assumption of palaeo-linguistic nature raised fierce opposition on behalf of the representatives of the neogrammarian school. They 'rummaged' among the roots of the family tree, actually, even much deeper than that, with regard to time, so that they *ab ovo* called into question the pervasive authority of the family tree. Thus, paradoxically, the neogrammarian school, that strived to discover the beginning, the origin of the language families, became the real opponent of the search of the protolanguage, dubbing as 'phantasmagoric' anything that lies deeper in time than the visible trunk of the tree – whereas, in fact, the roots of a tree would be nothing but that *status* of language preceding a given 'top-point convergence' (as discussed above).

To conclude: an early Schleicherian version of the tree, strengthened through the Darwin-Haeckel genealogy, became a triumphal march – although it is self evident that no parallelism whatsoever can be drawn between the nature of the biological development of a species and that of a language, because of the linguistic contacts and the

⁴ The HCM, since the last hundred and fifty years or so, has consistently separated the question of genetic relation among peoples from that of the genetic relation among languages, although the ideological origin of the discipline is still (partially) deeply rooted in that romantic climate within which the above mentioned principle – *gentem lingua facit* – remains one of the founding principles. Nowadays we are able to differentiate between the more recent, identity-forming truism that 'a nation lives in its language', and the romantic vision of the coherence and continuity of people and their language, coherence that would be deeply rooted in the past. To start with, the research on the Indo-European protolanguage meant the same as the question 'where did the Indo-European protocommunity come from, and what could have it been like'. It is not by chance that the Darwinian ideology came very handy to linguists operating in the Germanic area: in this context, there was a particular need for the romantic concept of ethno-linguistic identity, for the justification, and verification, of the *Kulturnation*, because, in the absence of a unified German state, there was not even (yet) the theoretical framework for the concept of a *Staatsnation*. If we examine the process of the strengthening of the neogrammarian ideology and its actual transformation into a dogma during the 1860's, in the light of the events that took place in 1871 (the unification of Germany), basically nothing appears to be just by chance.

socio-cultural character of the language itself (within the time frame of the last 40.000-120.000 years). In other words: anything could have happened in the period occurring between the formation of the human language and the formation of the (assumed) protolanguages of the various, modern daughter languages. This being the state of the art, Dixon (1997) and Pusztaý (2011: 29-111) tried to account for it in different ways, the former with the help of the *punctuated equilibrium* model (as mentioned), the latter with the assumption of the existence of a pre-Uralic linguistic phase, on the basis of argumentations of typological nature.

2.2. Linguistic divergence without convergence? And linguistic contacts?

The concept of a root-less family tree is problematic not only because it attributes to the human language, within the time frame of several thousand years only, those changes that, in fact, took place in periods of time of a greater order of magnitude (as discussed), but also because it neglects the existence of the process of convergence (see Künnap 1998: 9). Even accepting the premise that 'all the daughter languages of a given language family are to be traced back to a common mother / protolanguage', one must bear in mind that the proto language itself is nothing but a phase of the top-point of convergence (as already mentioned). It is even more important to realize that, theoretically, a given family tree⁵ is only capable of tracing back about 50% (at most) of the actual, continuous changes that might have taken place within a given linguistic area. In contrast, in its one hundred and fifty years or so of activity, the HCM has operated as if things proceeded in a continuous, mechanical, straightforward line of development, from the simplest status of a language to the most complex one. In other words, only divergence would exist, but not convergence (from this, in turn, the other implicit, tacit HCM assumption follows, according to which the only moment of convergence is the protolanguage phase). The family tree is silent on the fact that convergence plays an equally

⁵ This 50% estimate is, of course, only a 'theoretical optimum'. In reality, the family tree could not even reflect the series of the (possibly related) languages that became extinguished in the past and cannot therefore be known to us, or the intricate mass of convergent and divergent processes generated by complex linguistic contacts.

important role⁶ as divergence, as we know from those phases of languages we can actually ‘observe’ – as a matter of fact, the real changes are in the great majority of the cases the result of intricate, extensive processes of convergence and divergence. This complex interaction of processes, both in space and time, is such that the family tree cannot be able to represent them all in any way, not even through visual means. These issues were raised and properly summarized by the Estonian scholar Moora already in the 1950s, in his work on the ethnic history of the Estonian peoples (*Eesti rahva etnilisest ajaloost*). Moora (1956: 45)⁷ observes that neither languages nor communities and peoples develop along the lines of simple, regular processes of binary divisions and /or simple internal, linguistic/ethnic changes. Quite the contrary, they are influenced by a complexity of factors. He also observes that we forget that the family tree is nothing much than a mere frame, a mere aid to help us to organise the data of branching out languages.

2.3. The binary nature of language trees: far away from linguistic reality

An extraordinarily numerous amount of (more or less deeply) revisited versions of the family tree diagram have been proposed thus far, because of the wide spread awareness of the deep contradictions embedded in this metaphor. As a matter of fact, these contradictions continue to exist, even within the framework of nowadays models of computational phylogenetics. It is true to say that, in choosing the corpora of computational phylogenetics, several fundamental shortcomings embedded in the HCM can be avoided: a) we do not have to deal with a tautological corpus and, more in general, with the process of reconstruction of the corpus itself; b) the data in question are real, historical and /or synchronic data; c) there is no evaluation, but only investigation of the data base. However, there is a fundamental problem within this framework too: the binary

⁶ In this regard it suffices to refer to our modern age types of *lingua franca* to understand and appreciate the powerful weight exercised by linguistic convergence – or even to the painful fact that the linguistic diversity of the world keeps decreasing at a more and more rapid pace, whereby the languages that are loosing their voice are generally replaced by the various big languages of the world (English, Portuguese, Brazilian, Spanish, Russian, etc.), among the switching language peoples.

⁷ Wiik (1998: 19) considered Moora a particularly modern linguist, thanks to his views, as reported above. Wiik also saw in Moora’s views the first signs of modern ‘language-in-contact’ theories – with good reason.

way the computer itself operates, since the 'nodes' obtained through the phylogenetic diagram are, they too, the result of binary branching, binary divisions. Not being a biologist, I believe that the traditional taxonomy of the origin of species is certainly more realistic than cladistic taxonomy. Whatever the case, the founding principle of the binary model, that is, the binary division taking place from a top node, is particularly inadequate for describing social phenomena, and... languages *are* (also) a social phenomenon. The binary model stands far away from what we know about the reality of language formation and development, as shown, for example, in the wide European contact areas, and is equally estranged from convergence *vs* divergence processes – processes that form integral part of the nature of language. Numerous examples, well documented from the history of various languages, clearly demonstrate that a speech-community can really implode under particular, climatic or social pressure, and that numerous, parallel convergent phenomena can indeed concentrate, converge into one single language node – as is the case, more or less, with the formation of the Romance language family. As a matter of fact, the less extreme convergence and divergence processes are very rarely of binary nature. It is at this point worth observing that most division into branching of the traditional (Kettunen-Zongor type of) Uralic tree are also of binary nature (see in detail Salminen 1997); however, not even the technical circumstances could justify such a 'binary way of operating'.

The language, as a social phenomenon, is a fundamentally heterogeneous entity, an entity that hardly ever has anything in common with any phenomenon that is binary, or any sequence of otherwise regular repetitions. This being the state of the art, it would be better to visualize the real nature of language in terms of an endless, intricate network system, such as the subterranean rhizome of spearmint or the masses of *hyphae* ('webs') of mushrooms, known as *mycelium*. According to Connolly (1995: 94), instead of the irrelevant and unproductive "arbo-realism", a "rhizomatic" model (such as the *spearmint* rhizome) would represent a much better scientific metaphor for social phenomena,

A rhizome ceaselessly establishes connections between semiotic chains, organisations of power and circumstances to the arts, sciences, and social struggles

In my opinion, however, the most sensible and realistic metaphor could be the one that represents that mixture of olive oil with basil and balsamic vinegar that remains in our plates after we have eaten

grilled courgettes or aubergine. The dynamics of the division and related re-arrangement of the dots of balsamic vinegar, as well as the dynamics of their adhering to one another – dynamics that occurs, repeatedly, in new places in that stratum of olive oil where we might have first triggered the impulse to change with our fork – is for me the best metaphor to account for the nature as well as the unpredictability of linguistic change. Such a metaphor has the further advantage of not committing itself to any untenable methodological premises, unlike the language family tree, whose absolutism has tied down the methods of historical linguistics.

2.4. Biological species and daughter languages: The totally unfit metaphor

The weakest point of the evolutionary metaphor is the insistence on the unconditional parallel between the origin of species and the origin of language – equation that was practically established by Darwin, as we have seen in his long citation reported above. A biological species is a group of individuals that can breed with one another ('panmixia'); however, a certain group of individuals cannot breed with individuals of other groups, with other species, and these groups are normally reproductively isolated from other groups.

In contrast to biological species, among languages that rigid restriction to the bringing into life of new 'individuals', new 'entities', does not exist. Quite the contrary, there are no restrictions whatsoever, since any language(s) can enter into relation with any other language(s), independently of language family, language type or socio-cultural background. The mutual influence between two or more languages can trigger extensive 'contact phenomena', including multilingualism and the coming into being of new languages. Therefore, if we were to make any direct comparison between language relations and trees *vs* the origin of species (as Darwin advocated) that could prove adequate, then in the same way as language relations and daughter languages, any biological species should be able to come into contact and mate with any other species. In reality, as we know very well, from the encounter of a chimpanzee and a fish, for example, we shall never have a 'chim-fish'. Thus, judging from this simple example only, establishing any sort of equation between the origin of languages and that of species is fundamentally wrong.

To conclude, one can easily state that this Darwinian metaphor is totally inadequate to account both for the genetic relation among languages and their mutual contacts, as well as – one should add – any sociological phenomena. In other words, one can definitively state that, in this respect, this metaphor is ‘dead’.

3. Protolanguage and reconstruction

The adoption of the evolutionary metaphor has had other consequences for the HCM. In fact, the requirement of reconstructing the protolanguage derives (typically, although not necessarily) from the ‘family tree conception’. In turn, this idea of the existence of a single, relatively homogeneous protolanguage, coupled with the absence of historical corpora and real linguistic data, has given a free hand to the practice of reconstructing (at times highly) hypothetical forms, although in the last forty years or so several scholars have expressed their scepticism toward the soundness of this practice. For example, according to Künnap (1999), the reconstruction ‘obeys’ the family tree model also from the formal point of view, that is: the changes among the daughter languages are analysed exclusively in the light of the hypothetical branching, the hypothetical (reconstructed) nodes of the family tree itself, whilst, in reality, the real changes taking place among languages cannot be described without taking into account language contacts. Moreover, according to Künnap (1999: 441- 442) and Sutrop (1999: 645), the concept of reconstruction went too far, since protolanguages have been reconstructed that never actually existed. Künnap (*ibidem*) is also sceptic, a priori, regarding the possibility of having any insight whatsoever of any (past) status of language for which we do not have any (written) records.

We have already referred above to the contradictions embedded in this relatively unsystematic reconstruction pile⁸, observing that they have been regarded as a kind of ‘pre-categorical’ protolanguage, the description of the ‘amoeba status’ of language. It is in this spirit that many renown linguists wrote their work, quite surprisingly. Here we quote

⁸ Décsy (1990) tried to figure out a properly systematic, coherent model of the Uralic protolanguage at any level of language (lexicon, grammar, semantics, etc.), a model with all its difficulties and contradictions. Décsy even reconstructed a Uralic story of a fisherman family. Andrus Saareste (1952) proceeded in an even more arbitrary way: he reconstructed a ‘ghost-story’ for an un-documentable and undocumented period of the Estonian language, in three variants, without being able to support his reconstruction from the phonological, grammatical or semantic point of view.

some interesting examples from Uralic historical linguistic research, that is, the work by Klemm (1927), Ravila (1960), Balázs (1970) and Benkő (1970). In particular, Benkő, already in the 1970s, wrote about the ‘ambivalent nature of classes of words’; compare Benkő (1970: 37):

Indem wir den Gang der Sprachgeschichte zurückfolgen, vermindert sich der Unterschied zwischen Wort und Satz, Wortarten und Satzteilen immer stärker; am Ende ist er gewissermaßen nicht vorhanden...

Klemm’s (1927) and Ravila’s (1960) so-called protosentences too were definitively ‘pre-categorical’, and the supremacy of the noun over the verb, in the protolanguage, was also considered as plausible. In contrast, Hajdú (1966: 94-95) considered the ‘pre-category’ issues definitively as of palaeo-linguistic nature, and this is the way we look at them nowadays too. Actually, many scholars looked at these issues in this same way well before the formation of the Darwin-Haeckel metaphorical conception. Humboldt – who considered the analysis of native languages as of primary importance for the research of language development – stated that they are by no means inferior to the ‘educated’ languages of our times; compare the following quote (Humboldt 1822: 239):

Meine bisherige aber hat mir bewiesen, daß auch die sogenannten rohen und barbarischen Mundarten schon Alles besitzen, was zu einem vollständigen Gebrauche gehört

Even if we did acknowledge that the time frame of formation of the verbal coding of the *homo sapiens sapiens* (the verbal coding of the same nature as our modern languages) is much greater than the time frame of the reconstructed Indo-European *Ursprache*, or the Uralic protolanguage, it does not necessarily follow that the contradictions embedded in the practice of reconstruction of the protolanguage will automatically disappear. In fact, the most serious problem is not at all this chronological ‘somersault’, but the tautological nature of the method of reconstruction itself (as discussed below).

3.1. The tautological nature of the Neogrammarian reconstruction

To properly describe how the practice of reconstruction of the HCM heavily relies on the Neogrammarian phonological principles, it will suffice to quote some decisive thoughts by Schleicher (1861/62: § 243):

die Wissenschaftliche Darstellung der indogermanischen Declination hat die Aufgabe die im vorliegenden Stande der Sprache eng verwachsenen Elemente wieder aufzulösen; mit anderen Worten, *sie muss die älteren und ältesten Formen wieder herstellen, aus denen die Späteren erwachsen sind. Nur so können die Casusformen klar gefasst werden* [italics is added, P.P.]

Thus, according to Schleicher, on the basis of the currently available linguistic material, we should bring into life something out of older stage(s) of a given language, so as to be able, afterwards, to explain its current stage through its older one(s). What is this then, if not tautology? Nevertheless, this tautology, unfortunately, took ground, as shown, for example, by the fact that the otherwise excellent, Estonian dialectologist and linguist Mihkel Weske (1873: 38), literally re-proposes the above reported sentences by Schleicher in his work on historical, Balto-Finnic linguistics. Unfortunately, it is in this spirit that most text-books of Uralic studies have been written in the past, and are still written nowadays. In the sketch reported in the table below, I shall try to show how the systematic tautology proper of reconstructions came into being:

RECORDED LANGUAGE DATA	METHOD	HYPOTHETICAL LANGUAGE FORMS
DAUGHTER LANGUAGES	comparative reconstruction RETROSPECTIVE explaining real language data based on reconstruction FULLY TAUTOLOGICAL	PROTO- LANGUAGE
Grammar and vocabulary: natural with irregularities and grammatical variables		Grammar and vocabulary: totally regular and unnaturally homogenous, without grammatical variables

Fig 1. The main problems and the possible tautological nature of the linguistic reconstruction of protolanguages

In the absence of proper, historical corpora, by taking into account only the real data acquired from the (assumed) daughter-languages, the possibility to establish highly theoretical lexical, grammatical and

semantic reconstructions is on offer. However, these hypothetical protolanguages, protogramemes and protosememes can only play the role of just summarizing the state of the art, at most. In contrast, in the practice of historical Uralic phonology and morphology, statements such as ‘the outcome, the continuation of the protolanguage nasals in the cognate languages’, or ‘the outcome of genitive **-n* in the cognate languages’, etc., are a clear indication of how the HCM still operates nowadays along those tautological conceptions as embedded in Schleicher’s model. All this is quite a ‘non-sense’: whilst a reconstruction derives from real data, drawn from natural languages, the reconstruction itself is by no means a real entity, and it is for this reason that such an entity cannot account, literally, for the real data of real languages. Another major shortcoming is the fact that the reconstructed protolanguage, in contrast to spoken languages, is without exceptions, and too *homogeneous*, so that it is problematic also from a general, typological point of view (see Dixon 1997: 97-98). In other words, the concept of reconstruction is totally unable to account for one of the fundamental properties of natural languages: ‘variation’ – at any level of language (phonological, lexical, morphological, and syntactical level). The homogeneous, ‘normalized’ character of the protolanguage turns out to be even more unnatural, since protolanguages are assumed to go back to those ancient phases of development when there were no rigid state borders, no rigid language borders, and no norms dictated by the written language. Not to count the possibility that the typological nature of languages can change through time, for example from typically isolating to typically agglutinative, and back, in a cyclical manner. In a word, the traditional Uralic reconstruction (like any reconstruction) does not allow the slightest chance of reflecting and describing the variable nature of language.

3.2. Variation and reconstruction? The Lautgesetz dead end

It can be argued that in the past, more than in our 21st century⁹, variation has been a much more ‘visible’ property of languages, and that this

⁹ The drastic decreasing of nowadays linguistic variation across linguistic areas can be attributed to two linguistic and one extra-linguistic factors. The dramatic decrease of the number of languages and dialects spoken in the world is due to the influence of some, dominant languages. Also the local varieties within one single language are diminishing, because of the dominant role of the standard(ized) variants. These

visibility has disappeared from the horizon of linguistic research, because of the Neogrammarian absolutisms in focusing on sound change and its (assumed) regularity. The great majority of Neogrammarian linguists, unintentionally, projected the normative type of analysis of the late 19th century onto the historical phases of the unwritten languages of Eurasia. This was, at least at the time, linguistic ignorance, in the same way as it was ignorance the unconditional adoption of the *Lautgesetz* principle. It is worth observing at this point that the data contained in the extensive *German Dialectal Atlas* (Wenker 1878 & 1881) does not justify at all the pervasive dominance of the *Lautgesetz* conception. Quite the contrary, the dialectal work by Georg Wenker demonstrated¹⁰ that leaving aside phonological variation leads to a dead end approach in accounting for language change and correspondences (see Pajusalu *et. al.* 2002: 20). Nevertheless, instead of acknowledging that variation is a fundamental property of language and, therefore, of historical phonology, any instance of variation has been generally seen as abnormality or exception, fact that, already according to de Saussure, needed to be explained, since, in his opinion, sound changes are generally compulsory and non-conditional (see de Saussure 1967: 181-190). Regarding this rather arbitrary exercise – to speak politely – compare the following observations on Verner’s Law by Marcantonio & Brady (2012: 267), observation in turn inspired by Durkin’s (2009: 182) assessment of sound laws:

To take an example, when examining the evidence for a given law (not just Verner Law), the matches are cited in favour of the law, but the mismatches may be minimised or not counted: they are ‘explained’ to be the result of various factors, such as ‘conditioning environments’, ‘dialectal mixture’, ‘borrowing’, ‘secondary changes’ or ‘analogy’

It is worth observing at this point that the ‘retrospective’ point of view proper of the reconstruction could be achieved also without the Neogrammarian absolutism, if we took into account variation – in which case we would implement much more credible reconstructions.

phenomena are themselves greatly reinforced by an extra-linguistic factor: the ‘digital communication’, that nowadays constitutes an increasingly greater part of the overall human communication, exercising an increasingly more influential role. All these factors, together, contribute to diminish linguistic variation, both within single languages and at a global level.

¹⁰ The publication of the *Sprachatlas* by Wenker is not complete, although between 1876 and 1887 the author collected a corpus of 40.000 linguistic items from the German dialectal areas.

A different sort of question is, of course, whether or not the reconstruction of the protolanguage (in lexicon, grammar, semantics, etc.) is really necessary for the historical, comparative research. As a matter of fact, neither the antiquity nor the location of a reconstructed protolanguage can be really ascertained, so that there would be no much use in a methodological framework that tries to connect, hypothetically, a given protolanguage to a given linguistic area, and its contacts (even if we gave up the traditional tautological approach). Even so, the reconstruction would still fulfil its 'inventory function', inventory through which the results of reliable, historical linguistic research can be mapped and summarized. An example of such a type of positive¹¹, complex reconstruction can be considered the Uralic reconstruction proposed by Décsy (1990).

4. An approach without the evolutionary metaphor: old facts – old-new methods

To conclude this essay, it is worth investigating, briefly, the issue of what kind of research methods have been pushed aside, into the periphery of the received wisdom, by the evolutionary HCM model, the model that has indeed determined and controlled the direction of historical research for the last hundred and fifty years. It is not by chance that nowadays (after the period of dominance of the concepts of 'arborealism' and reconstruction), interest is mounting up again toward various analytical and methodological aspects of those schools of linguistics that have been thus far marginalized.

4.1. Sound law and variation: time and space meet in historical linguistics. A Hungarian phonological example

The dialectological research of the 19th-20th century has clearly demonstrated that a great number of similar phonetic/phonological changes can take place in many languages / language-families of the world, independently from one another. This means that one can call into question not only the tenet of the exception-less nature of sound

¹¹ In my opinion, the reconstruction is 'positive', because it distances itself from the tautological approach. However, an assessment of Décsy's reconstruction is outside the scope of this essay.

change, but also the tenet of the ‘specificity’ of language families (also with regard to their sound changes). On the basis of these dialectological studies it has also become self evident that phonological variables *can* and *do* coexist both in space and time – contrary to the Neogrammarian predicament (as specifically formulated by August Leskien (see Graefen & Liedke 2012) that ‘die Lautgesetze kennen keine Ausnahmen’. It is in this spirit that, for example, Horváth (2011) and Pomozi (2013: 159–170) illustrate the existence of phonological variations associated with what are typically considered to be instances of rigid Uralic (*vs* Hungarian) sound changes and correspondences. One could also re-formulate the issue as follows: phonological variation has its own time (historical) and space (dialectal) dimension, its own history. In other words, what we name, traditionally, as ‘sound change’, is, in reality, variation through time, historical variation; however, even this formulation can be quite misleading. As a matter of fact, unlike what assumed by the *Lautgesetz* tenet, the essential point here is that the changes in most cases cannot be reliably dated, so that the variation can be analysed not only from a static, synchronic point of view, but also dynamically – in a more limited way, however, because the variation that can be documented at the level of time and space could also have been implemented at the sociolinguistic level.

The issue of variation has been approached in an interesting and prolific way by the so-called ‘philosophical school of linguistics’ of the 19th century, as indicated in the following quote by Kövecses (2015: 3):

Two nineteenth-century Hungarian lexicographers, Gergely Czuczor and János Fogarasi, were instrumental in directing our attention to the importance of word roots in the organization of the lexicon of the Hungarian language. As they showed in their work, roots, provide the semantic basis for the majority of Hungarian words. They suggested that roughly 2,300 roots supply the core semantic element of over 100,000 Hungarian words

If, in a language, a corpus of at least 150.000 lexical items can be built out of 2.300 roots only, then, by logical necessity, in this language a significant portion of the possible (morpho-)phonological variables must be used. Within the framework of the investigation of the roots (or, more appropriately, stems), the mapping of the lateral genetic relations forms an integral part of the internal reconstruction. The Czuczor–Fogarasi (CzF) dictionary was, and still is, able to show the full range of phonological

variation existing in the Hungarian language¹². Through this method, it is also possible to enter into the sphere of internal reconstruction, in which case, for the etymological research of the related languages, the investigation can provide those extra perspectives of analysis that the neogrammarians were so eager to reject. Perhaps, we are dealing here with a branch of comparative studies that the conventional HCM (considering itself, seemingly, as the only legitimate representative of historical linguistics), rejected on methodological grounds¹³. As a consequence, the ‘sound laws’ conception, and related method of analysis, became dominant over what appears to me as being a more objective method: the one that takes (also) variation into account. Because of restriction of space it is not possible here to enter into the details of this relevant issue – relevant from the point of view of sound development, etymology and semantics (see some interesting examples in Pomozi (2013)). Nevertheless, let us take just one ‘crystal clear’ example of sound change that is widely considered to be without exceptions in the Uralic literature (including the TESz and SSA dictionary): the change $*p > f$, that would have taken place in Hungarian. This change is still considered to be valid and exception-less within Uralic historical linguistics, at the point that eventual ‘survival’ forms that could also be taken into consideration in this context – such as the correspondence $p \sim p$ or $p \sim b$ – are seen with great scepticism. In other words, only the data of example (1) – U $*p >$ Hungarian f – are considered to be comparable:

1) U $*p > f$:

Hungarian *fazék* ‘pot’ ~ Mansi *pōt*, *put* ~ Hanti *put* ~ Mari *pot* ~ Estonian *pada* ~ Finnish *pata* ~ Sami *batte-* (MszFE: 184); compare also the following, similar sets of correspondence, quoted from Hungarian only: *fagy* ‘frost’, to freeze’; *fél* ‘to be afraid’; *fej* ‘head’; *falu* ‘village’; *fan* ~ *fon* ‘to spin’; *fenyő* ‘pine (tree)’; *fészek* ‘nest’; *fiú* ~ *fi* ‘child, boy’.

However, there are numerous Hungarian etymologies for which the neogrammarian law $*p > f$ is anachronistic. If we looked at the ac-

¹² In the terminology of the 19th century, obviously, the term ‘variation’, as technical term, was not yet known.

¹³ At this point it is interesting to observe that the philosophical school of linguistics was not an opponent to the neogrammarians, neither it could have been, because it preceded them of about a century. It was the HCM that positioned itself as a fierce opponent to earlier linguistics concepts, branding them as speculative and non scientific.

tual data, it would be more appropriate to describe and establish the sound change / correspondences as follows:

other Uralic languages **p* vs Hungarian *f* ~ *p*

Let us have a look at the examples in (2), the case of *pulya*, and its (assumed) cognates as listed in (3):

2) *p* ~ *p*: *pulya* ‘boy’ ~ *pulyka* ‘turkey’:

According to TESz, *pulya* and *pulyka* are of unsure, or at least debated origin, and do not even belong together – although, according to the first proposed explanation (TESz 3: 308–309), both *pulya* ‘turkey, chick, young (of birds)’ and *pulyka* could actually belong together, *puly-ka* being the variant with *-ka*, a diminutive suffix referring to animals. However, if we consider the meaning of Hungarian *pulya* ‘young (of bird)’, and the Estonian and Finnish dialectal variants *puja* ~ *poj* (whereby *-ka* is indeed a diminutive suffix), we get the impression that the origin of the two words do deserve a new investigation, both from the Uralic and the dialectal / ‘linguistic geography’ point of view. The number of cognate words is high, and their semantic correspondence is also impeccable, the meaning being, consistently ‘chick, young (of bird)’ across all cognates:

3) cognates of *pulya* ~ *pulyka*:

Finnish *poika*, dialectal *poju* ‘boy, young (of birds)’; Veps *poeg* ‘somebody’s child’, young (of birds)’; Estonian *poeg*, dialectal *poja* ‘somebody’s child, young (of birds)’; Erzya (/Mordva ~ Mordvin) *bujo*, *pijo* ‘grand child’; Mari (/Cheremis) *pi-erre* ‘young man, lad’; Komi (/Zyrian) and Udmurt *pi* ‘(young) man, child’; Mansi (/Vogul) *püw* ~ *pir* ‘child, young (of birds)’; Eastern Hanti (/Ostyak) *pär* ‘young, child’, etc.

In my opinion, even more striking are the data relative to the cognates listed in (4) (see SSA 2: 455 for the Ob-Ugric data):

4) some more cognates:

Hungarian *forog* ~ *pörög* ‘to turn, revolve’; Finnish *pyörri-* ‘idem’, *pyörä* ‘wheel’; Estonian *pööra-* ‘idem’; Erzya Mordvin *нурдамс* ‘to turn (round) [intransitive]’; (South) Mansi *püwärt-* ‘turn’; (Eastern) Hanti *pəŋərət* ‘idem’

The current, authoritative etymological dictionaries do not even mention the possibility of a common origin for Hungarian *pörög* and (at least one of) the other, quoted forms in (4) – actually, TESz (1: 951–952) and SSA (2: 455) mention the root *for-*, but reject the etymological

connections suggested here. In contrast, Horváth (2011), in his analysis of the *pereg* ('to spin') word family, has a different opinion. The correspondence between Hungarian *forog* and the Erzya Mordvin form here quoted is impeccable, at any level. Thus, Hungarian *forog* can be etymologically connected with the Erzya cognate, a 'lateral' relationship, whilst *pörög* can be connected with the Balto-Finnic and Ob-Ugric forms. This, of course, applies if also on the front of the *p vs f* correspondence we can rely on what can be called a 'survival phenomenon', or, better, old, dialectal phonological variation, that has been preserved until now in the Hungarian lexicon. With regard to the examples in (4), it can be said that, overstating the role of the neogrammarian sound laws, as well as the maniac application of the evolutionary model, shuts the door to the acceptance of alternative analysis, because 'everything must develop into something', 'everything must change'.

4.2. A fruitful, historical-comparative linguistics approach: 'Folklore linguistics'

According to Sutrop (1999:649), the investigation of (old-style) folk songs lexicon (*rahvalaulude uurimine*), is the only pivotal field of research that makes it possible to reconstruct (the data of) a language without falling into the whirlpool of the family tree – a point of view that I share completely. As a matter of fact, not only the old style folk-songs, but also the old style folk-ballads typical of children games – thanks to their oral tradition, their genre and prosodic features – contain those old 'survival elements' that can be found also in the oldest written sources. Another promising area of research within the framework of the Hungarian folklore are the so-called apocryphal folk prayers. The archaic folklore genre represents a collective memory, and the extensive linguistic time codified, embedded in them certainly deserves attention (see Pomozi & Karácsony 2016: 30–31). For example, in the modern Hungarian language the grammatical category of 'evidentiality' is totally estranged, whilst in the eastern dialectal areas, in spontaneous discourse, one can still find traces of the 'two-choices ('witnessed' *vs* 'un-witnessed') evidential system', up to the 17th century (Pomozi 2014: 85–100). This very double choice evidentiality, grammatical system is reflected in some variants of the above mentioned apocryphal popular prayers, such as the Friday Prayer as recited by a *csángó-magyar* (the Hungarian csango dialect)

woman, Pezsán Jánosné – prayer written down in 1971 (Erdélyi 1976: 395). The detailed analysis of these traces of folkloric language and culture often can come very handy for the research of more remote historical and typological interconnections, because of the ‘real language’ data they reveal.

5. Conclusion

The evolutionary linguistic metaphor – according to which all the living languages develop from an arbitrarily established starting point toward some (unspecified) end point, through continuous, mechanical and systematic divergences – set itself a task whose achievement, within the given theoretical and methodological framework, was, *is* impossible *ab ovo*. In particular, appears impossible the task of describing the protolanguage, on the basis of the reconstructed protolanguage itself, and then sketching the process of formation of the (now living) languages of the language family in question, through intermediary, hypothetical, secondary protolanguages / nodes.

I personally believe that, in the case of linguistic investigation involving pre-historical times, the times before the availability of sufficiently old written documents (as within Uralic), relevant answers cannot be provided to the relevant questions exclusively by linguistic analysis. On the contrary, we can only progress if we adopt a holistic, interdisciplinary, cooperative approach, that includes also the results of disciplines such as population genetics and archaeology. In other words, it would be desirable to get rid of those aspects of the conventional method of analysis that have been harmful or useless in the last 150 years or so of HCM – and still are – and to look instead for interdisciplinary methods capable of promoting interaction and synergy.

References

- BALÁZS, J. 1970. Zur Ausbildung der ältesten uralischen syntagmatischen Verbindungen. In W. Schlachter (ed.), *Symposion über Syntax der Uralischen Sprachen*. Göttingen: Vandenhoeck & Ruprecht. 23-35.
- BENKŐ, L. 1970. Zur Frage der uralischen Nomenverba im Lichte der ungarischen Sprachgeschichte. In W. Schlachter (ed.), *Symposion über Syntax der Uralischen Sprachen*. Göttingen: Vandenhoeck & Ruprecht. 36-49.

- CONNOLLY, W.E. 1995. *The ethos of pluralization*. Minneapolis: University of Minnesota Press.
- CzF = *A magyar nyelv szótára I-VI (1862–1874)*. Czuczor G. & Fogarasi J. (eds.). Pest: Magyar Tudományos Akadémia.
- DARWIN, CH. 1859. *On the Origin of Species*. London: John Murray.
- DÉCSY, GY. 1990. *The Uralic Protolanguage. A Comprehensive Reconstruction*. Bloomington (IN): Eurolingua.
- DIXON, R. M. W. 1997. *The rise and fall of languages*. Cambridge University Press.
- DONNER, O. 1874. *Vergleichendes Wörterbuch der Finnisch-Ugrischen Sprachen*. Helsingfors: Druckerei der Finnischen Literaturgesellschaft.
- DURKIN, PH. 2009. *The Oxford Guide to Etymology*. Oxford University Press.
- FOX, A. 1995. *Linguistic Reconstruction: an Introduction to the Theory and Method*. Oxford University Press.
- GRAEFEN, G. & LIEDKE, M. 2012. *Germanistische Sprachwissenschaft. Deutsch als Erst-, Zweit-oder Fremdsprache*. 2. erw. Aufl. Tübingen-Basel: A. Francke Verlag.
- HAECKEL, E. H. 1874. *Anthropogenie oder Entwicklungsgeschichte des Menschen. Gemeinverständliche wissenschaftliche Vorträge über die Grundzüge der menschlichen Keimes- und Stammes-Geschichte*. Mit 12 Tafeln, 210 Holzschnitten und 36 genetischen Tabellen, 2., unveränd. Aufl. Leipzig: Engelmann.
- HAJDÚ, P. 1966. *Bevezetés az uráli nyelvtudományba*. Budapest: Tankönyvkiadó.
- HORVÁTH, K. 2011. Szócsaládok etimológiai kutatása a szinkronia-diakronia egységében: a pereg eredete es rokonsága. In E. Kádár & N. S. Szilágyi (eds.), *Szinkronikus nyelvtudomány és diakronia*. Kolozsvár: EME. 128-147.
- HUMBOLDT, VON, W. 1822. *Über das vergleichende Sprachstudium in Beziehung auf die verschiedenen Epochen der Sprachentwicklung*. Abhandlungen der Königlich Akademie der Wissenschaften zu Berlin, aus den Jahren 1820-1821. 239-260.
- KORHONEN, M. 1993. *Kielen syntty*. Porvoo: Werner Söderström OY.
- KÖVECSES Z. 2015. The Hungarian root *es-* in language and cognition. *Language and Cognition*. 1-35.
- KÜNNAP, A. 1998. *Breakthrough in present-day uralistics*. University of Tartu.
- KÜNNAP, A. 1999. *Mea culpa, aga 1500, eelsest eesti keelest ei tea me midagi. Keel ja kirjandus*, 52: 441-444.
- MARCANTONIO, A. 1996. La 'questione' uralica. *Rivista di studi Ungheresi*. 11: 5-44.
- MARCANTONIO, A. 2002. *The Uralic Language family. Facts, Myths and Statistics*. Oxford /Boston: Blackwell.
- MARCANTONIO, A. & Brady, R. 2012. The Evidence to Support Verner's Law. *Vedic Venues*. 2012/1. 262-306.
- MOORA, H. 1956. *Eesti rahva etnilisest ajaloost*. Tallinn: Eesti Riiklik Kirjastus.
- MSZFE = *A magyar szókészlet finnugor elemei I–III (1967, 1971, 1978)*. [main editor: L. György]. Budapest: Akadémiai Kiadó.
- MÜLLER, H. M. 2009. Sprachevolution. Die Entstehung der Sprache. In Müller (ed.), *Arbeitsbuch Linguistik*. Paderborn/München/Wien/Zürich: Ferdinand Schöningh.

- PAJUSALU, K. et al. 2002. *Eesti murded ja kohanimed*. Tallinn: Eesti Keele Sihtasutus. POMOZI P. 2013. Szótörténeti elvek a magyar nyelv szótárában – A belső rekonstrukció és nyelvhasználat viszonyához. In K. Horváth (ed.), *II. Czuczor–Fogarasi konferencia*. Budapest: Magyar Művészeti Akadémia. 139-182.
- POMOZI, P. (ed.) 2016. *Ott látok egy aranyágot... Archaikus népi imáink világa*. Sátoraljaújhely-Széphalom: A Magyar Nyelv Múzeumaért Alapítvány.
- PUSZTAY, J. 1995. *Diskussionsbeiträge zur Grundsprachenforschung (Beispiel: das protouralische)*. (VSUA 43) Wiesbaden: Harrasowitz Verlag.
- PUSZTAY, J. 2011. *Gyökereink*. Budapest: Nap Kiadó.
- SAARESTE, A. 1952. *Kaunis emakeel*. Lund: Eesti kirjanike Kooperatiiv.
- SALMINEN, T. 1997. Facts and myths about Uralic studies. *Sprachtypologie und Universalienforschung* 50: 85-95.
- SAUSSURE DE, F. 1967 (/1915). *Bevezetés az általános nyelvészetbe*. Budapest: Gondolat.
- SCHLEICHER, A. 1861/62. *Compendium der vergleichenden Grammatik der indogermanischen Sprachen*. (Kurzer Abriss der indogermanischen Ursprache, des Altindischen, Altiranischen, Altgriechischen, Altitalischen, Altkeltischen, Altslawischen, Litauischen und Altdeutschen). Weimar: H. Boehlau.
- SSA = *Suomen sanojen alkuperä* I-III (1992, 1995, 2000). [editors: E. Itkonen & U.-M. Kulonen]. Helsinki: Kotimaisten kielten tutkimuskeskus ja Suomalaisen Kirjallisuuden Seura.
- SUTROP, U. 1999. Eesti keele vanimatest perioodidest. *Keel ja Kirjandus* 52: 645-649.
- TAGGART, E. 2017. Comic Artist Maps the History of Languages with an Illustrated Linguistic Tree.
<https://mymodernmet.com/comic-artist-illustrated-linguistic-tree/>.
- TELEKI, J. 1821. *Egy tökéletes magyar szótár' Elrendelése, készítése módja. Jutalomfeleletek a magyar nyelvről a Magyar Nemzeti Museum 1815. 1816. 1817. esztendei kérdéseire*. Pest: Trattner János Tamás.
- TESZ = *A magyar nyelv történeti-etimológiai szótára* I-III. 1967, 1970, 1976 [main editor: L. Benkő]. Budapest: Akadémiai Kiadó.
- WENKER, G. 1878. *Sprach-Atlas der Rheinprovinz nördlich der Mosel sowie des Kreises Siegen*. Nach systematisch aus ca. 1500 Orten gesammeltem Material zusammengestellt, entworfen und gezeichnet von Dr. Georg Wenker. Marburg 1878.
- WENKER, G. 1881. *Sprach-Atlas von Nord- und Mitteldeutschland*. Auf Grund von systematisch mit Hilfe der Volksschullehrer gesammeltem Material aus circa 30.000 Orten bearbeitet, entworfen und gezeichnet von Georg Wenker. Abth. I, Lief. 1. Straßburg/London 1881.
- VERNER, K. 1877. Eine ausnahme der ersten lautverschiebung. *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der Indogermanischen Sprachen*, 23: 97-130.
- WESKE, M. 1873: *Untersuchungen zur vergleichenden Grammatik des finnischen Sprachstammes*. Leipzig: Breitkopf und Härtel.
- WIIK, K. 1998. *Lõunaeesti sünniteooriatest. Lõunaeesti keelest ja kirjakeelest*. Võro Instituudi Toimõtisõq 3. Võro Instituut: Tarto-Võro.

Differential Object marking in Eastern Mari and Permic: A look from the field*

Natalia Serdobolskaya

1. Introduction

One of the characteristic features of the Uralic languages is the phenomenon of differential object marking (DOM); see Bossong (1985), and the recent summary in Witzlack-Makarevich & Seržant (2017). For Uralic this phenomenon has been described in Wickman (1955); Collinder (1957); Bese, Dezső & Gulya (1970); Korhonen (1996).

Eastern Mari

(1)	Kütü	/	kütü-m	čumər-aš	məj
	herd	/	herd-ACC	gather-INF	I
	pij-lan		šüşkalt-əš-əm.		
	dog-DAT		whistle-NARR-1SG		

In order to gather the herd I whistled to the dog.

In Eastern Mari and in Permic languages, DOM displays the following tendencies¹:

- animate direct objects (DOs) are more likely to be marked with the accusative, while inanimate DOs are more likely to be unmarked;

* The work is supported by the RFH grant № 16-24-17003.

¹ It has been shown that in Balto-Finnic languages the distribution of the markers of the direct object is mostly verb-oriented, i.e. based on verbal aspect, tense, mood, or polarity (see Rätsep 1957, Tamm 2004; Heinämäki 1984, 1994). Some reference grammars of Udmurt also consider verbal categories among the basic factors that regulate DOM in these languages. However, in languages and dialects under discussion the verbal categories seem to play a minor role and are only relevant in some specific constructions and contexts.

- pronouns and proper names are more likely to be marked than common nouns;
- definite DOs are more likely to be marked than indefinite and non-specific ones;
- topicality is relevant, e.g. generic and universal DOs are preferably marked when topical, and unmarked if not.

Hence, at first sight it seems that the system of DOM in Uralic is organized along the same principles as discussed in de Hoop (2005) and de Swart (2007), and can be described in terms of typological scales including animacy scale, definiteness scale and topicality scale. Thus, it seems that the system of DOM is rather similar in the languages under discussion, and DOM can be synchronically analyzed as the pool of their common features (which can be ascribed both to diachronic factors and constant areal influence). However, I claim that the constructions of DOM differ among particular languages to a large extent, both on the morphological, syntactic, prosodic and semantic level. Moreover, DOM may work differently even in the dialects of one and the same language. Therefore, it must be admitted that the specific models of DOM are subject to fast changing and their layout cannot be easily predicted on the basis of their common origin.

I am going to present the data from fieldwork in Permic languages, Pechora (village of Eremeevo) and Izhma dialects (village of Muzhi), of Komi-Zyrian, Beserman dialect of Udmurt (villages of Shamardan and Vorca), and Eastern Mari (village of Staryi Tor'yal). Part of the data is taken from spoken language corpora compiled by our fieldwork teams, see http://beserman.ru/corpus/search/?interface_language=en and <http://web-corpora.net/KomiTexts/>. Such examples are marked with the tag (Corpus). The elicited examples are not marked. The information on Izhma is taken from Kashkin (2008), and Biryuk et al. (2010).

It must be specified that the present study is only focused on nouns in DO position. The pronouns in all the examined idioms show different constraints: for example, in Mari they only occur in the accusative, in Komi-Zyrian some of them can take the cumulative marker of accusative and possessivity, and some of them only take the accusative etc. Most of these constraints are lexically-based. As for adjectives and other modifiers that occur as a head of a noun phrase (NP), in all the idioms considered in this paper (except for Beserman) they obligatorily take case/possessive markers in case of the absence of a nominal head

(Simonenko & Leont'ev 2012). Thus, the differential object marking of non-nominals is a specific topic which must be dealt with separately.

The paper starts with the survey of DOM in Eastern Mari; then I pass to Permic languages: first I analyze DOM in Udmurt, and then I examine two dialects of Komi-Zyrian.

2. Eastern Mari

2.1. Morpho-syntax of DOM in Eastern Mari

In Mari the DO can either take a marker of the accusative $-(ə)m$ or be unmarked (1). In both cases it can take possessive markers (see Tužarov (1986) for examples with possessive suffixes of 1st and 2nd person), and nominal number morphemes (it must be noted that the possessive suffixes are very rare in case of the absence of the accusative marker; however, they are not ungrammatical in this case):

- | | | | | |
|-----|--------------|------------|---------------------|-----|
| (2) | Peʻa-n | kid-še | muš-m-əž-əm | məj |
| | Peter-GEN | hand-P.3SG | wash-NMLZ-P.3SG-ACC | I |
| | už-əm. | | | |
| | see-NARR.1SG | | | |

I saw Peter wash his hands.

- | | | |
|-----|---------------------|----------|
| (3) | tuvər- v lak | nal-aš |
| | shirt-PL | take-INF |

to buy shirts

Omission of the DO marker in Mari is restricted to dependent non-finite clauses² (Galkin 1964: 46, 85; Tuzharov 1998: 122), cf.:

- | | | | | |
|--------|---------------|-------------|-----|------------------|
| (4) a. | Č'odəra-šte | kaj-že | da | poŋg-əm / |
| | forest-INESS | go-OPT | and | mushroom-ACC |
| | *poŋgo | pog-əžo. | | |
| | mushroom | collect-OPT | | |

Let him go to the forest and collect mushrooms.

² Tužarov (1986) gives exceptions to this rule from Standard Eastern Mari. However, he shows that they are only limited to DOs with possessive suffixes of 1st and 2nd person.

b.	Məj	č'odəra-š	poŋgo	pog-aš
	I	forest-LAT	mushroom	collect-INF
	kaj-em.			
	go-PRS.1SG			

I am going to the forest to collect mushrooms.

Unmarked DOs must be adjacent to the verb and are prosodically united with it, which reminds of compounding, incorporation or pseudo-incorporation; see Serdobolskaya (2015) for details. However, Serdobolskaya (2015) shows that the unmarked DO does not form a compound with the verb (the unmarked DO + V complex does not yield to the vowel harmony rule observed in Mari; unmarked DOs can be coordinated with a conjunction; furthermore, morphologically the unmarked DOs are not treated differently from the accusative DOs, as both take the whole range of nominal grammatical markers).

The unmarked DO can only take a limited set of modifiers (adjectives, juxtaposed nouns, quantifiers, numerals). Restrictive relative clauses, universal quantifiers, demonstratives and indefinite pronouns are unacceptable with the unmarked DOs.

2.2. Semantics of the unmarked DOs

Most reference grammars argue that definiteness is most important for the choice of DO encoding in Mari (Galkin 1964, Anduganov 1991); see the critique in Tužarov (1984, 1986). Some specialists informally describe the semantic opposition of the accusative vs. caseless DO in terms of “singling out the DO” with the accusative vs. “singling out the action itself” without the accusative (Pengitov 1961: 67). Toldova and Serdobolskaya (2002) show that it is the information structure of the sentence that is crucial for the choice of DO encoding in Mari.

Indeed, indefinite/non-referential interpretation of unmarked DOs is the most frequent. However, they can also be specific and even definite:

(5)	Tide	istorij-že	takšəm	ške	odnoklassnik-šaməč'
	this	story-P.3SG	SO	REFL.GEN	classmate-PL
	kuze	vašlij-me	nergen	da.	
	how	meet-NMLZ	about	then	

This story is about a meeting of my classmates. {Two more sentences.}

Vot...	pogən-en-na		əle	kutər-en	kelš-en-na
well...	get.together-PST-1PL		be.3SG	talk-CONV	agree-PST-1PL
odnoklassnik-šaməč'		vašlij-aš.			
classmate-PL		meet-INF			

Well... we got together and arranged to meet our classmates.

The unmarked DO in (5) refers to an aforementioned discourse referent (the example contains two sentences taken from a narrative; the second sentence contains a DO referring to an antecedent which is in the first sentence).

Unique objects like 'sun' and 'moon', which are definite by their definition, can also be unmarked in DO position:

(6)	Keč'-əm /	keč'e		onč'-aš	jörat-əše
	sun-ACC	sun		look-INF	love-PTCP.ACT
	jen-vlak	er	kən'el-ət.		
	person-PL	early	get.up-PRS.3PL		

People who like to meet the sun get up early.

Some native speakers marginally allow proper names to be unmarked in DO position.

On the contrary, the accusative-marked DO can be non-specific:

(7)	Urem-əšte	pərəs-əm	už-aš	saj-lan	ogəl.
	street-INESS	cat-ACC	see-INF	good-DAT	NEG.PRS.3SG

It is a bad sign to see a cat / cats in the street.

As shown in Toldova & Serdobolskaya (2002), it is the information structure of the sentence that is crucial for the choice of DO encoding. Consider the following examples. In (8a) and (8b) the sentence appears in such a context where the DO and the verb have different information structure functions. The DO is focused in (8a), while the verb constitutes the background part. By contrast, in (8b) the verb is focused, while the DO is the topic. In both examples the accusative marker on the DO is obligatory.

- (8) a. – Tide materjal gəč' mo-m urg-aš
 this cloth out.of what-ACC sew-INF
 lij-eš? – **Tuvər-əm** / ***tuvər** urg-aš
 be.possible-PRS.3SG shirt-ACC shirt sew-INF
 lij-eš.
 be.possible-PRS.3SG

What can I make of this cloth? – You can make a shirt. (Toldova & Serdobol-skaya 2002: 117)

- b. {*What are you going to do with the underclothes?*}
 – **Vurgem-əm** / ***vurgem** šakal-aš
 underclothes-ACC underclothes hang-INF
 kül-eš.
 need-PRS.3SG

I have got to hang the underclothes.

However, if the whole verb-phrase (VP) with the DO and the verb is focused, the accusative marker is omitted (9). The same holds for cases when the whole VP is topical (10).

- (9) – Pe'a-lan [mo-m əšt-aš] kül-eš?
 Peter-DAT what-ACC do-INF need-PRS.3SG
 – Pe'a-lan [**pareŋge** erəkt-aš] kül-eš.
 Peter-DAT potato peel-INF need-PRS.3SG

What should Peter do? Peter should peel the potatoes. (Toldova & Serdobol-skaya 2002: 115)

- (10) Təj mo ergə-č'-ən **araka** jü-m-əž-lan
 you Q SON-P.2SG-GEN vodka drink-NMLZ-P.3SG-DAT
 kuan-et mo?
 rejoice-PRS.2SG Q

*(The speaker and the hearer are discussing the fact that the hearer's son drinks.)
 Do you approve of your son's drinking vodka?*

Hence, it is relevant whether the DO and the verb form a single information structure unit. The DO is unmarked if both the verb and the DO bear the same status in the information structure of the sentence, thus forming a single unit, a topical or wide-focused VP. The DO takes the accusative marker if the verb or the DO bears narrow (or contrastive) focus, or if the DO / V is topicalized.

Thus, DOM in Eastern Mari includes a pseudo-incorporation construction with a very specific information structure type (the DO and the verb form a single unit). Such a pattern is not observed either in Permic (as shown below), or in Mordvin languages (see Serdobolskaya & Toldova 2012 on the Shoksha dialect of Erzya-Mordvin and Toldova 2017 on Moksha-Mordvin).

3. Permic languages

In Permic languages, DO markers diachronically arise from the possessive suffixes, which are at different stages of grammaticalization (Maitinskaya 1979: 102; Rédei 1988: 382–383). Modern languages have a dedicated accusative marker, *-ez* in Udmurt (*-tə* in plural) and *-əs/-es* in Komi, and possessive markers that inflect for person and number. The possessive markers in Permic are often used beyond the domain of possession, signaling the definiteness/topicality/animacy of DOs; see Schlachter (1960), Fraurud (2001), Kuznetsova (2003), Suihkonen (2005), Winkler (2011). Let us consider two examples from Komi-Zyrian (Pechora dialect):

- | | | | |
|------|-------|----------------|--------------------|
| (11) | Pet'a | rad'ejt-ə | č'oj-sə. |
| | Peter | love-PRS.3(SG) | sister-ACC.P.3(SG) |

Peter loves his sister.

- | | | |
|------|--------------|-------------------|
| (12) | Boš't | n'an'-sə! |
| | take(IMP.SG) | bread-ACC.P.3(SG) |

Take this slice of bread! (The speaker is showing the slice.)

The example (11) illustrates that the possessive marker can be used to indicate the possession relation (namely, the relation of kin), while in (12) the same marker has deictic interpretation, “the slice that I am showing you”. Its actual possessor is 1st person (“I” or “we”), since the utterance is pronounced by a grandmother and addressed to her grandchild. In both Udmurt and Komi-Zyrian this interpretation is observed with the possessive markers of 2^d and 3^d person, whereas the markers of 2^d person usually develop very special discourse functions (for example, the ethical function, see below).

Hence, in Permic languages DOM shows the following ternary opposition: unmarked DOs vs. accusative DOs vs. possessive DOs (the possessive DOs are cumulative markers of case and possession, as shown below).

3.1. Permianic: Udmurt (Beserman)

3.1.1. Morpho-syntax of DOM in Beserman Udmurt

Let us consider the three DO marking strategies in Beserman Udmurt: an unmarked DO (13), a DO with the accusative marker (14) and a DO with the possessive (15):

(13)	Odig magaž'in-e	pār-i-m,	sur	baš'-t-ānā
	One shop-ILL	enter-PST-1PL	beer	take-SMLF-INF
	med-iš'ko-m	val.		
	want-PRS-1PL	be.PST		

We entered a shop, we were going to buy some beer. (Corpus)

(14)	So	ber-e	gine=n'i	so	kāl't'o-os-ez
	that	back-ILL	only=already	that	sheaf-PL-ACC
	kel'a-l'a-z-ā	zavod-jos-ā			
	send-ITER-3-PL(PST)	plant-PL-ILL			

{The women take the rubbish out of the flax sheaves.} Then they brought the sheaves to flax plants. (Corpus)

(15)	Sāre	š'i-i-z-ā	val-ze,
	then	eat-PST-3-PL	horse-ACC.P.3(SG)
	škura-ze	kel'-t-i-z-ā	
	skin-ACC.P.3(SG)	be.left-CAUS-PST-3-PL	

*Then they ate **his** horse, (and) left **its** skin. (Corpus)*

The first example (13) contains a non-specific DO, which is unmarked. In (14) the aforementioned DO takes the accusative marker *-ez*. The example (15) illustrates the use of the possessive DO marker in the function of possession.

The paradigm of possessive markers in Udmurt includes, apart from person and number distinctions, the distinction of the syntactic position of the head noun (DOs vs. non-DOs) and the distinction of (in)alienability, see Yedygarova (2010) and Winkler (2011).³

³ The *-m/-d/-z* variant is used after case markers ending in a vowel.

Person, number	DO set	Non-DO set: inalienable	Non-DO set: alienable
P.1SG	-me	-(j)ə̃ / -m ⁴	(j)e
P.2SG	-de / -te	-(j)ə̃d / -d	(j)ed
P.3SG	-ze / -se	-(j)ə̃z / -z	(j)ez
P.1PL	-mes		-(ə̃)mə̃
P.2PL	-des / -tes		-(ə̃)də̃ / -tə̃
P.3PL	-zes / -ses		-(ə̃)zə̃ / -sə̃

Table 1. The paradigm of possessive suffixes in Beserman Udmurt⁴.

As can be seen from Table 1, a separate set of markers is used with NPs in DO position. This set lacks (in)alienability distinction. Hence, it can be seen that in Beserman Udmurt possessivity and accusative marking are not independent nominal categories, unlike Mari, where DOs with and without accusative marking can both take possessive markers. I will therefore refer to the markers of the first column of Table 1 as “possessive accusative”, while the accusative marker *-ez*⁵ will simply be termed “accusative”.

As in Mari, both unmarked and accusative/possessive accusative DOs take plural morphemes, consider (14) and (16).

- (16) Kal' tože l'ebeda-**os** mar-jos mač'imač'exa
now also goose.foot-**pl** what-PL coltsfoot

okt -iš'ko.

gather-PRS.1SG

Now I also gather goose-foot, what else, coltsfoot (to feed the pigs). (Corpus)

Beserman Udmurt does not impose such severe restrictions on the use of modifiers with unmarked DOs, as Eastern Mari. In Udmurt, unmarked DOs can take all kinds of modifiers, including adjectives, participles, numerals, pronouns, and even demonstrative pronouns and genitive modifiers:

⁴ I am not considering the pronominal set of markers in *-im / -id / -iz*, since the present work is only focused on nouns.

⁵ Note that the accusative marker is homonymous to the possessive suffix of 3^d person singular of the alienable set.

- (17) Zək vōški-je so majal pun-em ber-e
 big keg-ILL that leaven put-NMLZ behind-ILL
- kak so vōški-jez kert-o-z-â val.
 how that keg-ACC tie-FUT-3-PL be.PST

When they have put this leaven into the keg, they tied up the keg. (Corpus)

- (18) I naprimer mil'-a-m priš'aga to mi
 and for.example WE-GEN1- P.1PL oath PTCL we
- baš'-t-i-m faktičeski uč'ebkaj-ân
 take-SMLF-PST-1PL actually training.unit-LOC

F'urstenval'd-ân.

Fürstenwalde-LOC

For example, we actually took our oath in the training unit in Fürstenwalde. (Corpus)

Example (17) includes two DOs with demonstrative pronouns; both refer to aforementioned participants; however, they are marked differently: the first occurs unmarked, while the second takes the accusative.

Note that unlike in Mari, the non-finite clause restriction does not hold in Udmurt (as well as in Komi-Zyrian and in Mordvin languages).

3.1.2. Semantics of unmarked DOs

Unlike in Mari, the main factor influencing the choice of DO marking in Beserman Udmurt is its definiteness, as has been claimed for Standard Udmurt in Perevoshchikov, Vakhrushev & Alatyrev (1962: 93), Csúcs (1990: 34), Winkler (2001: 20), Kondrat'eva (2002). Consider (14) where the aforementioned DO is marked with the accusative and (19) where an indefinite DO does not take any morphological marking:

(19)	Van'a	tāb-i-z	kāz	jāl-e,	kāz
	Vanya	walk.up-PST-3(SG)	firtree	top-ILL	firtree
	jāl-iš'en	so	až'-i-z	korka.	
	top-EGR	that	see-PST-3(SG)	house	

{Vanya and Peter were lost in the wood, they are looking for a place to stay at night.} Vanya climbed up the firtree, and from the top of the firtree he saw a house. (Corpus)

The example (14) contains an aforementioned participant, and the accusative is used. By contrast, the DO in (19) introduces a new participant in the discourse, and it is unmarked. Indefinite non-specific DOs are also unmarked, as ‘beer’ in (13).

Universal and generic DOs with kind reference allow all types of marking, depending on the information structure: DOs belonging to the topic are marked with accusative/possessive accusative, while focused DOs are unmarked. Consider two examples:

(20)	So	bāgat-e	d'eš'	pin'al	bin'-ānā.
	that	can-PRS.3SG	good	child	swaddle-INF

She knows how to swaddle children. (Serdobolskaya, Toldova 2012: 125)

(21)	Ėšk-em	jetān-ez	kert-āl-i-z-ā	kāl't'o-je.
	pull.out-NMLZ	flax-ACC	tie-ITER-PST-3-PL	sheaf-ILL

{A text about working with flax in the beginning of the 20th century. Context: Everybody was engaged while pulling out the flax – elderly people, women, children.} The flax that was pulled out was then tied in sheaves. (Corpus)

In (20) the sentence topic is ‘she’: it can be an answer to a question “What can she do? What is her strength?”, and the DO belongs to the wide focus. Hence, it is unmarked. By contrast, in (21) the whole sentence (and the whole text) is about flax and the works on flax in the village. Accordingly, the DO takes the accusative.

Hence, both the accusative and the possessive accusative are used to signal the definiteness of the DO (and the topicality of generic DOs).

The possessive accusative has some very special functions which are not observed with the accusative, e.g. it can have the “ethical” meaning as defined by (Spencer & Luís 2012: 169): “marking of the participant

that is outside of the argument frame of the verb". Most often, it is the addressee or the person who is in the focus of empathy (in the perspective, as defined in Sonnenhauser (2016)):

- | | | | |
|------|-------------------|---------------|------|
| (22) | Vina- de | ju-iš'ko-m | – o! |
| | vodka-ACC.P.2(SG) | drink-PRS-1PL | o |

{Context: People come to me, and I always treat them. They sing and dance. I sing, too, and drink vodka, my girl.} And how much **vodka** we drink – oh (quite a lot)! (Corpus)

For example, in (22) the 2^d person possessive introduces the addressee (the linguist who is recording the text), who does not have even an associative relation with the participant denoted by the head noun. This function has been observed in Schlachter (1960) in terms of “*Subjektivierung*” and in Winkler (2011: 66) in terms of “*associativer Zusammenhang*”: the speaker wants to represent a participant as important and in a way personally linked to the addressee.

Another function developed by the possessive markers is the activation of a semi-active participant or of the information that is already known or must be evident from the situation. Consider (23) and (24):

- | | | | |
|------|----------------------|--------------------|----------------|
| (23) | D'iš'-se | kâl'-i-z, | vu-ân |
| | clothes-ACC.P.3(SG) | take.off-PST-3(SG) | water-LOC |
| | gâl-t-i-z, | miš'k-i-z, | |
| | rinse-SMLF-PST-3(SG) | wash-PST-3(SG) | |
| | vaj | vâl-e | oš-i-z. |
| | branch | top-ILL | hang-PST-3(SG) |

{Context: The soldier saw a man fall into the swamp. He took him out of the swamp.} He took off **his** clothes, rinsed them in the water, washed them and **hang**ed them onto the branch of a tree.

- | | | | |
|---------------------|----------------|---------------------|----------------|
| Ad'ami-ze=no | | kut-i-z=no | |
| man-ACC.P.3(SG)-ADD | | catch-PST-3(SG)-ADD | |
| vu | puš-k-e | | dong-i-z. |
| water | inside-OBL-ILL | | pull-PST-3(SG) |

Then he caught **the man**, too, and pulled him into the water. (Corpus)

- | | | | | |
|---------|-------------|----------|--------------|---------------------|
| (24) a. | Kâsk-â | kal-ez – | es-ez | uš'-č'k-o-z. |
| | pull-IMP.SG | rope-ACC | door-P.3(SG) | open-DETR-FUT-3(SG) |

Pull **the rope**, and the door will open.

- | | | |
|----|-------------|------------------|
| b. | Kâsk-â | kal-ze! |
| | pull-IMP.SG | rope-ACC.P.3(SG) |

*Come on, pull **the rope!*** {Comment: *My hands are busy, I cannot do that myself and I'm saying: Come on, pull the rope, don't stand watching it.*} (Serdobolskaya, Toldova 2012: 134)

In (23) the bold-faced DO denotes the participant that has already been introduced into the discourse before, and is hence definite. However, in four preceding clauses this referent is not mentioned, hence, it acquires a 'semi-active' status in terms of Chafe (1994). In the second sentence of this example it is mentioned again, and the attention of the addressee is drawn towards this fact by using the possessive accusative (instead of accusative).

The example (24) represents a minimal pair illustrating the use of the accusative vs. possessive accusative. Both the sentences may be pronounced in front of the door of a house that both the interlocutors are entering. By using the accusative in (24a) the speaker simply asks the addressee to perform the action, while in (24b) he is irritated that the addressee is still not performing this action without being reminded, and the possessive accusative is used. I propose to explain this use on the same basis as the one in (23): the speaker wants to activate the information (it is somehow evident that the addressee must open the door: either s/he is younger or his/her hands are busy or s/he is standing closer to the door than the speaker) that is somehow already present in the situation (see Serdobolskaya, Usacheva & Arkhangelskiy (in press) on other uses of the possessive accusative in Beserman Udmurt).

Thus, the system of DOM in Beserman Udmurt is totally different from the one in Eastern Mari. The two systems are distinguished not only in the domain of morphology, but also of syntax and semantics. The basic opposition for Beserman Udmurt is that of definiteness: definite DOs are marked (either by the accusative or by the possessive accusative), while DOs of other referential types are unmarked (with the reservation on topical generics). The distribution of the two markers depends on the variety of factors, where the possessive accusative is reserved for special contexts of definiteness: bridging, ethical meaning, activation of known (evident) information etc.

The opposition of unmarked vs. accusative DOs in Beserman is functionally close to opposition of subject vs. subject-object agreement in Mordvin languages; see Toldova (2017) on Moksha-Mordvin and Serdobolskaya & Toldova (2012) on the Shoksha dialect of Erzya-Mordvin.

It could be expected that the system of DO encoding in Komi would be more similar to Udmurt than the Mari DOM. However, this is not the case, as shown below. I am considering two dialects of Komi-Zyrian, the Pechora dialect and the Izhma dialect.

3.2. Permic: Komi-Zyrian

Reference grammars claim that the most important factors of DO encoding are animacy (Bubrikh 1949 and Fedyuneva 1998), or definiteness (Lytkin 1955: 142; Lytkin & Timushev 1961: 864; Prokusheva 1984). However, the systems of DOM in the dialects of Komi-Zyrian differ to a great extent.

3.2.1. Morpho-syntactic properties of DOM in the Pechora dialect

As has been illustrated in the beginning of the section 3, the system of DOM in the Pechora dialect of Komi-Zyrian includes the ternary distinction of unmarked DOs vs. DOs with the accusative vs. DOs with possessive accusative markers. The paradigm of the possessive markers in the Pechora dialect of Komi-Zyrian involves the person/number distinction and the opposition of DO to other syntactic positions. The (in)alienability is not encoded in Komi. There is a separate marker for the 1st person singular possessors.

Person	DO set	Subject, 1SG	Other syntactic positions and person/number
1	-əs	-əj	-(y)m
2	-tə		-(y)d
3	-sə		-(y)s

Table 2. Possessive markers in the Pechora dialect of Komi-Zyrian

Plural number of the possessor is encoded by a special morpheme *-ny*, which is the same for all persons and syntactic positions. For nouns in the DO position, this marker occurs in combination with the person markers of the non-DO set and the possessive accusative, e.g. *vok-ny-m-əs* (brother-P.PL-P.1PL-ACC.P.1) ‘our brother’ (compare the singular possessor model *vok-əs* (brother-ACC.P.1SG) ‘my brother’, *vok-tə* (brother-ACC.P.2SG) ‘your brother’). In other syntactic positions this suffix is attached before the person marker, e.g. *stav-ny-s-lyš’* (all-P.PL-P.3-GEN2) ‘belonging to everybody’.

A crucial parameter for DOM in Komi-Zyrian is animacy: animate DOs can either take the accusative or the possessive accusative, and non-animate DOs can either take the possessive accusative or occur unmarked (see Klumpp 2014). For example:

- (25) Me aʒʒ'-yvl-i yvla vyl-yš' nyvka-əs /
 I see-ITER-PST(1SG) street top-EL girl-ACC
 nyvka-sə / * **nyvka.**
 girl- girl
 ACC.P.3(SG)

I saw a / the girl in the street.

- (26) Vok-əj pərəd-i-s **pu/** tonə sij-ə pu-sə
 brother-P.1SG chop-PST-3(SG) tree there that-NA tree-ACC.P.3(SG)

Brother chopped a tree / that tree there.

Animate DOs cannot occur unmarked, as shown in (25). Non-animate DOs do not occur with the accusative marker *-əs*. Note that the marker of the accusative is homonymous to the possessive of the 1st person in the DO set (see Table 2). Hence, with animate DOs this marker can be interpreted as accusative ('I saw a girl' in (25)), or the possessive accusative of the 1st person possessor ('I saw my girl/my daughter' in (25)). With non-animate DOs the marker *-əs* can only be interpreted as the possessive accusative of the 1st person possessor:

- (27) Kiš't təryt-ja jəl-sə /
 pour(IMP.SG) yesterday-ATTR milk-ACC.P.3(SG)
 jəl-əs.
 milk-ACC.P.1(SG)

*Pour out yesterday's milk / **my** yesterday's milk.*

Hence, non-animates cannot take the accusative, and animates cannot occur unmarked. This reduces the ternary opposition (unmarked DO vs. accusative vs. possessive accusative) into the binary one, as shown in Table 3. Hence, I will further refer to the two strategies of DOM in terms of '0/accusative' vs. 'possessive' strategy.

Animate	Inanimate
-sə ACC.P.3(SG)	-sə ACC.P.3(SG)
-əs ACC	no marking

Table 3. Animacy and DO encoding in Pechora Komi-Zyrian.

The noun denoting animals can either be treated as animate or non-animate, depending on the animal size and role in the situation. For example, (28) describes the act of fishing (where the properties of the fish are not relevant, and it does not take the accusative, as if it were inanimate). By contrast, in (29) the speaker wants to single out the unusual size of fish his father caught and thus uses the accusative.

- (28) Bat'-ə kyj-ə č'eri.
 father-P.1SG catch-PRS.3(SG) fish
Father fishes/is fishing.
- (29) Sij-a kyj-l-i-s zev yžyd kom-**jas-əs**
 that-NOM catch-ITER-PST-3(SG) very big grayling-PL-ACC
He caught very big graylings. (Corpus)

The noun *kaga* 'child, baby' can also follow the inanimate model:

- (30) Sij-ə viž'-ə **kaga**.
 that-NA look.after-PRS.3(SG) child
She looks after the child / children.

As in Mari and Udmurt, unmarked DOs can take the plural marker:

- (31) Me teč'-al-i **leč'-jas**, ružə men-am vəl-i.
 I put-DISTR-PST(1SG) trap-PL gun I.OBL-GEN1 be-PST
I used to put traps, and I had a gun. (Corpus)

In Pechora Komi-Zyrian, the DO encoding strategies impose severe restrictions on the modifiers of the DO. Unmarked DOs cannot take demonstrative pronouns and the universal quantifier. DOs with indefinite pronouns must take the 0/accusative strategy.

Both 0/accusative and possessive DOs can occur either pre-verbally or post-verbally.

3.2.2. Semantics of DOM in Pechora Komi-Zyrian

The main factor that influences the choice between 0/accusative and possessive DOs in the Pechora dialect is the information structure of the sentence. Topical DOs are marked with the possessive accusative (with minor exceptions), as well as DOs belonging to the tail (see Valduví 1992; “secondary topic” in terms of Nikolaeva 2001).

- (32) Kəni kəluj? – Kəluj-sə təl nun-al-i-s.
 where linen linen-ACC.P.3(SG) wind bring-DISTR-PST-3(SG)

*Where is the linen (that was on the rope)? – The wind brought **the linen** away.*

- (33) Me dərəm-sə vur-i, a
 I dress-ACC.P.3(SG) sew-PST(1SG) and
 e-g nəb.
 NEG.PST-1(SG) buy

*I HAVE SEWN **this dress**, I didn't buy [it].*

In (32) the topic of the sentence is ‘the linen’, as the sentence is an answer to the question about the linen, and the rest of the sentence ‘the wind brought away’ is a new information. In (33) the topic of the sentence is the subject, while ‘the dress’ constitutes given information outside of the focus domain. The verb is focused and bears the contrast. Thus, ‘the dress’ may be interpreted in terms of tail or secondary topic. In both (32) and (33) the DO takes the possessive accusative. By contrast, in the following examples it does not:

- (34) Obradujt-č'-a-s pervoj d'ert mat'erit-a-s
 rejoice-DETR-NPST-3(SG) first of.course scold.using.taboo.words-NPST-3(SG)
pon-j-əs pyšj-al-əm-yš'.
 dog-OBL-ACC run.away-DISTR-NMLZ-EL

(The dog was lost and was missing for a long time. Then it suddenly returned. The master) was glad; of course, first he scolded the dog using taboo words for having run away. (Corpus)

- (35) Myj ləš'-əd-i-s bat'? – Bat'
 what rough.hew-CAUS-PST-3(SG) father father
 ləš'-əd-i-s **potšəs**.
 rough.hew-CAUS-PST-3(SG) fence

What did father rough-hew? Father rough-hewed the fence.

In (34) the whole VP is focused, and the DO belongs to the wide focus (as can be seen from the context, the sentence is the continuation in the narrative and can be conceived as an answer to the question “What did the master do?”). In (35) the DO is in narrow focus, as it is an answer to the constituent question (‘wh-question’) “What did father rough-hew?”. In both cases the 0/accusative strategy is used. Note that in all four examples (32)–(35) the DO is definite; however, different DOM strategies may be used depending on the information structure of the sentence. Thus, for the Pechora dialect the information structure is a more important factor than the definiteness of the DO.

It must be specified that these rules may be violated: 15% of focused DOs take the possessive accusative (if aforementioned), and 20% of DOs in tail take the 0/accusative strategy (for topical DOs it is unacceptable). Still, in more than 80% of cases the main factor that determines the choice between the possessive vs. 0/accusative strategy is the information structure. In this respect the Pechora Komi-Zyrian system differs from Udmurt, where the main factor is the definiteness of the DO. It is also different from Mari. In this language the information structure is also important; however, the opposition is different than the one in Komi-Zyrian.

Hence, in the Pechora dialect all the three DO marking types are quite frequent; their distribution is mostly based on animacy and information structure. The system of DOM in the Izhma dialect is strikingly different.

3.3. Permian: Izhma dialect

3.3.1. Morpho-syntax of DOM in Izhma

In Izhma dialect (after Kashkin (2008) and Biryuk et al. (2010)), the ternary opposition generally observed in Permian (unmarked DOs vs. accusative vs. possessive accusative) is reduced to the binary one; however, the distribution is different than in the dialect of Pechora. In Izhma the accusative marker *-es* is only used with pronouns and proper names:

(36)	Maša	okišt-i-s	Vasilij-es.
	Masha	kiss- <small>PST-3</small> (SG)	Vassilij-ACC

Masha kissed Vassilij. (Kashkin 2008)

Hence, although all the three types of DO marking are possible, only two of them are opposed for common nouns: absence of marking vs. possessive accusative. The paradigm of possessive markers is given in the Table 4.

Person	DO set	Other syntactic positions
1sg	-es	-e
2sg	-te	-yd
3sg	-se	-ys
1pl	-num-es	-num

Table 4. Possessive markers in the Izhma dialect of Komi-Zyrian.

As in Pechora dialect, the possessor plural suffix *-ny* is attached to the stem before the person marker for the 2d and 3d person plural.

An interesting peculiarity of Izhma concerns the semantics of person distinction. Although there are markers of all persons, the speakers freely use the 3^d person possessive to refer to the possessor of 1st and 2^d person (Kashkin 2008). The 3^d person possessive markers can even co-occur with the genitive possessor of 1st/2^d person (which is not possible in the Pechora dialect):

- (37) Menam aj-ys vrač'.
 my father-P.3(SG) doctor

My father is a doctor. (Kashkin 2008)

- (38) Tenad aj-ys vrač'.
 your father-P.3(SG) vrač'

Your father is a doctor. (Kashkin 2008)

This means that the person semantics of this suffix is bleached and it can be used to denote merely possession, without specifying the person of the possessor.

As in other idioms discussed in this paper, the plural number suffix can be attached to unmarked DOs:

- (39) Eni nyy-jas ol-e-ny-s, šək paš'kem-jas
 now girl-PL live-PRS.3-PL-3 silk clothes-PL
 nool-e-ny-s.
 wear-PRS.3-PL-3

Nowadays the girls wear silk clothes (lit. the girls live and wear). (A folk song: Chastushka. Biryuk et al. 2010: 260)

As well as in the Pechora dialect, unmarked DOs cannot take demonstrative pronouns (40), while DOs with possessive accusative cannot take indefinite pronouns (41).

- (40) Vas'a təd-e etae **mort-se /** * **mort.**
 Vas'a know-PRS.3(SG) this man-ACC.P.3(SG) man
Vas'a knows this man. (Biryuk et al. 2010: 237)

- (41) Maša ad'd'il-ema kuč'em-ke **mort /** * **mort-es.**
 Maša meet-PST₂ which-INDEF man man-ACC
(I heard that) Masha has met a man (but I don't know him). (Biryuk et al. 2010: 228)

3.3.2. Semantics of DOM in Izhma

The possessive accusative is obligatory in the context of deictic use:

- (42)a. Əbes-**se** sipty.
 door-ACC.P.3(SG) close.IMP.SG
 b. * Əbes sipty.
 door close.IMP.SG
 a.=b. *Close the/this door.* (Kashkin 2008)

Kashkin (2008) and Biryuk et al. (2010) show that 'aforementioned-ness' is a sufficient condition to trigger the possessive accusative in Izhma, cf.:

- (43) Me mun-i ulič'a kuz'a i ad'd'-i
 I go-PST(1SG) street along and see-PST(1SG)
 pon. **Ponm-ys** / * **pon** kuč'-is uut-ny.
 dog dog-P.3(SG) dog begin-PST.3SG bark-INF
When I was going along the street I saw a dog. The dog started barking.
 (Kashkin 2008)

The examples with aforementioned DOs are not provided; however, the authors claim that the restriction is the same in the DO position. In this respect, the Izhma DOM differs from Pechora: as shown above (3.2.1.), in Pechora, the definiteness (including 'aforementionedness') of the DO does not obligatorily trigger the presence of the possessive

accusative. Definite DOs are usually marked in case they are topical or belonging to the tail. Unlike in the Pechora dialect, the topicality factor plays a minor role in Izhma: Kashkin (2008) states that generic DOs are preferably marked when topical and unmarked when focused; however, this is a mere tendency.

Hence, in Izhma Komi-Zyrian we observe the oppression of the accusative and the expansion of the possessive accusative of 3^d person onto contexts of definiteness in general: according to Kashkin (2008) it is obligatory used to encode unique NPs, aforementioned NPs and NPs with definite determiners. This is supported by the fact that the person semantics of the discussed marker is subject to semantic bleaching: it can be used with genitive possessors of 1st and 2^d person. Thus, Kashkin (2008) and Biryuk et al. (2010) conclude that the marker of the 3^d person possessive accusative in Izhma has grammaticalized as a marker of definiteness.

4. Conclusion

In many typological studies and descriptions of specific languages the label DOM is used as a cover term for a wide range of constructions that sometimes do not have much in common. It is well-known that the Uralic languages demonstrate DOM, or, more specifically, DOE (differential object encoding), since in Mordvin, Ugric and Samoyedic it is a matter of choice between two conjugations, subject vs. subject-object, that is, verbal agreement with the object. Many scholars reconstruct the accusative in **m* for Proto-Uralic, with the reservation that it was probably used for definite DOs only (Wickman 1955: 145-149; Maitinskaya 1974: 241-246; Rédei 1975; Raun 1988). A different solution is proposed by J. Gulya and A. Künnap: Künnap (2006) suggests that discourse activation must have been more important than actual definiteness in Proto-Uralic; a similar idea is expressed by Gulya (1995: 96-97), who proposes to consider object conjugation as a marker of transitivity, rather than of semantic factors as definiteness or topicality. Farkas (1956) and Vértes (1960) claim that the discussed marker had primarily a deictic function (rather than that of a definiteness marker) in the proto-Uralic period. For proto-Hungarian Kiss (2013) proposes that verbal object agreement appeared as a marker of topicality, rather than definiteness, while the definiteness function arose in Old Hungarian.

Although the common diachronic reasons for developing a DOM system are not to be doubted, the systems of DOM are very different in modern languages, and in many of them the definiteness factor is relevant only partly. Even if the research is narrowed on Mari and Permic languages, it is observed that in each language the syntactic and prosodic patterns in DOM differ to a large extent (pseudo-incorporation in Mari vs. ternary opposition in Udmurt vs. two animacy-based binary oppositions in the Pechora dialect of Komi-Zyrian vs. binary opposition in Izhma Komi-Zyrian). The functional distribution of the accusative and non-accusative DOs is also subject to variation. Even the Permic languages (and the dialects of one and the same language), where the common diachronic origin of accusative marker is beyond doubt, develop different systems of DOM. Two binary distinctions in Pechora Komi-Zyrian are based on animacy and topicality, whilst Izhma Komi-Zyrian shows the side-lining of the accusative and the expansion of the possessive accusative. This actually leads to a binary opposition based on definiteness. Definiteness is also important in Beserman Udmurt (as well as in Mordvin languages); however, unlike in Izhma, it is not a fully grammaticalized distinction, and in most contexts of definiteness the accusative marker can be dropped (even with demonstrative pronouns and genitive modifiers). The possessive accusative in this language is used to encode some discourse functions, as the ethical function, bridging, reactivation of semi-active topics and others.

The conclusion then can be made that although the label DOM is used to characterize the Uralic languages in general, it must be admitted that specific semantic, morpho-syntactic and prosodic features of the constructions under discussion must be considered as innovative features. It can be speculated that such features (the specific forms of DOM) are subject to recent and very fast language changes, and they therefore can only be used with caution in reconstructions of the model of DOM in Proto-Uralic.

Abbreviations

DO	direct object
DOM	differential object marking
NP	noun phrase
ACC	accusative
ADD	additive particle
ATTR	attributive
CAUS	causative
CONV	converb
DAT	dative
DETR	detransitive
DISTR	distributive
EGR	egressive
EL	elative
FUT	future
GEN	genitive
GEN1	genitive 1 (used with non-DOs)
GEN2	genitive 2 (used with DOs)
ILL	illative
IMP	imperative
INDEF	suffix of indefinite pronouns
INESS	inessive
INF	infinitive
ITER	iterative
LOC	locative
NA	nominative/accusative form of pronouns
NARR	narrative past
NMLZ	nominalization
NOM	nominative
NPST	non-past (present and future)
OBL	oblique stem
OPT	optative
P	possessive
PL	plural
PRS	present
PST	past
PST2	evidential past
PTCL	particle
PTCP.ACT	active participle
REFL	reflexive
SG	singular
SMLF	semelfactive

References

- ANDUGANOV, J.D. 1991. *Istoričeskaya grammatika marijskogo jazyka. Sintaksis*. Vol. 1: *Vvedenie. Substantivnye slovosočetačija*. Joškar-Ola: Marijskoe knižnoe izdatel'stvo.
- BESE, L., DEZSŐ, L. & GULYA, J. 1970. On the syntactic typology of the Uralic and Altaic languages. In L. Dezső & P. Hajdú (eds.), *Theoretical problems of Typology and Northern Eurasian Languages*. Budapest: Akadémiai Kiadó. 113-128.
- BIRYUK, et al. 2010 = Biryuk, O.L., Kashkin, E.V., Kuznetsova, A.I. & Usacheva, M.N. 2010. *Slovar' muzhevskogo govora izhemsckogo dialekta komi-zyryanskogo jazyka*. Ekaterinburg: Basko.
- BOSSONG, G. 1985. *Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Gunter Narr.
- BUBRIKH, D.V. 1949. *Grammatika literaturnogo komi jazyka*. Leningrad: LGU.
- CHAFE, W. 1994. *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. Chicago: University of Chicago Press.
- COLLINDER, B. 1957. *Survey of the Uralic languages*. Stockholm: Almqvist & Wiksell.
- CSÚCS, S. 1990. *Chrestomathia Votiacica*. Budapest: Tankönyvkiadó.
- DE SWART, P. 2007. Cross-linguistic variation in object marking. Ph.D. diss. Nijmegen: Radboud Universiteit.
- EDYGAROVA, S. 2010. *Kategoriya posessivnosti v udmurtskom jazyke*. Tartu: Tartu Ülikooli Kirjastus.
- VON FARKAS, J. 1956. Der Genitiv und der Akkusativ in der uralischen Grundsprache. *Ural-Altäische Jahrbücher* XXVIII: 5-17.
- FEDYUNOVA, G.V. (ed.) 1998. *Komi jazyk. Enciklopediya*. M.: DIK.
- FRAURUD, K. 2001. Possessives with extensive use: A source of definite articles? In I. Baron, M. Herslund & F. Sørensen (eds.), *Dimensions of possession (Typological Studies in Language)*. Amsterdam / Philadelphia: Benjamins. 243-267
- GALKIN, I.S. 1964. *Istoričeskaja grammatika marijskogo jazyka. Morfoložija*. Vol. 1. Joškar-Ola: Marijskoe knižnoe izdatel'stvo.
- GULYA, J. 1995. Transitivität, Determiniertheit und objektive Konjugation. *Congressus Internationalis Finno-Ugristarum VIII*. Pars III. 96-99.
- HEINÄMÄKI, O. 1984. Aspect in Finnish. In: C. de Groot & H. Tommola (eds.) *Aspect Bound: A voyage into the realm of Germanic, Slavonic and Finno-Ugrian aspectology*. Dordrecht/Cinnaminson: Foris. 153-177.
- HEINÄMÄKI, O. 1994. Aspect and Boundedness in Finnish. In C. Bache, H. Basbøll & C.-E. Lindberg (eds.), *Tense, Aspect and Action. (EALT 12)*. Berlin /New York. 207-233.
- KASHKIN, E.V. 2008. Posessivnoe sklonenie v muzhevskom govore izhemsckogo dialekta komi zyryanskogo jazyka. *Vestnik YANAO* 6: 33-46.

- KISS, K.É. 2013. From Proto-Hungarian SOV to Old Hungarian Top Foc V X. *Diachronica* 30 (2): 202-231.
- KLUMPP, G. 2014. Identifiability, givenness and zero-marked referential objects in Komi. *Linguistics* 52 (2): 415-444.
- KONDRAT'ÉVA, N.V. *Vyrazhenie pryamogo ob'ekta v udmurtskom yazyke (v istoriko-sopostavitel'nom plane)*. Avtoref. dis. ... kand. filol. nauk. Izhevsk: Udmurtskii gos. un-t, 2002.
- KORHONEN, M. 1996. Remarks on the structure and history of the Uralic case system. In T. Salminen (ed.), *Typological and Historical Studies in Language by Mikko Korhonen. A Memorial Volume published on the 60th Anniversary of his Birth*. Helsinki: *Mémoires de la Société Finno-Ougrienne* 223. 219-234.
- KÜNNAP, A. 2006. On the Uralic (*)*m*-accusative. *Linguistica Uralica* XLII (3): 161-166.
- KUZNETSOVA, A.I. 2003. Kumulaciya grammaticheskix znachenij v agglutinativnyx pokazatelax: deikticheskiye funkicii possessiva v uralskix jazykax. In P. Suihkonen & B. Comrie (eds.), *Collection of papers of International symposium on deictic systems and quantification in languages spoken in Europe and North and Central Asia*. Izevsk / Leipzig. 249-259.
- LYTKIN, V.I. 1955. *Sovremennyj komi jazyk*. Vol. 1: *Fonetika, leksika, morfologiya*. Syktyvkar: Komi knizhnoe izd-vo.
- LYTKIN, V.I. & Timushev, D.A. 1961. Kratkij ocherk grammatiki komi jazyka. In Lytkin V.I. (ed.), *Komi-russkij slovar*. Moscow: Gos. izd-vo inostrannyh i nacional'nyh slovarej. 839-923.
- MAITINSKAYA, K.E. 1974. Sravnitel'naya morfologiya finno-ugorskikh jazykov. In *Osnovy finno-ugorskogo jazykoznanija (voprosy proiskhozhdeniya i razvitiya finno-ugorskikh jazykov)*. Moscow: Nauka, 214-382.
- MAITINSKAYA, K.E. 1979. *Istoriko-sopostavitel'naya morfologiya finno-ugorskikh jazykov*. Moscow.: Nauka.
- NIKOLAÉVA, I. 2001. Secondary topic in information structure. *Linguistics* 39 (1): 1-49.
- PENKITOV, N.T. (ed.) 1961. *Sovremennyj marijskij jazyk*. Vol. 2: *Morfologija*. Joškar-Ola: Marijskoe knižnoe izdatel'stvo.
- PEREVOSHCHIKOV, P.I., Vakhrushev, V.M. & Alatyrev, V.I. (eds.) 1962. *Grammatika sovremennogo udmurtskogo jazyka. Fonetika i morfologiya*. Izhevsk: Izhevskoe knizhnoe izdvo.
- PROKUSHEVA, T.I. 1984. Kategorija opredelennosti / neopredelennosti v komi jazyke. In G.I. Tiraspol'skij (ed.), *Vzaimodejstvie finno-ugorskikh i russkogo jazykov*. Syktyvkar: Komi filial AN SSSR,
- RÄTSEP, H. 1957. Aspektikategorii eesti keeles. *Emakeele Seltsi Aastaraamat* III: 72-77.
- RAUN, A. Proto-Uralic Comparative Historical Morphosyntax. In D. Sinor (ed.) *The Uralic languages. Description, history, and foreign influences*. Leiden / New York / København/Köln: Brill. 555-571.

- RÉDEL, K. 1975. Der Akkusativ in den uralischen Sprachen (Zum Kasussystem der uralischen Grundsprache). *Congressus Internationalis Finno-Ugristarum III. Pars I.* 133-138.
- RÉDEL, K. 1988. Geschichte der Permischen Sprachen. In D. Sinor (ed.), *The Uralic languages. Description, history, and foreign influences.* Leiden /New York/ København/ Köln: Brill. 351-395.
- SCHLACHTER, W. 1960. *Studien zum Possessivsuffix des Syrjänischen.* Berlin: Akademie Verlag.
- SERDOBOLSKAYA, N. 2015. Pseudo-incorporation analysis of Unmarked Direct Objects in Mari. In B. Gehrke & O. Borik (eds.), *The Syntax and Semantics of Pseudo-incorporation (Syntax and Semantics, v. 40).* Leiden / Boston: Brill:2015, 298-330.
- SERDOBOLSKAYA, N. & Toldova, S. 2012. Differencirovanoe markirovanie pramogo dopolneniya v finno-ugorskix jazykax. In A.I. Kuznetsova, N.V. Serdobol'skaya, S.Yu. Toldova, S.S. Sai & E.Yu. Kalinina (eds.), *Finno-Ugorskije jazyki: fragmenty grammatičeskogo opisaniya. Formal'nyj i funkcional'nyj podxody.* Moscow: Jazyki slavyanskikh kul'tur. 59-142.
- SERDOBOLSKAYA, N., Usacheva, M. & Arkhangelskiy, T. (in press). Grammatization of possessive markers in the Beserman dialect of Udmurt. In L. Johanson, I. Nevskaya & L.F. Mazzitelli (eds.), *Linguistic possession. New insights from the languages of Europe and North and Central Asia.* Amsterdam: Benjamins.
- SIMONENKO, A.P. & LEONT'EV, A.P. 2012. Morfosintaksis imennogo kompleksa v finno-permskikh jazykakh: analiz v ramkax programmy minimalizma. In A.I. Kuznetsova, N.V. Serdobol'skaya, S.Yu. Toldova, S.S. Sai & E.Yu. Kalinina (eds.), *Finno-ugorskije jazyki: Fragmenty grammatičeskogo opisaniya. Formal'nyj i funkcional'nyj podxody.* Moscow: Jazyki slavyanskikh kul'tur. 262-340.
- SONNENHAUSER, B. 2016. Presenting narration: The perspectival construction of narrativity in Bulgarian and Macedonian. In N. Igl & S. Zeman (eds.), *Perspectives on Narrativity and Narrative Perspectivization.* Amsterdam /Philadelphia: Benjamins. 43-62.
- SPENCER, A. & LUÍS, A.R. 2012. *Clitics: An Introduction.* Cambridge: Cambridge University Press.
- SUIHKONEN, P. 2005. On the categories and functions developed from the possessive and deictic suffixes in Udmurt. In C. Hasselblatt, E. Koponen & A. Widmer (eds.), *Lihkkun lehkos! Beiträge zur Finnougristik aus Anlaß des sechzigsten Geburtstages von Hans-Hermann Bartens.* Wiesbaden: Harrassowitz Verlag. 401-432.
- TAMM, A. 2004. *Relations between Estonian verbs, aspect, and case.* Doctoral dissertation. Budapest.
- TOLDOVA, S.Ju. 2017. Kodirovanie pryamogo dopolneniya v mokshanskom jazyke. *Acta Linguistica Petropolitana.* Saint-Petersburg: Nauka. (In press).

- TOLDOVA, S. & SERDOBOLSKAYA, N. 2002. Nekotorye osobennosti oformlenija prjamogo dopolnenija v marijskom jazyke. In O.A. Kazakevich (ed.), *Lingvističeskij bespredel*. Moscow: Nauka. 106-125.
- TUŽAROV, G.M. 1984. Problema nemarkirovannogo imeni v marijskom jazyke. *Sovetskoje finno-ugrovedenie* 20 (4): 282-289.
- TUŽAROV, G.M. 1986. Problema nemarkirovannogo akkuzativa v marijskom jazyke. *Sovetskoe finno-ugrovedenie* 22 (2): 99-107.
- TUŽAROV, G.M. 1998. *Grammatičeskije kategorii imeni suščestvitel'nogo v marijskom jazyke*. Diss. hab., Joškar-Ola.
- VALLDUVÍ, E. 1992. *The Informational Component*. New York: Garland.
- VÉRTES, E. 1960. Beitrage zur Frage des finnisch-ugrischen bezeichneten Akkusativobjekts. *Acta Linguistica Hungarica* X: 181-194.
- WICKMAN, B. 1995. *The Form of the Object in the Uralic Languages*. Uppsala / Wiesbaden: Almqvist & Wiksells Boktryckeri.
- WINKLER, E. 2011. *Udmurtische Grammatik*. Wiesbaden: Harrassowitz Verlag.
- WITZLACK-MAKAREVICH, A. & SERŽANT, I.A. 2017: Differential argument marking: Patterns of variation. In I.A. Seržant & A. Witzlack-Makarevich (eds.), *The Diachronic Typology of Differential Argument Marking*. Berlin: Language Science Press. <http://langsci-press.org/>.

The 'impossible' Comparison between the Creatures and Figures of the Hungarian Folklore and the Finnish Mythology

Elisa Zanchetta

1. Introduction

Unlike the other Uralic (U) peoples, the Hungarians did not develop folk beliefs rich in mythical creatures. Even a comparison with the mythology of the Voguls (/Mansi) and the Ostyaks (/Khanty), who are (supposedly) most closely related to them from a linguistic point of view, turns out to be impossible, due to the scanty Hungarian material. As a matter of fact, Vogul and Ostyak mythology is so rich in gods and spirits that it is impossible to establish precisely both their number and their names. The Finnish mythology is equally rich and articulated. The reasons for this difference are unknown at the present state of research, though we may re-state what we have already read in our text books, that is: this peculiarity of the relatively 'poor mythological world' of the early *Magyars* may be traced back to a combination of the following factors: their (supposedly) early separation from proto-U and then from the Ugric branch (explanation that is, however, circular); the introduction of the Christian faith, that surely must have eradicated most of their 'pagan' beliefs; the historical events the Magyars underwent, that is, their new geographical position in the Carpathian Basin and related cultural influences from their new neighbours, etc. Dömötör (1983: 83) also hypothesized that the Hungarians preferred to believe in the ability of skilled humans to improve their own destiny, rather than expecting support and protection from the spirits.

As widely described in Castren's work (Castren 1853: 1, 4), the Finns peopled nature with spirits since they became aware of the fact that men could not and did not have absolute control over the powers that influenced their lives. Similarly, the shaman played a key role in

acting as an intermediary between the human world and the world of the spirits. Thus, the more or less rich range of spirits present in the two cultures may be due (also) to the different ways the Hungarians and the Finns tried to explain what happens to both nature and man – of course, different cultures produce different responses to the necessities arisen from natural environment and historical events.

The aim of this article is to outline the state of the art of the Hungarian and Finnish mythological world. Furthermore, I shall propose ideas for future comparative, contrastive research involving also the (Northern) Germanic world. Hungarian and Finnish are two languages that are considered to be linguistically related, but whose folklore is different, especially as far as the amount of mythical beings is concerned. Nevertheless, I believe, what is important is not to set aside the possibility of ‘reconstructing’ Hungarian folklore through mythological comparison (as much as possible), and to further advance linguistic as well as anthropological research. As a matter of fact, ‘mythological reconstruction’ might help to shed light on linguistic and philological aspects of U studies that are still unknown.

2. Mythology or folklore?

I will use the term ‘mythology’ also when dealing with Hungarian, since my research is based on the general definition of mythology given by Hoppál (1989: 147), who stated that he would deal with Hungarian mythology conceived as: a “system of notions, reconstructed thanks to folk beliefs, folk tales, sagas, sayings, ancient prayers, incantations, folk arts”. As a consequence, Hungarian folk tales acquire more prominence. According to Steiner (1980: 28), there are two reasons why these folk beliefs, tales etc. should be closely analysed in order to reconstruct the mythology of the heathen Hungarians. Firstly, folk tales offer an almost pure vision of the primordial elements on which a culture is built: since the fantastic is to be located beyond reality, it can’t be modified by historical events. Secondly, folk tales originate when a community abandons its ancient religion and adopt a new faith. Consequently, their ancient beliefs lose their doctrinal content and survive as secular, immutable tales whose only aim is to entertain. Nevertheless, folk tales characters should always be analysed separately from those portrayed in folk beliefs, since they may sometimes differ, though they have the same name. An example taken from Dömötör (1983: 85) may help to

clarify this statement. In Hungarian folk beliefs the *táltos* is only a human being, whereas in folk tales this noun identifies the horse the hero acquired after managing to reach the top of the *égig érő fa* ('tree reaching the sky'), and serving the dragon or the *vasorrú bába* ('old woman with the iron nose'). Nonetheless, we must notice that both concepts are closely related to one another, since the horse represents the *táltos* drum that enables the *táltos* to travel through the different levels of the cosmos, riding its sound (Steiner 1980: 77).

Before trying to compare Hungarian and Finnish mythical beings, it is worth taking a look to their cosmological conception, in order to have a clear idea of the places where these creatures operate, as well as to outline similarities and differences that may be useful to the reconstruction of the Hungarian and Finnish cosmological conception. In doing so, Hungarian folk tales play a key role, since the mythological world depicted in them coincides with that portrayed in folk beliefs, as pointed out in various numbers of the folklore magazine *Magyar Néprajz*.

3. The Cosmological Conception of the Hungarians and the Finns

Both the Hungarians and the Finns imagined the Earth as a flat disc surrounded by water. For example, the *Magyar Néprajzi Lexikon* presents the world of Lajos Ámi's¹ tales as totally surrounded by a vast expanse of water which may coincide with the so called *Óperenciás tenger*² 'Óperencia Sea', a recurrent mythological *topos* in folk tales opening formulas. The Finns thought the Earth was surrounded by water conceived as a boiling river (Corradi 1983: 52). Another recurrent expression of Hungarian folk tales opening formulas presented in *Magyar Néprajz* is *Üveghegyek* 'glass mountains', situated on the edge of the world where the huge tent-like celestial vault is so low that it doesn't let sunlight get through. The Glass Mountains represent the boarder beyond which the Other World stretches vertically.

¹ Lajos Ámi (1886-1962) was a Hungarian story-teller of half-tzigane origin.

² In Hungarian folk tales the expression *Óperenciás tenger* refers to a far off territory situated beyond the boundaries of the World. An example taken from the folk tale entitled *A mindent járó malmocska* 'the small all-grinding mill', may give a clue of its function in depicting the imaginary world of folk beliefs: *Volt, hol nem volt, még az Óperenciás-tengeren is túl, ahol a kis kurta farkú malacska túr...* 'Once upon a time there was or there was not, even beyond the Óperencia Sea, where the short-tailed piglet grubs...'.

In the Finnish mythology too we encounter the World's Mountain: it is situated in the Northern Region, and the dead peoples have to reach it, after going through a terrible whirl called *kurimus*. The dead peoples arrive at the Walls of Pohjola, whose gates are so smooth that they would manage to climb them only if they had kept the nails they had cut during their lifetime. A conception shared by both the Finns and the Hungarians is that of the low sky, reaching the earth: this place, located at the edge of the Earth, could be inhabited only by small creatures, such as birds. The Finns regarded it as the southern, warm *Lintukotola* 'country of birds', that is the abode reached by the dead in the shape of bird, as well as the place inhabited by the dwarfs (Corradi 1983: 52). This is a clue of the close connection between dwarfs and dead peoples, concept shared also by the Germanic tribes: the elves were called *álfar* in Old Norse, they were worshipped like the dead and sacrifices were made on a heap reminding a burial mount (Turville-Petre 1964: 303-308).

The World's Tree – known also as the World's Pillar among the Finns (Corradi 1988: 48) – rose up from the World's Navel (*föld köldöke* in Hungarian), that is from the middle of the Earth. The branches of the World's Tree support the Sky, this being conceived as a capsized cauldron revolving around the world by the Hungarians (Kiszely 1996: 512), and as a huge dome scattered with stars – *kirjokansi* – by the Finns (Corradi 1983: 51). The World's Tree could be represented by a cherry tree and by a castle revolving on the claw of a bird (cock, duck or goose) in Hungarian folk beliefs and tales (Corradi 1988: 50-51, 59), while in Finnish mythology it could be also a huge mountain or a Big Oak – *iso tammi* (Corradi 1983: 54). The Hungarian World's Tree, imagined as a castle, is well described in the folk tale *Hajnal, Vacsora, Éjfél* ('dawn, dinner, midnight') collected by the writer Grandpierre Emil Kolozsvári, and included in the volume entitled *A csodafurulya. Magyar népmesék* ('the magic flute. Hungarian folk tales'; Kolozsvári 1954):

Alig nézett körül odabent, hát egy tizenkétemeletes palotát látott. Nem a földön állott ez a palota, hanem egy kakaslábon / récelábon / lúdlábon. S hogy véletlenül le ne csússzék róla, aranylánccal szorosan hozzákötötték, a lánc másik végét pedig egy aranycsillagra hurkolták. Ez a palota mindig arra fordult, ahonnan az áldott nap sütött (Kolozsvári 1954: 348, 350, 352)
 He had just looked around under there when he saw a twelve-storied palace. This palace did not stand on the earth, but on the claw of a cock / duck / goose. It had been firmly fasten with a golden chain, the other

end of the chain was instead knotted to a golden star, so that it couldn't slip down accidentally. This palace kept on revolving towards the direction where the blessed sun shone.

The Finns thought that the Milky Way had roots and crown, and was called *linnun rata*, the 'way of the birds' (Corradi 1983: 53), while the *Magyars* knew various designations for it: *tündérek útja* 'fairies' way', *szépasszonyok útja* 'fair ladies way' and *szépasszonyok vászna* 'linen of the fair ladies' (in Transylvania, for instance), all alluding to ancient mythological conceptions now forgotten (Dömötör 1982: 225). Celestial bodies are found among the branches of both the Hungarian and the Finnish World's Tree: the Sun – imagined like a golden wheel – on the Finnish Big Oak (Corradi 1988: 48); the Sun, the Moon and the Stars on the Hungarian World's Tree (Kiszely 1996: 512). In both cultures the Sun and the Moon had their own houses: the Hungarians believed they lived in golden and silver castles, the Finns in golden and silver rooms (Corradi 1983: 50, Hoppál 1989: 149). The World's Tree connected the upper, the middle and the lower world; its roots went down into the underworld traditionally inhabited by fish, frogs and snakes – in Hungarian folklore also by lizards (Diószegi 1967: 15). The concept of the Other World presents a dichotomy in Hungarian folklore, creating a horizontal pair of oppositions: it can be referred to as *alvilág* ('the world underneath'), thus denoting the dwelling place of the damned, or as *másvilág*, *túlvilág* (the 'other world'), that denotes instead a more neutral concept, simply identifying the realm of the deceased (Hoppál 1989: 149-150). According to *Kalevala* mythology, the Finnish land of the dead was *Manala* ('under the earth'), or *Tuonela* ('the abode beyond'), where all the dead went, notwithstanding the cause of their death (Lönnrot 1985 XVI, 178-180), after crossing the black river of *Tuonela*. The access to the hereafter seemed to be placed in a northern country called *Pohjola*. For example, the suitors of Louhi's daughter were put through difficult trials that involved also the capture of *Manala*'s animals (Lönnrot 1985, XIV: 373-380; XIX: 101-110, 155). Furthermore, *Märkähattu karjanpaimen* ('the shepherd with the wet hat') ran out of *Pohjola*'s hut, crossed the courtyard and reached *Tuonelan joki* 'the river of Tuonela' (Lönnrot 1985, XII: 495-500). We may suppose that *Pohjola* corresponds to the Germanic *Niflheimr*³ 'fog-world'

³ In Germanic mythology *Niflheimr* is the ancient underworld that existed before the creation of the world and functions also as abode of the dead.

(as described in *Gylf*⁴ 49 and *Bdr*⁵ stanza 2). Thus we may say: *Pohjola* is to *Manala* as *Niflheimr* is to *Hel*⁶. According to Hungarian folklore, the shepherd with shamanistic skills could descend to the Other World through a hole at the base of the World's Tree (Hoppál 1989: 149): this opening is indicated by the word *lik* 'hole' in folk tales (but it could be a well in ruins as well), and could find correspondence both in Sami and Finnish mythology, in particular with the black hole in the centre of the Sun drawn on Sami shaman drums to indicate the 'hole of the spirit', and the Finnish *lovi* 'stone opening' (Piludu 2007: 97). As explained in *Magyar Néprajz*, in Hungarian folk beliefs the paradise was a bright abode in the sky, while in Finnish mythology we seldom find its description (Pentikäinen 1999: 204). The festivities of *Päivölä* may be the portrait of the bright Finnish paradise (Pentikäinen 1999: 199), presenting analogies both with the shining abode of the fallen Germanic warriors *Valhöll* 'carrion-hall' (*Grm*⁷ 8), and with that of the damned called *Náströnd*⁸ 'shore of the corpses', whose walls were made of weaved reptiles (*Gylf* 52).

4. Spirits and Gods

The cosmos so conceived was thought to be inhabited and ruled by a variety of gods and spirits. In ancient times the Finns worshipped celestial bodies and natural elements, such as stones, fire, water, trees and the earth (Di Luzio & Giansanti 2014: 53). Natural and human beings were protected by both a supreme God and a *haltia*, that is a guardian spirit that took care of them until their death (Comparetti 1898: 183). Also the Hungarians peopled nature with spirits, that, however, didn't have any power over the the creatures of the world, but

⁴ *Gylf* = *Gylfaginning* 'deluding of Gylfi'; it is the first section of the *Snorra Edda*, that is a manual of poetics composed by the Icelandic politician and learned Snorri Sturluson around 1220.

⁵ *Bdr* = *Baldrs Draumar* 'Baldr's dreams'; it is an Eddic poem that describes Óðinn's rode to *Hel* in order to ask a dead seeress about the meaning of his son's evil dreams.

⁶ *Hel* is assumed to be the old Germanic conception of the netherworld. According to the Eddic composition entitled *Sigrdrífumál* 'Sigrdrifa's song', *Hel* was the abode of those who died of illness.

⁷ *Grm* = *Grímnismál* 'words of Grímnir', is a poem found in the *Poetic Edda*. *Valhöll* is the paradise of the dead heroes.

⁸ In Germanic mythology *Náströnd* is the huge hall of the damned situated far from the sun and with its door towards the North.

simply scared people passing by (Hoppál 1989: 151), thus resembling the Finnish spirits of the deceased called *Männingäiset* and *Kööpeli*. In order to describe them, Castrén refers to Cristfried Ganander (18th century Finnish compiler of folk culture) in his work: according to the latter, the *Männingäiset* were shadows, church ghosts which could be found in the hollows along mountain ridges or in isolated places in order to scare passers-by. The *Kööpeli(s)* represented the spirits of those who had committed serious crimes, consequently they weren't buried adequately, thus remaining by their corpse on earth (Castrén 1853: 124-125).

Finnish mythological creatures can be distinguished between good and evil: air gods, female deities of wood and the majority of water gods were benevolent, while spirits and gods related to agriculture didn't enjoy much worship (Castrén 1853: 72, 92). Yet Hungarian mythical beings can't be classified as such, since the majority of them are malevolent or behave ambiguously. Furthermore, a clear distinction between groups of spirits or gods ruling over a specific area (air, water, earth gods, deities of the 'Other World') is impossible to make, because Hungarian mythical beings are capable of metamorphosis, or may exist simultaneously in different shapes; for example, the *lidérc*, that is, the *ignis fatuus*, can take the shape of a chicken or of a human being, and the dragon – *sárkány* – can also appear in human form (Dömötör 1992: 85).

As stated by Hoppál (1989: 151), an "important mythological feature" of Hungarian folk beliefs is that creatures are not grouped in families with a house and a court, as in Finnish mythology – recall *Ahto*, the male water god, who lived in *Ahtola* together with his wife *Wellamo*, and the *ahtolaiset*, referring to both their children and servants (Castrén 1853: 72-79).

4.1. Air gods

Among the Finns the Sun /Day (*Päivä*), the Moon (*Kuu*), the Star (*Tähti*) and the North Star (*Otava*) were imagined as male deities with a wife, a daughter and a son (Castrén 1853: 51-55). In Hungarian folklore mention is made of the Sun, the Sun's Mother (*Nap anyja*), the Moon and the Moon's Mother (*Hold anyja*), dwelling on the World's Tree, and the Wind Mother (*Szélanya*), imagined as an old woman guarding the wind in a cave (Hoppál 1989: 151). Both the Hungarians and the

Finns believed in the existence of a creature devouring planets and causing eclipses, known as *Markoláb* in Hungarian folklore (Dömötör 1992: 223), and as *Kapeet* among the Finnish beings listed by the bishop Agricola (see Di Luzio & Giansanti (eds.) 2014: 71).

4.2. Water gods

Hungarian water gods cannot be labelled by a common term used all over the country, quite the contrary: only the communities living along water courses adopted several names (Dömötör 1992: 105). Like in many other cultures, also Hungarian water deities were evil, while the Finnish praised them highly, except for a few cases, such as: *wesi-Hiisi* 'water-Hiisi', *wetehinen* 'the one that abides in the water', *Tursas*, *iku-Turso* 'the eternal *Turso*', *meri-Tursas* 'sea-*Tursas*', etc. (Castrén 1853: 83-85). Malevolent water creatures appear in Hungarian folk tales: in *A zöldszakállú király* ('the green beard king'), the devil (*ördög*) pulls the poor man or the king into water while they are drinking.

4.3. Earth gods

Among the earth deities, the Hungarians worshipped the corn spirit in the shape of a bear, a wolf or some mythical forefather, whilst the mythical beings of the forest were known mostly in mountainous areas (Dömötör 1992: 107). Instead, in Finnish mythology the gods of wood were the most worshipped among earth divinities (Castrén 1853: 91-92). This difference is to be traced back to the climate of the territories where these peoples settled down: the warmer climate enabled the Hungarians to practise agriculture, whereas the Finns had to struggle in order to live in a freezing cold region with an infertile soil.

4.4. Diseases and death

Illnesses and death were thought to be caused by some supernatural forces: for the Hungarians an illness was the personification of a given disease, or it could also be attributed to the *lidérc* 'ignis fatuus' (see above; see Dömötör 1992: 89). The Finns blamed *Loviatar*, daughter of the death god *Tuoni*, for the birth of the nine plagues torturing man (Lönnrot 1985, XLV: 39-44, 147-176).

5. Conclusion

As I stated at the beginning, a comparison between Hungarian and Finnish mythology seem to be at first 'impossible', but this doesn't mean that trying a comparative, contrastive analysis is a nonsense task. As a matter of fact, in addition to highlighting the diversity among the two mythological cultures, we may also notice overlaps, though – it may be objected – they are very few and can be found also in the mythology of other populations.

In conclusion, my speech intended not only to highlight the difficulties in comparing the mythology of two (supposedly) related peoples, but also to stress the necessity for further comparative research within the mythological world of the U languages, taking into account also Hungarian folk tales and *gyermekijesztők* ('expressions for frightening kids'; Hoppál 1989: 152), as well as the need for collaboration with experts of non-U mythology, such as Germanic and Nordic philology.

References

- CASTRÉN, A. 1853. *Vorlesungen über die finnische Mythologie*. Im Auftrage der Kaiserlichen Akademie der Wissenschaften aus dem Schwedischen übertragen und mit Anmerkungen begleitet von A. Schiefner. St. Petersburg: Buchdruckerei der Kaiserlichen Akademie der Wissenschaften.
- COMPARETTI, D. 1898. *The Traditional Poetry of the Finns*. London /NewYork / Bombay: Longmans, Green & Co.
- CORRADI, C. 1983. *I Finni*. Parma: Palatina Editrice.
- CORRADI, C. 1988. *Sciamanesimo e flora sacra degli ugrofinni in una prospettiva indouralica ed amerindia del Nord*. Roma: Carocci.
- DI LUZIO, F. & GIANANTI D. (eds.) 2014. *Mikael Agricola. Gli dèi di Finlandia e di Carelia. Hämäläisten ja karjalaisten jumalat*. Viterbo: Vocifuoriscena.
- DIÓSZEGI, V. 1967. *A pogány magyarok hitvilága*. Budapest: Akadémiai Kiadó.
- DÖMÖTÖR, T. 1983. *Hungarian Folk Beliefs*. Budapest: Athenaeum Printing House.
- HOPPÁL, M. 1989. Hungarian Mythology: Notes to Reconstruction. In M. Hoppál & J. Pentikäinen (eds.), *Uralic Mythology and Folklore*. Budapest-Helsinki: MTA Sokszořitó. 147-154.
- KISZELY, I. 1996. *A magyarság őstörténete (Mit adott a magyarság a világnak)*. I-II. Budapest: Püski.
- KOLOZSVÁRI, G.E. 1954. *A csodafurulya. Magyar népmesék*. Budapest: Ifjúsági könyvkiadó.

- LÖNNROT, E. 1985. *Kalevala. Das finnische Epos*. Stuttgart: Philipp Reclam.
- Magyar Néprajz: <http://www.tankonyvtar.hu/hu/tartalom/tkt/magyar-neprajz-magyar/ch26s06.html>.
- ORTUTAY, Gy. 1979. *Magyar Néprajzi Lexikon*. Budapest: Akadémiai Kiadó. (<http://mek.niif.hu/02100/02115/html/>)
- PENTIKÄINEN, J. 1999. *Kalevala Mythology. Expanded edition*. Bloomington / Indianapolis: Indiana University Press.
- PILUDU, V.M. (ed.) 2007. *Sulle tracce della renna del cielo. Scritti sullo sciamanesimo nordico di Juha Pentikäinen e Anna Leena Siikala*. Roma: Bulzoni Editore.
- SCARDIGLI, P. 1982. *Il Canzoniere Eddico*. Milano: Garzanti.
- STEINER, A. 1980. *Sciamanesimo e folklore. Elementi sciamanici nelle favole ungheresi*. Parma: Edizioni all'Insegna del Veltro.
- STURLUSON, S. 1975. *Edda*. Milano: Adelphi.
- TURVILLE-PETRE, E.O. G. 1964. *Religione e miti del Nord*. Milano: Il Saggiatore.
- VON SEE, K. 2000. *Kommentar zu den Liedern der Edda. Band 3: Götterlieder*. Heidelberg: Universitätsverlag C. Winter.

Contributors and abstracts

Giuseppe (Pino) Cossuto

Studied Islamic studies, History and Anthropology at Sapienza University of Rome, and Ottoman History and Language in Bucharest and Istanbul. In 2002 he obtained his Ph.D. in 'Islamic Civilization: History and Philology', defending a dissertation titled: *The human and political history of mirza Kan Temir and the juridical status of Moldavia, Wallachia and Crimea: vassal states of the Ottoman Empire*. Cossuto has published several volumes about the 'steppe peoples' of Eurasia – their history, culture and traditions – concentrating on analysing and interpreting the sources that refer to them. More information on the author's activities can be found on: <https://sites.google.com/site/pinocossuto/>.

Abstract

A 'steppe nomadic culture' vs a 'forest language': Modern identity dissonance in the history of the Magyars

The socio-economical culture of the Magyar Confederation of the Conquest Period (9th century AD) firmly belongs to the steppe nomadic system, but the language of the Magyars is classified as belonging to the Uralic language family. In this brief article, I shall trace a concise historical account of the origin of the theories affirming the 'identification' of the Magyars with other peoples of the past that practised a way of life different from that of the steppe peoples. However, the Magyars consistently claimed their affinity with the Huns, Turks and other populations that practised equestrian nomadism. I shall argue that, to understand and solve this contradiction, one has to look into the socio-economical world and way of life of these nomadic peoples, rather than just investigating one single aspect of their 'culture', that is, their language(s).

M.M. Jocelyne Fernandez-Vest

Has worked at several universities, such as: Professor of Scandinavian Languages and Literatures (Rennes II), Finno-Ugric Linguistics (Sorbonne Nouvelle-Paris III), General Linguistics (École Pratique des Hautes Études ((EPHE)-Sorbonne), and Centre National de la Recherche Scientifique (CNRS). She is head of the UMR OSTERLITS: Uralic Linguistics, Multilingualism and Translation theory. She is Professor *Emerita* since 2013 at Laboratoire LACITO (Langues et Civilisations à Tradition Orale), etc. She has numerous publications in various fields of linguistics, including: typology, cognitive semantics, Scandinavian and Finno-Ugric languages, pragmatics, 'information structure', 'orality', etc. For further information see: <http://lacito.vjf.cnrs.fr/membres/fernandez.htm>.

Abstract

Information Structuring and typology: Finnic and Samic word order revisited through the prism of orality

The present demonstration is a contribution to the Information Grammar of Orality theory. It is based on two types of constructions: 'Detachment Constructions' and 'Non Finite Constructions', and adopts an 'Information Structuring' methodology. Analyzing a Sami-Finnish translation of two current Non Finite Constructions, it turns out that these constructions have a higher frequency in Finnish than in Sami. The Information Structuring articulation of the Sami sentence is generally binary, whereas the synthetic structure of the Finnish sentence is enunciatively opaque: a temporal Non Finite Construction that functions as an adverbial clause rather than a separate Theme-clause seems to imply a cognitive reductionism typical of normalized language. This could partly explain the rarity of this 'Uralic' construction in Sami, a language whose register is still mainly oral.

Cinzia Franchi

Is associate professor of Hungarian Language & Literature at the Dpt. of 'Studi Linguistici e Letterari (DiSLI)' of the University of Padua, dealing both with Modern and Old Hungarian. Other fields of study are: the translation of Hungarian literature into Italian; the cultural relations between Hungary and Italy and between Hungary and Romania, in particular Transylvania. Since 2012, she regularly organizes the 'Seminario di Traduzione Ungherese' ('Hungarian translation seminar').

Among her translations of Hungarian texts, one can mention: *Lettere dalla Turchia* ('letters from Turkey), by Kelemen Mikes (Lithos, 2006). For the series *Leo*, she has published the monograph: *L'arancia ungherese. La letteratura in Ungheria negli anni Cinquanta* (2014) ('the Hungarian orange. The literature in Hungary in the 1950s). e-mail: cinzia.franchi@gmail.com

Juha Janhunen

Is since 1994 professor and chair of 'East Asian Languages and Cultures' at the University of Helsinki. He was trained in Finland, Hungary and Japan in the fields of Uralic and Altaic Studies. He has carried out linguistic field work in various parts of North-east Asia, including Siberia, Mongolia, Manchuria, and the Kuku Nor region of Tibet. Most recently, he has been working in the Amur-Sakhalin region of the Russian Far East. Apart from descriptive and comparative linguistics, his field of interests comprises issues of ethnic history, language ecology and linguistic endangerment. E-mail: asiemajeure@yahoo.com.

Abstract

Issues of comparative Uralic and Altaic Studies (4): On the origin of the Uralic comparative marker

The paper surveys the distribution, origin and typology of the comparative marker **-mpA* in the Uralic languages. This marker is attested in three branches of Uralic – Finnic, Samic and Hungarian – and the comparative forms thus marked are normally considered to be of a denominal origin. In view of Turkic parallels it is, however, possible to defend a deverbal origin instead. This idea was first presented by G.J. Ramstedt already a century ago, but it has been ignored by later scholars. New data from Turkic make it look even more probable that we are indeed dealing with an originally deverbal form.

László Marác

Is associate professor at the Department of European Studies of the University of Amsterdam, and vice-coordinator of the MIME-consortium (www.project-mime.org.) that won the European FP7-tender in 2013 under the reference of 'SSH Call 2013.5.2-1: *Multilingual Challenge for the European Citizens*'. In 2015, he has been awarded the title of 'honorary professor' at the 'L.N. Gumilyov Eurasian National University' in Astana, Kazakhstan.

He has published on a range of topics, including European studies, historical linguistics and cultural and linguistic diversity. His recent publications include: *Towards Eurasian Linguistic Isoglosses: the Case of Turkic and Hungarian*. Astana: International Turkic Academy (2016). See <http://www.uva.nl/en/profile/m/a/l.k.maracz/l.k.maracz.html>. for further information.

Abstract

Revisiting the theory of the Hungarian vs Chuvash lexical parallels

In this article I shall revisit the vexed question of the presence of numerous words of Turkic origin in the Hungarian language. According to received wisdom, these words have all been ‘borrowed’ into Hungarian – despite the fact that a significant percentage of them belongs to basic lexicon, typically resistant to borrowing. More precisely, this process of borrowing would have taken place almost entirely from one particular Turkic language, Chuvash, the only surviving languages belonging to the so-called West Old Turkic branch (as against East Old Turkic, or Common Turkic). This claim in turn is based on the observation that Hungarian and Chuvash, but not Common Turkic, share certain specific isoglosses, mainly rhotacism and lambdacism. However, this scenario is not as clear cut as typically depicted, and, consequently, the conventional model is unable to account for a number of counter examples. In this article I shall propose a different interpretation of the data, suggesting that rhotacism, lambdacism and other isoglosses are better accounted for if interpreted in terms of ‘Eurasian isoglosses’, going back to the Early Central Asian linguistic area.

Angela Marcantonio

Is associate professor of Linguistics, Finno-Ugric & Hungarian Studies, at Sapienza University of Rome. Her main fields of research are historical, descriptive and typological linguistics, applied to the Indo-European languages, the Finno-Ugric (/Uralic) languages, as well as the (assumed) linguistic connections between the respective proto-languages. In particular, she has done research on Finnish, Hungarian and the Ob-Ugric languages, as well as on the linguistic correlations between Hungarian and Turkic, research whose results are now published in the volume: *The nature of the Hungarian vs Turkic linguistic correlations: is Hungarian really a ‘proto-typical’ Uralic language?* (Monograph n. 16 of *Rivista di Studi Ungheresi*; Rome, 2017). For further information see: <http://sites.google.com/site/angelamarcantonio/Home>

Borbála Obrusánszky

Studied History and Mongolian at the Eötvös Loránd (ELTE) University in Budapest. She continued her studies on the Mongolian language and culture at the 'Mongolian State University' in Ulaanbaatar, where she defended her Ph.D. thesis (1999) on issues regarding the state organisation of Nomadic peoples. She is currently an 'Adjunct' at the 'Károli Gáspár Reformed Church University' in Budapest. She is the editor-in-chief a journal of history titled *Kelet Kapuja* ('the gate to the East'), and serves as president of the 'Association of the Ancient Hungarian History'. Several of her publications deal with the history and socio-military culture of the Hunnic peoples, both in Asia and Europe. Her main areas of research include the Eurasian nomadic society and the Mongolian language. E-mail: borbala.obrusanszky@gmail.com.

Abstract

Are the Hungarians Ugric?

In the 19th century the language family tree model was created, and several language families were established. However, there are instances where the validity of the tree model, and, consequently, the traditionally established families, may be called into question, especially if the peoples /languages classified as belonging to the same family have had no actual connection in any way, in any historical period, and, presumably, in any pre-historical period either. A case in question is the classification of Hungarian. On the basis of the relevant historical sources, including the ancient Hungarian *Chronicles*, I shall argue that, even if the Uralic proto-speech-community had existed, the conventional thesis of the belonging of the Hungarian language, and peoples, to this proto-community is not supported by the necessary extra-linguistic evidence, the linguistic evidence itself being rather wanting too.

Péter Pomozi

Is associate professor and member of the 'Doctoral School of Linguistics' at the Eötvös Loránd University (ELTE) of Budapest. He studied, among others, Hungarian and Italian language and literature and Finno-Ugric philology. His main research fields are: historical linguistics, sociolinguistics, language planning and the linguistics of folklore. He is member of the editorial board of several scientific journals of folklore and linguistics in Estonia and Hungary, editor of many book

series in Finno-Ugric languages and author of numerous scientific and popular publications of linguistics and folklore. Pomozi is also the promoter of a large exhibition-series on the linguistic archaism of Hungarian folk ballads, songs and prayers. E-mail: pomozi.peter@btk.elte.hu. Website: finnugor.elte.hu.

Abstract

The impossibility of the evolutionary metaphor: Neogrammarians, family trees and linguistic affinity

The historical, comparative method, and related practice of reconstruction, relies on the application of neogrammarian postulates, including the ‘sound laws’, principle that is, however, not supported by the evidence. The other major problem embedded in this method is the exaggerated role attributed to the process of reconstruction and to the ‘family-tree way of thinking’ – this in turn representing the embodiment of the a-critical application of the Darwinian metaphor to languages, despite the fact that this metaphor is not fully adequate in the latter case. In this essay I shall raise some still unanswered questions regarding the issue of the applicability, or otherwise, of the evolutionary metaphor to languages and the issue of the limits of reconstruction, whilst also highlighting some essential elements of the historical context within which the neogrammarian /evolutionary concept developed. Finally, I shall argue that, being the relations between people and languages highly intricate, only an interdisciplinary approach can produce meaningful results.

Natalia Serdobolskaya

Is associate professor at the ‘Russian State University for the Humanities’ and ‘Head of Laboratory at the Moscow State University of Education’. Her main areas of research are ‘differential object marking’ in Finno-Ugric languages and ‘complementation’. In addition to articles on differential object marking and various aspect of the syntax and semantics of complement clauses, the author is one of the editors of the volume *Finno-ugorskie yazyki: fragmenty grammaticheskogo opisaniya. Formal’nyi i funktsional’nyi podkhody* (‘Finno-Ugric languages: fragments of grammatical description. Formal and functional approaches’), Moscow, Yazyki slavyanskikh kul’tur (2012), and one of the authors of the on-line dictionary of the Beserman dialect of Udmurt (beserman.ru). e-mail: serdobolskaya@gmail.com.

Abstract

Differential Object marking in Eastern Mari and Permic: A look from the field
The phenomenon of differential object marking (DOM) is characteristic of the Uralic languages, and many scholars reconstruct for Proto-Uralic the model of DOM with definiteness as the main relevant factor. In this paper, I consider the models of DOM in two Permic languages and in Eastern Mari, arguing that the systems of DOM are very different in the modern languages, and in many of them the definiteness factor is relevant only in part. Even the Permic languages and the dialects of one and the same language develop different systems of DOM. This leads to the conclusion that specific semantic, morpho-syntactic and prosodic features of DOM must be considered as innovative features of the Uralic languages and their dialects.

Elisa Zanchetta

Graduated in *Modern Languages, Literatures and Cultures* (B.A.) at the University of Padua in 2012, defending a thesis entitled: *Finnische Lehnwörter aus den alten germanischen Sprachen*. In 2014 she obtained her M.A. degree in *European and American Languages and Literatures* at the University of Padua, specializing in German and Hungarian. She has translated and commented fourteen Hungarian *mitikus mesék* ('tales of magic'), taken from Benedek Elek's collection. She has published an article entitled *L'Aldilà boreale* ('The boreal other world') in the journal *Lingue e Letterature d'Oriente e d'Occidente* – the full text is available in the following link: <http://www.fupress.net/index.php/bsfm-lea/issue/view/1464>. e-mail: elisazanchetta110890@gmail.com.

Abstract

The 'impossible' Comparison between the Creatures and Figures of the Hungarian Folklore and the Finnish Mythology

Hungarian and Finnish are claimed to belong to the same linguistic family, but it seems that their mythological world hardly share any common traits, this being yet another contradiction embedded in the Uralic theory. Thus, a comparison of Uralic mythological themes may seem at first 'impossible', also because of the scanty Hungarian material. However, I shall argue that comparing at least Hungarian and Finnish mythology is not a 'nonsense' task, and that further research is necessary in order to try to 'reconstruct' ancient Hungarian folk believes. This, in turn, will contribute significantly to philological, anthropological and linguistic research in Uralic, and, possible, to the task of clarifying the position of Hungarian within the family.

COMITATO EDITORIALE
SAPIENZA UNIVERSITÀ EDITRICE

Coordinatore

GIUSEPPE CICCARONE

Membri

GAETANO AZZARITI
ANDREA BAIOCCHI
MAURIZIO DEL MONTE
GIUSEPPE FAMILIARI
VITTORIO LINGIARDI
CAMILLA MIGLIO

COMITATO SCIENTIFICO
SERIE PHILOLOGICA

Responsabili

VICENÇ BELTRAN, FRANCO D'INTINO, ARIANNA PUNZI (Roma, Sapienza)

Membri

FABIO FINOTTI (Pennsylvania)
LEONARDO FUNES (Buenos Aires)
SABINE E. KOESTERS GENSINI (Roma, Sapienza)
LUIGI MARINELLI (Roma, Sapienza)
SNEŽANA MILINKOVIC (Beograd)
RYSZARD NYCZ (UJ Cracovia)
JUAN PAREDES (Granada)
PAOLO TORTONESE (Paris III)
JAMES VIGUS (London, Queen Mary)
FABIO ZINELLI (Paris, Ecole pratique des hautes études)

COMITATO SCIENTIFICO
MACROAREA E

Coordinatrice

CAMILLA MIGLIO

Membri

VICENÇ BELTRAN
MASSIMO BIANCHI
ALBIO CESARE CASSIO
EMMA CONDELLO
FRANCO D'INTINO
GIAN LUCA GREGORI
ANTONIO IACOBINI
SABINE KOESTERS
EUGENIO LA ROCCA
ALESSANDRO LUPO
LUIGI MARINELLI
MATILDE MASTRANGELO
ARIANNA PUNZI
EMIDIO SPINELLI
STEFANO VELOTTI
CLAUDIO ZAMBIANCHI

Il Comitato editoriale assicura una valutazione trasparente e indipendente delle opere sottoponendole in forma anonima a due valutatori, anch'essi anonimi. Per ulteriori dettagli si rinvia al sito: www.editricesapienza.it

COLLANA CONVEGNI

Per informazioni sui precedenti volumi in collana, consultare il sito:
www.editricesapienza.it

20. Giorgio Bazzichelli. L'uomo e lo scienziato
Autori vari
21. Attuazione e sostenibilità del diritto alla salute
Roberto Nania
22. Épicurisme et Scepticisme
Stéphane Marchand & Francesco Verde
23. I musei di chimica e la chimica nei musei della scienza
Luigi Campanella e Valentina Domenici
24. Digital Humanities
Progetti italiani ed esperienze di convergenza multidisciplinare
Fabio Ciotti
25. Atti della Giornata in ricordo di Federico Caffè
Mario Tiberi
26. Information Technologies for Epigraphy and Cultural Heritage
Proceedings of the First EAGLE International Conference
Silvia Orlandi, Raffaella Santucci, Vittore Casarosa, Pietro Maria Liuzzo
27. Oltre i confini
Studi in onore di Giuseppe Burgio
*Raimondo Cagiano de Azevedo, Claudio Cecchi, Angela Magistro
Giorgio Milanetti, Giuseppe Sancetta, Donatella Strangio*
28. Novità nella ricerca archeologica a Veio
Dagli studi di John Ward-Perkins alle ultime scoperte
Roberta Cascino, Ugo Fusco, Christopher Smith
29. La letteratura e il male
Atti del Convegno di Francoforte, 7-8 febbraio 2014
Gianluca Cinelli e Patrizia Piredda
30. La Facoltà di Scienze dell'Università di Roma
dall'Unità alla prima guerra mondiale
Giornata di lavoro e discussione
Enrico Rogora
31. Paul Celan in Italia
Un percorso tra ricerca, arti e media 2007-2014
Atti del convegno (Roma, 27-28 gennaio 2014)
Diletta D'Eredità, Camilla Miglio, Francesca Zimarri

32. Per un *corpus* dei pavimenti di Roma e del Lazio
Atti della Giornata di Studi, Roma 24 novembre 2014
Sapienza Università di Roma
Claudia Angelelli e Stefano Tortorella
33. Europa Concentrica
Soggetti, città, istituzioni fra processi federativi e integrazione politica
dal XVIII al XXI secolo
Alessandro Guerra e Andrea Marchili
34. Edizioni Critiche Digitali/Digital Critical Editions
Edizioni a confronto/Comparing Editions
Paola Italia e Claudia Bonsi
35. La famiglia da concepire
Il benessere dei bambini e delle bambine con genitori gay e lesbiche
Roberto Baiocco, Nicola Carone, Vittorio Lingiardi
36. Digital and Traditional Epigraphy in Context
Proceedings of the EAGLE 2016 International Conference
Silvia Orlandi, Raffaella Santucci, Francesco Mambrini, Pietro Maria Liuzzo
37. I Filosofi del diritto alla 'Sapienza' tra le due Guerre
Atti del Convegno Internazionale Roma, 21 e 22 ottobre 2014
Gianpaolo Bartoli
38. Critica clandestina?
Studi letterari femministi in Italia
Atti del convegno - Sapienza Università di Roma, 3-4 dicembre 2015
Maria Serena Sapegno, Ilenia De Bernardis, Annalisa Perrotta
39. REMSHOA
l'Italia, la Shoah, la memoria. La deportazione degli ebrei in Grecia
Luca Micheletta
40. Human Nature
Anima, mente e corpo dall'antichità alle neuroscienze
Nunzio Allocca
41. The state of the art of Uralic studies: tradition vs innovation
Proceedings of the 'Padua Uralic seminar' - University of Padua,
November 11-12, 2016
Angela Marcantonio

This volume contains the Proceedings of the ‘Uralic Studies’ Seminar: *The State of the Art of Uralic Studies: Tradition vs Innovation*, held in Padua (Italy), November 12-13, 2016. The seminar was organized by the Department of ‘Studi Linguistici e Letterari’ of Padua University and the ‘Facoltà di Lettere e Filosofia’ of Sapienza University of Rome. The aim of the seminar, and of this volume, was/is to bring together linguists working on the Uralic languages from different perspectives, with the purpose of increasing the exchange of ideas and fostering mutual influences on each other field and methods of analysis. In addition to presenting the current ‘state of the art of Uralic studies’ – for specialists, general linguists and general public – the volume also addresses some issues related to the so-called ‘Ural-Altai theory’, nowadays often referred to as the ‘Ural-Altai linguistic belt, unique typological belt’. The contributors to the volume are renown scholars of Uralic, and also Altaic languages, from various European universities, such as Moscow, Helsinki, Paris, Budapest etc.

Angela Marcantonio is associate professor of Linguistics, Finno-Ugric & Hungarian Studies, at Sapienza University of Rome. Her main fields of research are historical, descriptive and typological linguistics, applied to the Indo-European languages, the Finno-Ugric (/Uralic) languages, and the (assumed) linguistic connections between the respective proto-languages. In particular, she has conducted research on Finnish, Hungarian and the Ob-Ugric languages, and the linguistic correlations between Hungarian and Turkic, now summarized in the volume: *The nature of the Hungarian vs Turkic linguistic correlations: is Hungarian really a ‘proto-typical’ Uralic language?* (Rivista di Studi Ungheresi 16 Rome, 2017).

ISBN 978-88-9377-066-8



9 788893 770668

