

# **Reconstructing Historical Sociolinguistics and the Typology of Loanwords:**

## **The Case of ERIC Loans in the Balkans**

**BRIAN D. JOSEPH**

The Ohio State University

**LAGB Summer School Day Lecture**

**6 September 2021**



## Main points for today:

- loanwords typically imply a social connection between peoples, via face-to-face interaction and contact that make the loans possible
- they thus offer a basis for looking back to earlier language and social states and of reconstructing the conditions under which the loans could have occurred
- but usual typologies for loanwords miss the full picture of the earlier social conditions
- drawing on collaborative work with Victor Friedman and data from the Balkans, I propose here a new class of loanwords that focuses on the conversational interactions that speakers had and discuss what such loans mean for historical sociolinguist(ic)s

# THE BALKANS



## **From a linguistic standpoint, what is notable about the Balkans:**

- **multilingualism** (note languages with a special affinity:  
**Albanian; Balkan Romance (Aromanian, Meglenoromanian, Romanian; Judezmo); Balkan Slavic (Bulgarian, Macedonian, Torlak Serbian); Balkan Turkish; Greek; Romani — as well as other smaller languages with deep(ish) roots in the area, e.g. Hungarian (in Romania), Circassian (in Kosova, until 1994), etc.**
- **language contact**

- sprachbund (also: “linguistic area”) = a geographic zone where languages, through intense and sustained contact in a mutually multilingual society, have come to converge with one another structurally and lexically and to diverge from the form that they held previously

- structural aspects of Balkan convergence have dominated the literature, e.g. **a volitionally based future tense, the use of finite subordination instead of infinitives, the merger of genitive and dative cases, an enclitic (postposed) definite article, the encoding in the verbal system of a distinction between real and presumed information-source (evidentiality)**, etc.
- but, given the structural focus of most work on the sprachbund, the lexicon has almost always been treated as somehow beside the point (and in any case, culturally based loans seem not to tell us much about the sprachbund as they are found in routine non-sprachbund contact situations)
- that view may be short-sighted, as the lexicon actually has much to offer

**Some traditional types of insights from loanwords in the lexicon:**

**--as to domains of interaction**

**e.g. Greek religious terms in Bulgarian and Romanian with spread of Christian orthodoxy**

**--as to historical phonology**

**e.g. chronology of sound changes, as with (Tosk) Albanian rhotacism (VnV → VrV, cf. \*woinos ‘wine’ > *verē*) that affects Latin loans (e.g. *inimicus* ‘enemy’ > *armik*) but not Slavic loans, and so can be dated to post-Roman, pre-Slavic contact period (4<sup>th</sup>-7<sup>th</sup> cent.)**

I go beyond such insights from loanwords here and instead use loanwords as a basis for reconstructing the social conditions of contact in the past.

**A key distinction in Balkan language contact (as shown through the lexicon):**

- **object-oriented** interaction (with a particular goal in mind, such as obtaining something (such as information), engaging in a commercial transaction (buying and selling), etc.)
- **human-oriented** interaction (no particular goal beyond the interaction itself, involving (generally) amicable interaction on a personal level)

**There is plenty of evidence of the former, but it is the latter that is interesting from the perspective of historical sociolinguistics.**

**Note that my focus here is on the nature of the surface material that passes between languages, on the loanwords themselves**

**--it is thus on the first element in various contact-related dichotomies in the literature:**

- **borrowing** vs. language shift (Thomason & Kaufman 1988)
- **matter borrowing** vs. pattern borrowing (Matras & Sakel 2007)
- **borrowing** vs. replication (Heine & Kuteva 2010, in their “contact grammaticalization” schema)

## Some standard typologies of loanwords

- i) focusing primarily on the FORM of the loan (cf. Winford 2003: 41-46):
  - Haugen (1950): **importation** vs. **substitution** (“based on the presence or absence of foreignness markers”)
  - Haugen (1953): **lexical borrowings** (“imitation of some aspect of the donor model) vs. **creations** (“entirely native [with] no counterpart in the donor language” (but based on some nonnative material))

**ii) focusing primarily on the CONTENT of the loan:**

- Bloomfield 1933 – **cultural borrowings** (those arising via the exchange, often mutual, of terminology between speakers of different languages, representing different cultures) vs. **intimate borrowings** (those not obviously linked to cultural objects)

iii) focusing primarily on the MOTIVATION for the loan:

- Hockett 1958 -- **need borrowings** (essentially Bloomfield's cultural type, though the motivation of “needing” a word for a (new) cultural item is at issue) vs. **prestige borrowings** (where the motivation is the “prestige” that the borrowing language speakers accord to material from the donor language)

## Some failings with these typologies:

- a. by focusing on form, Haugen's does not build in the social context for the loans
- b. the types listed above are not necessarily discrete – a cultural/need loan might be undertaken for reasons of (Hockettian) prestige or be associated with (Bloomfieldian) intimate contact
- c. noncultural/nonneed loans are not always a matter of prestige, at least not obviously so; e.g.: Albanian diminutive *-zə* in Megara Greek *Iyaza* ‘a little’

d. moreover, borrowing implies interaction between/among speakers, but ... “non-interactive” borrowing does occur, e.g.:

- learned borrowings, as with Latin into Romance, or Old Church Slavonic into Bulgarian and Russian, or even Modern Greek *katharevousa* (high-style, Ancient-Greek-based) words adopted into *dimotiki* (low-style, colloquial Modern Greek)

- constructed neologisms (as with lots of western medical and other technical terminology, generally with Greek or Latin roots, e.g. *encephalography* ‘imaging of the brain’, *electroencephalography* ‘electronic encephalography’, and *electroencephalographologist* ‘a specialist who studies electroencephalographs’)

## **What these typologies are missing:**

- **the full dynamics – the full historical sociolinguistic picture -- of the social environment in which the borrowing occurs especially as to the MEDIUM through which the borrowing takes place**
- **this is a particular concern for the Balkans, since the lexical side of the Balkan sprachbund is only one dimension to the contact-related effects, inasmuch as there is massive structural convergence evident too (in the form of the future tense, in the use of finite subordination instead of infinitives, in the merger of genitive and dative case, etc.)**

## **Motivating a New Type of Loanword (based on work with Victor Friedman):**

- Recognize a type of loan phenomenon which is consistent with what is known about contact in the Balkans, the contact that gave rise to the structural convergence (and thus revealing the sprachbund)**

## **Our hypothesis:**

- sprachbunds arise under conditions of **sustained, intense, intimate** contact among speakers, with **mutual, multi-directional, multi-generational multi-lingualism** (**our “4-M model”**)

**Note:** this is distinct from the “4-M Model” of Myers-Scotton & Jake 2000 and Jake & Myers-Scotton 2021, which refers to a typology of morphemes involved in codeswitching

That is, there is/are:

- **intimacy**, yes, but more than that
- **prestige**, yes, but more than that
- multiple **cultures** (and thus “need” in some sense), but more than that

**Therefore, we need: a loan type that ...**

- **corresponds to these conditions**

**and**

- **is consistent with sprachbund formation**

## **Our suggestion:**

- **base loan typology on the mutual interaction, specifically on conversational interaction, between/among speakers**

Therefore, we need to recognize a type of loans that are:

**Essentially Rooted In Conversation**

i.e.:

**E.R.I.C. loans**

These are loans that depend crucially on speaker-to-speaker interaction of an on-going and sustained kind, the sort of contact that can be characterized as intense and at the same time intimate, as opposed to occasional and casual, human-oriented rather than (solely) object-oriented.

## **Motivating the acronymic notion, formally and conceptually:**

- **formally, the acronym is a suitable homage to Eric P. Hamp, the dean of Balkan linguistics, a long-time champion of the study of language contact in general and in the Balkans in particular (and a mentor to both Friedman and myself)**
- **conceptually, the notion offers a way to distinguish between loans that take place under sprachbund-conducive conditions and those that take place under casual contact situations.**

- and, it can extend to some sorts of “borrowing” that is not lexical in nature, e.g. the diffusion of expressive phonology (as with the expressive value of [ts]/[tʃ] (and voiced counterparts) in the Balkans (Joseph 1984, Curtis 2008) or gestures (as with the upward head nod for ‘no’)

## To elaborate:

- **face-to-face interaction, of the sort that would necessarily have occurred under the intense and on-going contact among speakers in the Balkans, is essential for creating and propagating the structural convergences typically taken as diagnostic of a sprachbund**

- and, that certain kinds of loanwords occur in such a social milieu is a bonus of sorts, and means that the loan words can be both an *indicator* that contact conducive to the formation of a sprachbund is taking place and a *by-product* of such contact.
- that is, these are loans that tell us something about speaker contact and about the sociolinguistics and the socio-history of the region

- this view draws on the notion of “degrees of contact”, as recognized explicitly in the Thomason & Kaufman 1988 “scale of borrowability”, where the borrowing of different types of linguistic material is claimed to correlate with different levels of intensity of contact among speakers.

- our contention is that, consistent with this scale, certain types of loanwords, especially those embedded in discourse and in conversational use and those that go beyond simple exchange of information, correlate with the intense, sustained, and intimate contact that is necessary for the formation of a linguistic area with structural convergence, i.e. a sprachbund.

- these loans are “**sprachbund-consistent**”, since they represent those lexical elements that most directly reflect the sort of language contact that is consistent with the emergence of a sprachbund, namely sustained intense contact on a day-to-day basis in a multilingual milieu.
- many ERIC loans are members of **closed lexical classes**, representing vocabulary domains that are generally held to be somewhat resistant to borrowing, and yet they are borrowed

**A strikingly thorough-going case in the Balkans -- the entry of  
Turkish words into Macedonian (Friedman 1986):**

- virtually all categories of Macedonian lexical items, covering virtually all sectors of the vocabulary, have been affected by Turkish borrowings, but note that many fall into ERIC classes:

“The large number of Turkish lexical borrowings belong to all levels of vocabulary and almost all parts of speech, e.g. *džeb* ‘n. pocket’ (*ceb*), *bendisa* ‘v. please’ (*beğen-*), *taze* ‘adj. fresh’ (*taze*), *badijala* ‘adv. for nothing’ (*bâdihava*), *ama* ‘conj. but’ (*amma*), *karşı* ‘prep. opposite’ (*karşı*), *ič* ‘pron. nothing’ (*hiç*), *sikter* ‘excl./interj. scram’ (*siktir*), *keški* ‘part. if only’ (*keşke*). The only Macedonian traditional part of speech lacking Turkisms is the numeral, although there are Turkisms in numerical expressions, e.g. *čerek* ‘quarter’, and Turkish numerals in other parts of speech, e.g. *bešlik* ‘five-grosch silver coin’ ... Turkish vocabulary has penetrated every facet of Macedonian life: urban and rural, e.g. *duk’an*, ‘shop’ (*dükkân*), *sokak* ‘street, alley’ (*sokak*), *ambar* ‘barn’ (*hambar*), *endek* ‘ditch, furrow’ (*hendek*); man-made and natural, e.g. *tavan* ‘ceiling’ (*tavan*), *šiše* ‘bottle’ (*şîşe*), *zumbul* ‘hyacinth’ (*zümbül*), *taftabita* ‘bedbug’ (*tahtabiti*); intimate and abstract, e.g. *džiger* ‘liver, lungs’ (*ciger*), *badžanak* ‘brother-in-law (wife’s sister’s husband)’ (*bacanak*), *rezil* ‘disgrace’ (*rezil*), *muabet* ‘conversation’ (*muhabbet*).”

**Classes of ERIC loans (with examples to follow from all around the Balkans), going beyond Friedman's Turkish → Macedonian instances**

- Kinship terms
- Numerals
- Pronouns
- Adpositions
- Negatives
- Complementizers
- Discourse elements (interrogatives, connectives, attitudinal expressives, interjections, gestures)

- Vocatives
  - Onomatopoeia
  - Reduplication (especially of an expressive nature)
  - Expressive phonology
  - Idioms (and phraseology more generally, even shared proverbs)
- • •

## Examples of Loans in these Classes (a sampling, not exhaustive)

- **Kinship terms**

Turkish *baba* ‘father’ → Albanian, Aromanian, Greek *baba* ‘father’ (and labeled an archaism in contemporary Macedonian and Bulgarian sources, suggesting that it was in wider use in earlier times)

Greek *patera* ‘father’ → Aromanian (in Greece) *patera*

Turkish *nene* ‘mother’ → Albanian *nënë*, Greek *nené*

Turkish *dayı* ‘maternal uncle’ → Alb *dajë*, Mac *daja* (and Grk *daís*, with semantic shift to ‘bully’)

Greek *tsatsa* ‘aunt’ → Aromanian *țăță* (surely connected to Turkish *çaça* ‘woman who keeps a brothel’ but that is said to be from Greek)

- **Numerals**

**Romany borrowing of '7'/'8'/'9' (from Greek)**

‘11’ – ‘19’ as ‘DIGIT-on-TEN’, e.g. Albanian *pesë-mbë-dhjetë*  
‘15’ (widely cited in the literature as Slavic in origin but  
Albanian and Romanian disagree with Slavic in the gender of  
‘ten’ in various formations, so this pattern may have a  
different origin and may in fact reflect some very early  
Albanian (“Albanoid”) input (cf. Hamp 1992))

Turkish numerals in Balkan languages (e.g. in Pomak (Balkan  
Slavic) in Greece today, digits for ‘five’ and above are all  
Turkish).

- **Pronouns**

Turkish *hiç* → Mac *ič*, Alb *hiç*, Jud *hič*, Rmi *hič*, Arom *hiçi*,  
Romn *hici* ‘nothing’

Greek *mu* ‘my’ → Aromanian *-m* (vs. native *-n'i* (from Latin  
*mihi*, presumably via \**mnihi*)

Turkish *bu* ‘this’, *o* ‘that’, *kim* ‘who?’ → Ottoman-era  
Adrianople Greek *bu kim* ‘who is this?’, *kim o* ‘who is that?’

Turkish *kim* ‘who?’ → Bulgarian *kimi...kimi* ‘some...others’ and  
in combination with native interrogative pronouns, e.g. *kim koj*  
‘someone’

Greek *kaθénas* ‘each one’ → Ayia Varvara Romani

- **Adpositions**

Turkish *karşı* (*karşı* in Balkan Turkish) 'opposite' → Albanian *karshi* (also in Aromanian, Bulgarian, Ottoman-era Adrianople Greek, and Macedonian (in Arom & Grk as a postposition)

Turkish *gibi* 'like' (postposition) → Ottoman-era Adrianople Greek *gibi* 'like' (postposition)

Greek *me* 'with', if a borrowing from Albanian *me*, since derivation of Greek from Ancient Greek *metá* is problematic (requiring several ad hoc assumptions)

- **Negatives**

Greek *oxi* ‘no’ → Southern Aromanian *ohi* (Vrabie 2000)

Turkish *hiç* ‘mere trifle; nothing’ → Ottoman-era Adrianople  
Greek *hitʃ* ‘never; (not) at all’

Trk *ne...ne* ‘neither...nor’ → Agia Varvara Romani

Trk *hayır* ‘no!’ → Agia Varvara Romani

Greek *ba* ‘ah well; unh unh; no way’ → Arom, Romn *ba*

Blg *njama* ‘there is not’ → Wallachian Romn *neam* ‘not at all’

Trk *yok* ‘there is not’ → Grk, Alb, Balkan Romance, Balkan  
Slavic emphatic negative [jok] (“No way!”), and via calquing is  
basis for Alb (of Tetovo in Macedonia) use of *nuk* ‘not’ as  
general negative utterance and negative existential

**Greek *mi* ‘prohibitive negator’ → Vardar Macedonian,  
Bulgarian, and Aromanian *mi***

**Spread as a syntactic borrowing, note spread of one-word  
prohibitive use of prohibitive negator: Grk *Mi!* Alb *Mos!*,  
Romani *Ma!* (and extension to Slavic *nemoj/nim*) ‘Don’t!’**

**Upward head nod for negativity (found at least in Greek, Romanian (older generations), and Turkish) — Greek is a likely source, given what is known about Ancient Greek gestures and the fact that the distribution especially in Italy coincides with geographic limits of Magna Graecia)**

**Cf. Salmons 1990: “the wholesale adoption of English discourse markers in Texas German as part of the overall convergence of communication patterns, including gestures”**

**Also: ingressive voiceless dental affricate or alveolar click (the clucking noise accompanying head nod) — conventionally spelled *tsk* in English, *cq* in Albanian, *ck* in Macedonian and Bulgarian, *τσουχ* in Greek, *țăț* in Romanian, and *cik* (rarely *çik*) in Turkish, but all phonetically [t̪]**

‘enough’ + VERB = ‘do not VERB’ in northern Greece, in Kastoria region:

- Greek: *ftan* ‘is enough’ → *ftan’ krents* ‘you spoke enough, don’t speak’
- Slavic: adverb *dosta* ‘enough’ → *dosta zborvi* ‘enough speaking, do not speak’
- Aromanian: *duri/dure* (from Trk *dur* ‘stop’ → *duri zburets* ‘stop speaking, do not speak’.  
(= shared (calqued) phraseology in the domain of negation that is somewhat grammatical in nature)

- **Complementizers**

Greek *(h)oti* ‘that’ → Macedonian *oti* ‘that’

Greek *(h)o, ti* ‘for that reason’ → Macedonian/Bulgarian *oti*  
‘because’

Turkish *zira* ‘because’ (from Persian) → Bulgarian *zerem*

- **Discourse elements (interrogatives, connectives, attitudinal expressives, interjections, gestures)**

**NB:** These are ERIC forms, since they serve as the "glue" of everyday interactions between people, and include markers of common discourse strategies (e.g. questioning) as well as frequent discourse “particles” and indicators of an individual's status relative to other interlocutors (e.g. solidarity, distancing, etc.) but also those that serve a purely expressive purpose (i.e., elements that add "color" to conversation):

## INTERROGATION

Trk *mi/mü/mi/mu* ‘(focalizing) marker for yes-no questions’ →  
Ottoman-era Adrianople Greek *mu* ‘marker for yes-no  
questions’

- round back harmonic form borrowed (adopted without adaptation)
- postpositive in OAG, as in Trk, though slightly different (phrase-final in OAG, post-focus in Trk)
- has other, more discourse-based functions, in OAG: “it offers many nuances that accompany interrogativity : irony, doubt, challenging, swearing (an oath)” (Ronzevalle 1911)

--examples:

burís mu  
can/2sg QN  
‘Can you (do it)?’  
(Standard Greek: *borís*;

θa ‘rt’s mu  
FUT come/2sg QN  
‘Will you come?’  
(Standard Greek: *θa ‘rθis*;

## ATTITUDES / AFFIRMATION

provincial Turkish (nonharmonic) *de* → Greek *de* (signaling impatience), Albanian *de* (emphatic with imperatives), Macedonian *de* ‘c’mon’ (or possibly of Slavic origin, cf. verbal root *dě-* ‘put’ (or ‘say’?)

Slavic *da* ‘yes’ → Bitola Aromanian *da*

Albanian *po* ‘yes’ → Aromanian (dialectal) *po*

Greek *malista* ‘yes (indeed)’ → Aromanian (dialectal) *malista*

## CONNECTIVES

Mac *i* ‘and’ → Aromanian (in Macedonia)

Trk *hem* ‘and, too, and yet’ → OAG *hem* ‘and; too’

Trk *hem ... hem* ‘both ... and’ → Alb, Blg, Mac, Romani

Trk *ya ... ya* ‘either ... or’ → Alb, Armn, Blg, Grk, Mac, Romani

Mac *ili* ‘or’ → Trk (dialect spoken in Macedonia)

*ama/ami/ma/mi* ‘but (and various adversative nuances of meaning’ (whatever its origin, it has spread all over the Balkans, found in each of the languages; cf. Fielder 2010)

Trk *ancak* ‘but, on the other hand’ → Alb (*anxhak*), Armn (*anđeac*, though in the meaning ‘almost, finally’), Blg (*andžak*)

Trk *illâ ve lâkin* ‘but on the other hand’ → Alb (*velakin*), Blg (*illja veljakim / illjakim*), Armn (*eleafim / ileafim*)

Trk *me(đe)r* ‘but; however’ → Blg *meger / mer*

Trk **acaba** ‘I wonder if; oh indeed!’ Blg *adžaba/adžeba* ‘I wonder; is it so?’, Mac *adžaba*, Grk (of Edirne) *adžiba*, Armn *hağiba/agęeba* ‘is it so?’, Megl *adžaba*

Grk **aliθeia** ‘truly? really?’ Armn *aliθea/alihea*

Trk **alis** ‘really, indeed, actually’ Blg *alis*

Trk **artık** ‘now; well then; not anymore’ Blg *artăk* ‘finally; really; in fact’, Armn *artic* ‘finally’, Grk (of Edirne)

Trk **bari(m)** ‘at least; for once’ Alb *bar/bare(m)/bari*, Armn *báre/bári/bárim*, Blg (dialectal) *bar/bare/bárem/barém/bári/bári/bárim/barím*, DRmn *barem*, Grk (dialectal) *barim*, Rmni *barem*

Trk **belki(m)** ‘perhaps, maybe’ Alb *belqim*, Armn *belchi*, Blg *belki* ‘maybe; probably; as if’, Grk *belki(m)* (dialectal)

Trk **değil mi** ‘isn’t it so?’ Alb *dilmi* ‘since; because; after’, Armn *delme* ‘since’, Blg (dialect) *delmi/dilmi/dilma* ‘isn’t it?’

Trk **elbet(te)** ‘certainly, surely’

Alb *(h)elbet(e)*, Armn *elbet(e)* ‘possibly; assuredly’, Blg *(x)elbetel/elbetta/xelbette/xelbet(t)ja*, DRmn *(h)elbet*

Trk **galiba** ‘probably, presumably’

Alb *galiba* ‘perhaps’, Blg *galiba*, Grk (of Edirne) *galiba*

Trk **gerçek** ‘real; really, in truth’

Blg (dialectal) *gerček*

Trk **gûya** ‘as if; supposedly’

Alb *gjoja/gjyja*, Armn *ghio(i)a/ghio(i)am/ghioae*, Blg (+ dialect) *g'óa/g'oé/gjúva/g'oј/g'óve*, Grk (of Edirne) *γ'a*

Trk **korsan/m** ‘as if; supposedly’  
**korse(m)**

Alb *korsëm/korrse* (dialectal), Blg *korsem/kórsa/kórsak/kórsum*

Trk **sahi** ‘really, truly’

Alb *sahi*, Armn *saí* ‘exact’, Blg *saí* (usually followed by the Turkish interrogative particle *mi* to render ‘Really?’)

Trk **samsahi** ‘really really’  
(intensive reduplication of *sahi*)

Blg *samsai* ‘obviously; indeed’

Trk **sanki(m)** ‘as if’

Arom *sanchi*, Blg & Mac *sankim* ‘actually; that is to say; as if’, DRmn *sanche/i*, Grk (of Edirne) *sangim*

Trk **sözde** ‘so-called; supposed(ly), as if’

Blg *sjuzde* ‘supposedly (indicating disbelief)’; Grk (of Adrianoupolis) *seüzde*

Trk **yani** ‘that is to say’

Alb *ani* ‘well now!; never mind’, Grk (of Edirne) *γ'a'ni*

Trk **zahir** ‘apparently, clearly, evidently’

Armn *zări* ‘evidently, certainly, therefore’; Blg *zer* (particle reinforcing affirmation)

Trk **zaten** ‘essentially; already’

Blg (dialectal) *zata(n)* ‘indeed, really’; Jud *zatén* ‘indeed’ (Bunis 1999: 434); Grk (of Edirne) *zatin* ‘naturally; also’

Alb **mbase** ‘perhaps; maybe’

Grk (*m*)*bas (ke)* ‘perhaps’

Grk **táxa** ‘as if’

Armn *taha*

## OTHER SENTENCE ADVERBS / MITIGATING MODIFIERS

Trk **bile** ‘even; already’ Alb **bile** ‘even; in fact’, Armn **bile**, Blg **biljá(m)/bilé(m)**, Rmni **bila(m)/bilim**

Trk **hemen** ‘almost, nearly’ Blg **hemen**; Grk (of Edirne) **emén**

Trk **sade** ‘only’ Alb **sade**, Armn **sade**, Blg **sa(a)dé**, DRom Grk (of Edirne) **sadé**, Rmni **sáde/sadé**

Trk **salt** ‘only’ Alb **sall(a)/sallde/sallte** (dialectal), Blg **sal/sált(e)**, MRom **sal/săl**, Rmni **saltə'**

Trk **tamam** ‘just right; there you have it!’ Alb **tamam/taman**, Armn **tamam/tamamá/tamamaná**, Blg **tamám/tamán**, DRom (dialectal) **taman**, Grk (dialectal) **tamam(i)**, MRom **tamam**, Rmni **tamami**

Rmn **mai** almost Blg **mai** (Banfi 1985: 100)

**‘unceremonious term of address’ (lots of variants, almost all ultimately from Greek, where dialectally some 55 different forms are found, almost all stemming from AncGrk *mōré* (vocative of ‘foolish’), cf. Joseph 1997):**

**Turkish:**      *bre, bire, be*

**Albanian:**      *o, ore, or, mor, more, moj, ori, mori, moré, mre, voré, bre*

**Romanian:**      *bre, mă, măi*

**Bulgarian:**      *more, mori, bre*

**Macedonian:**      *more, mori, bre*

**Serbian:**      *more, mori, bre*

**Judezmo:**      *bre*

**Greek:**      *moré, bre, vre, re, aré, maré, mari, oré, voré, etc.*

## EXHORTATIVES

Turkish *(h)ay de* → Romanian: *haide*, Serbian *hajde*, Albanian *hajde*, Greek *aide* 'c'mon'

Greek *ela* 'c'mon' → Bulgarian *ela*

Trk *ha* 'a call to action' → OAG *ha*

## GREETINGS

Trk *merhaba* ‘hello’ → Blg *maraba*, OAG *merhaba*

Grk *geia sou* ([ja su]) ‘hello’ (lit.: “health to-you”) → Armn

Trk *uçurlar olsun* ‘good luck! good journey’ (literally “good-  
omens may-there-be”) → OAG *urular olsun* ‘goodbye’

Trk *oğlu-m* ‘son-my’ (term of endearment) → OAG *oylum*

Trk *can-im* ‘soul-my’ (term of endearment) → OAG *džanım*

‘bird’ in terms of endearment (direction unclear): Alb *zogu im*,  
Grk *poulí mu* (both “bird my”), Blg *pilence*, Armn *puiłǔ*,  
DRom *puiule* (all diminutives, “little chick”).

- **Vocatives**

**Blg preservation of VOC versus loss of all other cases (VOC as quintessentially conversational “case”), and note too coincidence of Grk VOC.SG of MASC *o*-stem nouns, e.g. *Petre* ‘(O) Petros!’ with Blg VOC.SG of MASC *o*-stem nouns, e.g. *Ivane* ‘(O) Ivan’ and the role that may have played in Blg VOC-preservation**

**Albanian *biro* ‘O son!', *Agimo* ‘O Agim!', with *-o* apparently from Slavic vocative**

**Romanian vocative endings influenced by Slavic, cf. especially *-e-le* (FEM.PL), apparently with Slavic vocative particle *-le***

- **Onomatopoeia** (*vel sim.*)

‘dog noise’: Albanian *ham-ham*, Romanian *ham*, Greek *γανγαν*,  
Turkish *hav hav*, Macedonian *av av*, Romani *hau-hau*,  
Bulgarian *bau-bau*

‘knocking’: Albanian *tak-tak*, Greek *tak tak*

‘noise for attracting a cat’: Greek *ψι ψι ψι*; Bulgarian,  
Romanian *ps ps ps*

- **Reduplication** (especially of an expressive nature)

Turkish affective *m*-reduplication (e.g. *kitap-mitap* ‘books and such’) → various languages, e.g.:

Alb *shiri-miri* ‘confusion’

*cingra-mingra* ‘trivia’

*çikla-mikla* ‘tiny bits and pieces; crumbs; trivia’

Blg *knigi-migi* ‘books and such’

*xăr-măr* ‘grumbling, tiff, discord, disagreement’ (Trk *hir mir* ‘quarrel’)

Mac *kal-mal* 'mud or whatever'

Rmn *ciri-miri* ‘confusion’

Jud *livro mivro* ‘books and such’  
*sapatos mapatos* ‘shoes, shmoes’ (Bunis 1999)

Grk *tzantzala mantzala* ‘rags and such’  
*ta sandala ke ta mandala* ‘stuff and things’ (OAG)  
*surðu murðu* ‘topsy-turvy’ (Levkas dialect)  
*i sara ke i mara* ‘Tom, Dick, and Harry; ragtail and bobtail’  
*ara mara* ‘who cares?’  
*ares mares (kukunares)* ‘nonsense’

**Turkish-style CVC- intensive prefixing (e.g. *beyaz* ‘white’ / *bem-beyaz* ‘pure white’):**

**Greek *tsiplakis* ‘naked’ / *tsir-tsiplakis* ‘stark naked’**

**Bulgarian *baška* ‘other / *bambaška* ‘very particular’  
*gol-goleničák* ‘stark naked’**

**Macedonian *gol-goleničok* 'stark naked'**

**Romanian *gol-goluț* 'stark naked'**

**BCS *go-golest* 'stark naked'**

- Expressive phonology
  - one or more of [ts]/[tʃ] and voiced counterparts [dz]/[dʒ] can be shown to have special affective ("allolinguistic") value in Greek (Joseph 1984) and Albanian (Curtis 2008), as measured partly by their *lexical distribution*, found primarily in onomatopoeia, in highly connotative words (that add “color” to conversation), and in borrowings (thus outside the native system per se)

- Turkish <ç> (= [tʃ]) also figures in onomatopoeia and other such “evocative” words (Marchand 1953)
- thus, the development of the Greek and Albanian expressive sounds may well have been influenced by Turkish here (and cf. the generally low-stylistic value accorded Turkisms (Kazazis 1972))

- **Idioms** (and **phraseology** more generally, even shared proverbs)

Friedman 1986 on Macedonian:

“the use of *jade* 'eat' to mean 'be subjected to' as in *jade k’otek* 'get a beating' (literally 'eat a blow') is obviously based on Turkish *kötek yemek* 'ibid.', and so *jade dožd* 'get soaked' ('eat rain', Tk *yağmur yemek*) or *jade gomno* 'say something stupid/embarrassing' ('eat excrement', Tk *bok yemek*, cf. English *to put one's foot in it*) can likewise be identified as calques.”

--Various everyday expressions (directionality unclear):

Greek *pos ise? (ime) kala* ‘how are-you? I-am well’ (thus, ‘be’ with ADVERB)

= Albanian *si je? (jam) mirë* (*mirë* as adverb with ‘be’, not adjectival form),

= Macedonian *kako si? dobro* (adverbial form)

= Bulgarian *kak si? dobre* (adverbial form)

= Romani *sar sijan? shukar*

= Turkish *nasılsın? iyi.*

Alb. *sa bën* ‘how much does it cost?’ (literally “how-much does-it-make?”) = Greek *poso kani* ‘idem’

Greek *to ksero ap' ekso* 'I know it by heart' (lit: "it I-know from outside") = Agia Varvara Romany (Messing 1988: 61) *dzanav-les avral* (*avral*= 'from outside, from abroad')

**Note too that there are dozens of shared proverbs, e.g., for the need to economize and plan for “rainy days”, one can say (more or less) “(Save) white money for a black day”:**

- Alb    ruaj paran e bardhë për ditë të zezë (= “preserve money white for day black”)
- Aro    bani albi pentru zile negre        (= “money white for days black”)
- Blg    beli pari za černi dni (= “white money for black days”)
- Rmn stringe bani albi pentru zile negre da (= “gather money white for days black”)
- Grk    t' áspra ja tes mávres méres (= “the white (things, i.e. coins) for the black days”)
- Mac    beli pari za crni denovi    (= “white money for black days”)
- Trk    ak akça kara gün içindir (= “white silver-coin black day for-is”)
- Rmi pharne pares mižinav kales dives resav (= “white money I.conceal, (for a) black day I.grasp”)

## Others:

“Not all flies make honey” (for expressing contempt for the lazy and for those who shirk duties)

“If a baby does not cry, it will not suck (i.e., nurse)” (for *The squeaky wheel gets the grease*)

“Water sleeps but an enemy does not” (for *Still waters run deep*)

These are a basis for viewing the region as a “Kultur-bund” (culture zone), consistent with human-oriented interaction

**One final clearly conversational example (also overlapping with “culture zone” notion (Joseph 2011)):**

- an expression used by some Greeks (maybe obsolete now but it did exist) in the game of “peek-a-boo” that adults and older children play with young children:

*buli buli buli buli ... dza*

- the *buli* part (repeated four times) comes when the face is covered up by one’s hands and the *dza* part comes when the hands open up to reveal the face.

## Where does this come from?

- in Greek, *buli* is just a nonsense word
- *dza* (with variant *tsa*)= an interjection marking surprise or indicating something like “here I am (somewhat unexpectedly)”

## But where does that *dza* come from?

- the only dictionary to comment on its source says it is a “nursery word”, and its use and form do make sense in that regard, based on Greek [ts/dz]’s special expressive (“*allolinguistic*”) status – a marked functional status for linguistic elements involving being on the margins of “core” information-oriented communication (Joseph 1982, 1984, 1994 and elsewhere).
- still, even with an *allolinguistic* rationale for *tsa/dza*, one has to ask why, if the word has a nursery-related origin, it has the particular form that it does.

- But, there is a compelling source for *tsa/dza* from a language outside of Greek:
- Albanian has an interjectional word spelled < xa > (phonetically [dza]) that, as listed in Mann 1948, has a meaning ‘here you are’
  - this presentational meaning fits in well in the game of peek-a-boo as the sound that accompanies revealing of the face

--so, if this source of the Greek utterance is accepted, this form would have entered Greek either through direct contact with Albanians, e.g. in northwestern Greece, or through contact with Arvanitika, the Albanian dialects spoken mainly in Central Greece, Attica, parts of the Peloponnesos, etc.

## But what about the rest, the *buli* part?

- that too has a compelling source in Albanian:
  - [buli] can be related to the Albanian verb *mbyll* ‘close, shut’ (3sg past)
- thus, the phrase in its etymological meaning would be:
  - “[when the hands cover the face] (it-has-)closed, (it-has-)closed, (it-has-)closed, (it-has-)closed ... [when the hands open up to reveal the face] Here-it-is!”.

**Phonetically the Greek [buli] would be an expected rendering of Albanian *mbyll*:**

- the initial ***b*-** for Albanian [mb] conforms to prevailing Greek phonotactics at the likely time of borrowing (with voiced stops without a nasal “prop” being allowed in word-initial position)
- so also Greek [ l ] for Albanian “dark” <ll>
- the [ u ] for the Albanian front rounded [ y ] vowel (Arvanitika has [ i ] for general Tosk [ y ] but i > u in labial + sonorant contexts occurs sporadically across all of Greek.)

**This would have been borrowed only through use, presumably by Arvanitika speakers using it to play with Greek babies (and thus made available to Greek speakers) or by Greeks over-hearing Arvanitika speakers using it with their own babies.**

**Either way, the conversational element and ‘intimate’ contact are paramount here.**

**We can thus conclude, based on these ERIC loans, that peoples in the Balkans in the Ottoman era (earlier too, but the Ottoman era was formative for the sprachbund) were talking to and interacting with one another, and doing so under human-oriented conditions.**

**These loanwords thus offer us a glimpse into the social conditions in the Balkans some 500 years ago, thus fulfilling a goal of historical sociolinguistics.**

**Moreover, this reconstruction is confirmed by the evidence of documented interaction conditions in bilingual villages and sectors in towns and cities in the 19<sup>th</sup> century and first half of the 20<sup>th</sup> century, e.g.:**

- Aromanian-Greek bilingualism in Metsovo (Récatas 1934, updated by Koufougiorgou 2004; Capidan 1940)
- Bulgarian-Turkish bilingualism in eastern Bulgaria (Mirčev 1952, Stojanov 1952)
- Greek-Turkish bilingualism in Ottoman-era Edirne (Ronzevalle 1911)

Еυχαριστώ

Falemnderit

Благодарјам

Благодарю

Mulțumesc

Teşekkür edirim

•

•

•