Journal of Language Contact

Journal of Language Contact

Evolution of Languages, Contact and Discourse

Instructions for Authors

Please make sure to consult the Instructions to Authors prior to submission to ensure your submission is formatted correctly. These instructions can be downloaded from the journal homepage at brill.com/jlc.

The JLC submission guidelines cover the following items:

1. Aims and Scope; 2. Submission of articles; 3. Originality of the submitted work ;4. General instructions for submission; 5. Language of contributions; 6. Transliteration and fonts; 7. Article organization and style; 8. References; 9. Illustrative matter; 10. Abstract and keywords; 11. Proofs; 12. Consent to publish; 13. Off-prints

Appendix: 'The Leipzig Glossing Rules: Conventions for interlinear morpheme-by-morpheme glosses'

Contact Information

Please send any inquiries regarding manuscript submission to: Robert Nicolaï, nicolai@unice.fr.

Books for review, book reviews and offers of review articles should be sent to: Robert Nicolaï; 7, rue Neuve; F-06510 Le Broc; France.

Open Access

This is an open access journal distributed under the terms of the CC-BY-NC license at the time of publication, which permits any non-commercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited. This journal is fully sponsored by Knowledge Unlatched, therefore the Article Publication Charge is waived. The author(s) sign(s) a special Brill Open Consent to Publish. More information on Brill's policy on Open Access can be found on brill.com/brillopen. The Brill Open Consent to Publish will be sent by the responsible production editor.

BRILL

In and Around the Balkans: Romance Languages and the Making of Layered Languages

Francesco Gardani
Professor of Romance Linguistics, Institute of Romance Studies, Faculty of
Arts and Social Sciences, University of Zurich, Zurich, Switzerland
francesco.gardani@uzh.ch

Michele Loporcaro
Full Professor of Romance Linguistics, Institute of Romance Studies, Faculty
of Arts and Social Sciences, University of Zurich, Zurich, Switzerland
loporcar@rom.uzh.ch

Alberto Giudici
Assistant in Romance Linguistics, Institute of Romance Studies, Faculty of
Arts and Social Sciences, University of Zurich, Zurich, Switzerland
alberto.giudici@uzh.ch

Abstract

The languages of the Balkans are a rich source of data on contact-induced language change. The result of a centuries long process of lexical and structural convergence has been referred to as a 'sprachbund'. While widely applied, this notion has, however, increasingly been questioned with respect to its usefulness. Addressing the linguistic makeup of the Balkan languages, the notion of sprachbund is critically assessed. It is shown that a) the Balkan languages and the Balkan linguistic exclaves (Albanian and Greek spoken on the Italian peninsula) share similar contact-induced phenomena, and b) the principal processes underlying the development of the Balkan languages are borrowing and reanalysis, two fundamental and general mechanisms of language change.

Keywords

1   Introduction

The languages of the Balkans – principally including the Indo-European (sub)branches of Albanian, Greek, (South) Slavic and (Eastern) Romance – are a gold mine when it comes to instances of contact-induced language change; they are a paramount example of linguistic and socio-cultural processes which over the centuries have led to high levels of lexical and structural convergence. As is well known, the result of this process of convergence has been referred to as a 'sprachbund', a term, coined by Trubetzkoy (1923) (see Section 2.2), that both has been present in the literature for approximately one hundred years and has been applied to a number of cases of linguistic convergence cross-linguistically (Campbell, 2017: 20–22). However, the notion of sprachbund (also known as 'linguistic area' and 'convergence area') has increasingly been questioned with respect to its usefulness.

The present paper introduces a special issue of the *Journal of Language Contact* focusing on Romance languages as they participate in the making of layered languages in the Balkans; the scope of this issue, however, extends beyond the geographic boundaries of the Balkans as it includes Albanian and Greek linguistic exclaves on the Italian peninsula, in contact with Italo-Romance varieties, as well as the Greek dialects Heptanesian, Pontic, Aivaliot and Cypriot in contact with Turkish and varieties of Romance.[1] In this introductory paper, we shall address the linguistic makeup of the Balkan languages, focusing on the characteristics that led linguists to develop the idea of linguistic area, and summarize critical claims by Dahl (2001), Stolz (2002), Campbell (2006; 2017), among others, on the notion of sprachbund. More specifically, we will show that a) the Balkan languages and the Balkan linguistic exclaves share similar contact-induced phenomena; and b) the principal processes underlying the development of the Balkan languages are borrowing and reanalysis, two fundamental and general mechanisms of language change.

The paper is organized as follows: Section 2 sketches the linguistic makeup of the Balkans (Section 2.1) and provides a critical assessment of the notion of sprachbund (Section 2.2). Section 3 discusses the occurrence of the same phenomena generally considered key properties of the Balkan languages, outside the Balkans. Section 4 summarizes the seven contributions to the special issue.

---

1   Into the bargain, the (northern) geographic boundary itself of the Balkan has been a matter of unrelenting debate (cf., e.g., Reed, Kryštufek and Eastwood, 2004: 9–10; Vezenkov, 2017: 116–119).

2   The Balkans

The Balkan peninsula (Fig. 1) is the land mass bounded by the Black Sea to the east, by the Sea of Marmara, the Aegean, Mediterranean, Ionian Seas to the south, by the Adriatic Sea to the west and by the rivers Danube and Sava to the north (but see fn. 1). Here, four distinct subgroups and several varieties of the Indo-European language family coexist: (i) Albanian, (ii) Greek, (iii) the South Slavic languages Bulgarian, Macedonian, and some dialects of the Bosnian-Croatian-Montenegrin-Serbian complex, (iv) the Eastern Romance languages Romanian, Istro-Romanian (spoken in Istria),[2] Megleno-Romanian (spoken in a small area in northern Greece and the Republic of North Macedonia) and Aromanian (spoken in northern and central Greece, southern Albania, the Republic of North Macedonia and south-western Bulgaria) (for Eastern Romance, cf. Maiden, 2016). Some authors such as Friedman and Joseph (2017: 55) count in also the co-territorial dialects of (Indic) Romani and to some extent the co-territorial dialects of Ibero-Romance Judezmo (spoken in Saloniki and Istanbul) and of Turkic (especially West Rumelian Turkish and Gagauz, both belonging to the Oghuz subbranch). In the following subsections, we shall outline the linguistic makeup of the Balkans (Section 2.1) and summarize the vivid debate sparkling around the notion of sprachbund, in particular as it relates to the Balkan languages (Section 2.2).

2.1   *Linguistic Makeup*
Starting from the early nineteenth century, scholars observed the spread of grammatical features across the Balkans: besides implicit thoughts by Leake (1814: 380), the first linguistic works explicitly concerned with the diffusion of grammatical structures in the Balkans are Kopitar (1829: 86), Miklosich (1861), Schuchardt (1884), Seliščev (1925) and, crucially, Sandfeld (1902) and Sandfeld (1926).[3] The property which attracted the interest of linguists the most was a high degree of structural convergence, as it is perhaps best exemplified by the following statement by Kopitar (1829: 86): "*nur eine Sprachform herrscht, aber mit dreyerley Sprachmaterie* [there is one grammar with three lexical materials, our translation]." While Kopitar's statement is blatantly exaggerated and should rather be taken as a slogan, instances of structural parallelism

---

2   Note that Istro-Romanian is generally not included in the alleged 'Balkan sprachbund'; alternatively, some authors consider it the least Balkan of the four Daco-Romance languages and claim that it shares a little amount of Balkanisms with the core Balkan languages (Zegrean, 2012: 43).

3   The 1926 paper, written in Danish, is better known as Sandfeld (1930) in French.

FIGURE 1   Map showing the major political boundaries, topography and rivers of the Balkan Peninsula (source: Reed, Kryštufek, and Eastwood, 2004: 10; reprinted by permission from Springer Nature Customer Service Centre GmbH)

are indeed observable. A case in point is the 'verbal complex', a string of material ordered in a template-like fashion including markers for negation, tense, modality, argument structure and a verb, as exemplified by the data in (1), adapted from Friedman and Joseph (2017: 56).[4]

---

4   Convergence of this kind, however, should be considered as surface similarity. As Sims and Joseph (2019: 101) have shown, "the internal structuring of the Balkan verbal complex differs from one language to another". Also Friedman (2011: 279), with respect to the feature of object doubling, speaks of the "differences in conditioning factors, which in turn indicate different degrees of integration into the grammar".

(1)	a. Albanian	*s'*	*do*	*te*	*ja*		*jep*
		NEG	FUT	SBJV	3.SG.DAT/3.SG.ACC		give.1.SG
	b. Daco-	*nu*	*o*	*să*	*îl*		*dau*
	Romanian	NEG	FUT	SBJV	3.SG.DAT/3.SG.ACC		give.1.SG
	c. dialectal	*δe*	*ϑe*	*na*	*tu*	*to*	*δoso*
	Greek	NEG	FUT	SBJV	3.SG.GEN	3.SG.ACC	give.1.SG
	d. dialectal	*ne*	*ḱe*	*da*	*mu*	*go*	*davam*
	Macedonian	NEG	FUT	SBJV	3.SG.DAT	3.SG.ACC	give.1.SG
	e. Serbian	*neću*		*da*	*mu*	*ga*	*dam*
		NEG.FUT.1.SG		SBJV	3.SG.DAT	3.SG.ACC	give.1.SG

'I will not give it to him.'

Miklosich (1861) was the first scholar to identify a series of features common to the Balkan languages, although the number of features to be included in the Balkan sprachbund, often called 'Balkanisms') (cf. Sandfeld, 1930; Joseph, 1983; 1992; 2010; Friedman and Joseph, 2017; 2022) varies consistently among authors and has led to a distinction between primary (or 'core') and secondary (or 'peripheric') Balkan languages. Here is a tentative list (including, whenever available, Romance examples).

a.   presence of a (stressed) mid-to-high central vowel in Albanian, Romanian, Bulgarian, some dialects of Macedonian and Serbian, some Romani varieties, and Turkish;

b.   presence of *i-e-a-o-u* in the vowel inventory without phonological contrasts in quantity, openness, or nasalization in Greek, Tosk Albanian, Romanian, Macedonian, Bulgarian, some dialects of Serbian, and Romani;

c.   devoicing of word-final stops in Bulgarian, Macedonian, Megleno-Romanian, Modern Greek, some Romani dialects, South Montenegrin and Torlak Serbian, and Turkish;

d.   voicing of voiceless stops after nasals (NT > ND) in Albanian, Greek, Aromanian;

e.   presence of the voiced/voiceless interdental spirants $\delta \vartheta (\gamma)$ in Aromanian, Albanian, Greek, (dialectal) Macedonian;

f.   presence of a future tense built with a reduced, often invariant, form of the verb 'want' in Greek, Tosk Albanian, Daco-Romanian, Istro-Romanian, Aromanian (2), Macedonian, Bulgarian, Serbian, and Romani;

(2)   Aromanian (Capidan, 1932: 466)

> *va*   *s-dormu*
> FUT   SUBJ-sleep.1.SG
> 'I will sleep'

g.   morphologically realized distinction of witnessed vs reported (including admirative and dubitative) in Albanian, Bulgarian, Macedonian, and Turkish,[5] and to a lesser extent in Romani, Serbian, and Aromanian (3);

(3)   (Gorna Belica Frasheriote) Aromanian (Friedman, 1994)
> *Abe,*   *munduem*   *ka*   *Silja*   *kăntac-ka!*
> hey   I.thought   that   Silja   sing.PART-ADMV
> 'Hey, I thought Silja is singing!'

h.   reduction in the nominal case system, especially a falling together of genitive and dative cases in Greek, Albanian, Daco-Romanian (4a-b), Istro-Romanian, Megleno-Romanian, Aromanian, Bulgarian, and Macedonian;

> Daco-Romanian (adapted from Tomić, 2006: 132)

(4)   a. *Petru*   *i=a*   *dat*   *fetei*   *o*   *floare*
>   Petru   3.SG.DAT=has   given   girl.F.SG.OBL   INDEF   flower
>   'Petru has given a flower to the girl.'
>   b. *floarea*   *fetei*
>   flower.DEF.F.SG   girl.DEF.F.SG.OBL
>   'the girl's flower'

i.   enclitic definite article in Albanian, Daco-Romanian,[6] Istro-Romanian, Megleno-Romanian (5), Aromanian, Macedonian, Bulgarian, and south-eastern (Torlak) Serbian;

(5)   Megleno-Romanian (Capidan, 1925: 149)
> *casi-li*
> house.PL-DEF
> 'the houses'

---

5 The most likely source of the evidentiality distinction in the Balkans is Turkish (cf. Section 2.2), where reported information (also referred to as 'non-firsthand' and 'indirective') on past events is encoded by the suffix *-mIş*, yielding, e.g., *gelmiş* 'obviously came' vs unmarked *geldi* 'came' (cf. Johanson, 2000: 81).

6 For Romanian, Ledgeway (2017) has convincingly shown that the definite article is no longer enclitic but inflectional.

j.   analytic comparative for adjectives in Greek, Albanian, Daco-Romanian, Istro-Romanian (6), Megleno-Romanian, Aromanian as well as Romani and Turkish;[7]

(6)   Istro-Romanian (Kovačec, 1971: 108)
         mài   bùr
         more   good
         'better'

k.   object clitic doubling in Albanian, Greek, Daco-Romanian (7), Megleno-Romanian, Aromanian, Bulgarian, and Macedonian, dialectally in Serbian, and to a limited extent in Romani;

l.   personal direct objects introduced by a preposition in Daco-Romanian (7), some Aromanian dialects (Sobolev, 2008: 117), and in southern Macedonian dialects;

(7)   Daco-Romanian
         l-am   văzut   pe/(pe)   Ana*
         OBJ-have   seen   to   Ana
         'I saw/have seen Ana.'

m.   double determination in deixis, that is, the co-occurrence of a demonstrative adjective with a definite article and a noun in Greek (8) and in Albanian and Slavic varieties;

(8)   Greek
         Idha   afton   ton   andhra
         saw.1SG   this   DEF   man
         'I saw this man.'

n.   use of enclitic oblique pronouns as possessive markers in Greek, Slavic, Daco-Romanian, Istro-Romanian, Megleno-Romanian, Aromanian (9);

(9)   Aromanian (Capidan, 1932: 415)
         sora=tsi
         sister=POSS.2.SG
         'your sister'

---

7 Bulgarian and Macedonian are generally included in the group of languages sharing the use of an analytically build comparative (cf., e.g., Joseph, 2010: 622). However, both languages use prefixation, e.g., Bulgarian *slab* 'weak', *poslab* 'weaker'.

o.   decreased use of a non-finite verbal complement replaced by finite complement clauses in Greek, Macedonian, Bulgarian, Serbian (especially the Torlak dialects), Romani and also occurring in Albanian (especially Tosk) and Eastern Romance (10);

(10)   Daco-Romanian (Tomić, 2006)
           Evită   *să*   *te*   *vadă*
           avoid.3.SG   SUBJ   2.SG   see.3.SG.PRS.SUBJ
           '(S)he avoids seeing you.'

The occurrence of the features listed in *a-o* (and of some others) in (some of) the Balkan languages led linguists to elaborate the notion of sprachbund, i.e., a group of languages whose similarity is not due to (narrow) genealogic relatedness, that is, inheritance from a common ancestor language, but to "historical development" (Boas, 1929: 7), that is, mutual influences over time. In the following section we introduce the notion of sprachbund, especially as it relates to the Balkans.

2.2   *Contact-Induced Change in the Balkans*

According to a recent definition, a sprachbund is "a geographically delimited area including languages from two or more language families (or subgroups) which share significant traits. Most of these traits are not found in languages from the same families outside the area, and can be considered area-defining" (Aikhenvald, 2018: 149). This definition evidently echoes the first explicit proposal of the concept of sprachbund that Trubetzkoy had formulated 95 years earlier:

> besides such genetic grouping, languages which are geographical neighbors also often group independently of their origin. It happens that several languages in a region defined in terms of geography and cultural history acquire features of a particular congruence, irrespective of whether this congruence is determined by common origin or only by a prolonged proximity in time and parallel development. We propose the term language union [*jazykovyj sojuz*][8] for such groups which are not based on the genetic principle (Trubetzkoy, 1923: 116, quoted from Toman, 1995: 204).

_____

8   The term sprachbund appears for the first time in Trubetzkoy's 'Proposition 16': "Groups composed of languages which show a high degree of similarity with respect to syntax, a similarity in the principles of morphological construction, and which offer a large number of common culture words, sometimes also an outward similarity in the phonological

A comparison of the two definitions shows that the intension of sprachbund has stayed unchanged for decades in its core elements, viz. structural convergence and geographic contiguity. However, an intense debate has sparked concerning some defining criteria of sprachbund, such as, for example, the minimum number of shared grammatical features, oscillating between a single trait (e.g., Jakobson, 1931) and several ones (Thomason, 2001: 101), and the minimum number of languages to which the trait(s) extend(s), e.g., at least three according to Schaller (1975: 58) (see Campbell, 2006: 7–10; 2017, for excellent overviews). As a result of such divergent opinions, the number of definitions of sprachbund "is almost coextensive with the number of linguists working in the field of areal linguistics" (Stolz, 2006: 33). Also the key criterium of geographic contiguity (Boas, 1929: 6) has met with general criticism (cf. Stolz, 2006: 36). The assumption, implicit in the topography-based approach, that geographic proximity is a proxy for linguistic contact has been argued to be inadequate and regarded as "*post hoc* attempts to impose geographical order on varied conglomerations of these borrowings" (Campbell, 2006: 1). Another property which some linguists have considered specific to sprachbünde is multilateral diffusion, as opposed to unilateral diffusion such as that occurring in substratum interference. For example, Tosco (2000: 359) weighs the existence of an Ethiopian linguistic area and concludes that it is not tenable on the basis of the fact that "[n]o multilateral contact is observable, but only unilateral diffusion in the form of a shared substratum". While seemingly promising, this property, too, is not conclusive: as Aikhenvald and Dixon (2001: 11) put it, "depending on the historical events, the direction of diffusion can suddenly change […]; this creates a 'historically' multilateral area, every synchronic 'cut' of which can be considered unilateral."

Finally, a central question in the sprachbund debate is that asked by Dahl concerning the "reality" of a linguistic area: "In the end, we are led to the following more far-going question about the notion of area: to what extent do areas […] have a reality of their own and to what extent are they just convenient ways of summarizing certain phenomena?" (Dahl, 2001: 1458). Admittedly, the question concerning the reality of linguistic areas is hard to answer, and we

---

inventories, –but which possess neither systematic sound correspondences, nor ha[ve] any correspondences in the phonological makeup of the morphological units nor any common basic lexical items– such languages groups we call Sprachbünde" (Trubetzkoy, 1928: 17–18, translated by Campbell, 2006: 3).

cannot answer it here. For all the reasons surveyed thus far, first and foremost
the non-specificity of structural convergence to linguistic areas, some promi-
nent scholars (in particular, Dahl, 2001; Campbell, 2006; 2017; Stolz, 2006) have
seriously questioned the validity of the notion, concluding that "[a]n area is
then simply the sum of many such binary ["one language influences another"]
relationships" (Dahl, 2001: 1458) and that "linguistic areas boil down merely to
a study of local linguistic borrowing and its history, and little else" (Campbell,
2006: 2).[9]

Whether one believes or not in sprachbund, in the end, it is unquestion-
able that the Balkan languages display a high degree of structural conver-
gence. The Balkan languages that share the most Balkan features are Albanian,
Macedonian, Bulgarian, and Romanian (Schaller, 1975: 100). As far as Romance
is concerned, according to this approach, Aromanian, Megleno-Romanian,
and Romanian are Balkan languages *stricto sensu*, whereas extinct Dalmatian
(Bartoli, 1906; Maiden, 2020), Istro-Romanian (see Loporcaro et al., 2021), and
Judezmo (see Friedman and Joseph, 2021a) are languages of the Balkans, that is,
languages which are spoken within the geographic boundaries of the Balkans.

Structural convergence is a frequent outcome of borrowing as the princi-
pal mechanism of contact-induced language change (Gardani, forthc.). In the
case of the Balkans, the source of borrowing is known in some cases: for exam-
ple, the loan verb marker -(*i*)*s*- (12) is ultimately Greek and the evidentiality
distinction (cf. feature *g*) was most likely borrowed from Turkish (Friedman,
1999: 521). Often, however, the exact origin of a spread Balkan trait is hard to
trace[10] and, even when we do not know it, it appears likely that the structural
parallelisms found in the Balkan languages have resulted from a stratification
of several processes of change (both contact-induced and internal), through-
out a turbulent history characterized by socio-political circumstances leading
to complex population movements and, during some periods, to ethnic and
linguistic intermingling (Banfi, 1991; Calic, 2019). Consequently, the source of
borrowing is not necessarily a single dominant language (Lindstedt, 2014).
During the Middle Ages, the languages of power in the Balkans were–at var-
ious times–Greek, Slavic, and Latin/Romance (especially Balkan Latin, i.e.,
the Latin variety used in the territory of Roman Dalmatia (cf. Skok, 1915) and

---

9   Still other scholars focusing on socio-historical aspects of language contact claim that the
      notion of sprachbund "remains a useful heuristic referring to the results of historical and
      social processes of language contact" (Friedman, 2011: 275).
10   For a discussion on the issue of determining directionality of change, exemplified with the
      diffusion of object doubling, see Friedman (2011: 283).

Venetian); during the Ottoman Empire, Turkish dominated the Balkan peninsula as the language of administration, trade, and the military; for centuries Church Slavonic was the language used in religious service in Bulgaria, Macedonia, Wallachia and Moldavia, while Greek enjoyed prestige among Christians as the language of the Greek Orthodox church; in particular, Greek was the language of the culture in the Balkans, and also a language of trade (cf. Friedman, 2006: 669–670). However, Greek is less Balkanized a language than Balkan Slavic, Eastern Romance, and Albanian (Lindstedt, 2014) and its influence was particularly strong in the southern Balkan regions, south of the so-called Jireček line (Jireček, 1901: 13–14), separating the influence of Greek from that of Latin, to the north). Besides borrowing, at least two more general mechanisms are responsible for spread of features and convergence, namely reanalysis and contact-induced grammaticalization, that is, a grammaticalization process which is transferred from a source language to a recipient language (Heine, 1994; Heine and Kuteva, 2003; 2005; Gast and van der Auwera, 2012). A combination of borrowing and reanalysis is evident in the diffusion of the formative -(*i*)*s*-, originally borrowed from Greek where it forms the perfective. As Breu (1991a; 1991b) has convincingly demonstrated, this formative has become the general loan verb integration suffix throughout the Balkan languages.[11] For example, given the Modern Greek verb *agapo* 'to love', -*s*- marks the perfective as it attaches to the perfective stem *agapi*-, yielding, e.g., the past form *agápisa*. Through the contact with other languages in the Balkans, this formative has been reanalyzed and refunctionalized as a loan verb marker (Gardani, 2016). Thus, in Arvanítika, a variety of Tosk Albanian spoken in Greece that has been involved in a four centuries lasting contact with Greek (Tsitsipis, 1998: 1), Greek *agapo* has been integrated as *agapís* 'I love'. The ensuing generalization of the borrowed formative as a loan verb marker, in terms of what Breu (1991a: 42) calls *analogische Ausweitung des Entlehnungsverfahrens*, is shown by the fact that the formative also applies to Greek verbs that do not display the sigmatic perfective stem. For example, in (11a) and (11b), the formative -*s*- occurs on a Greek-origin deponent verb *sképtomai / skéftomai* 'to think, reflect'. In Greek, deponent and medio-passive verbs are asigmatic, thus, the perfective stem of *sképtomai / skéftomai* is *skept- / skeft-* (data from Haebler, 1965: 166).

---

11   It is curious that this widespread trait has never been counted in as a Balkan sprachbund feature in any well-known feature list.

		Arvanítika (Albanian)	Greek
(11)	a.	*ʃcep-s-em*	*sképtome*
		think-LVM-PRS.1.SG	
		'I think'	
	b.	*u-ʃcep-s-ʃə*	*skéfthika*
		MEDP-think-LVM-PST.1.SG	
		'I thought'	

Similarly, in Bulgarian, verbs borrowed from Turkish are integrated by adding -(*i*)*s*- to the Turkish preterit morpheme -*DI*- (realized as -*dı, -di, -du, -dü, -tı, -ti, -tu, -tü*), which itself serves as a loan verb marker. The form *bojadisvam* in (12a) is made up of the Turkish *boyadı* (12b), preterit of *boyamak*, to which a loan verb marker and the inflectional formative are suffixed (data from Breu, 1997: 159).

(12)	a.	Bulgarian	b.	Turkish
		bojad-is-vam		*boya-dı*
		paint-LVM-1.SG		paint-PST.PFV
		'I paint'		's/he painted'

3   Balkan Languages Outside the Balkans

Beyond the Balkans, contact between Romance languages and other non-Romance Balkan languages has occurred in three main Balkan exclaves in the south of Italy. These involve the contact of Italo-Romance with Slavic (cf., e.g., Breu, 1998), Albanian (cf., e.g., Altimari and Savoia, 1994), and Greek (cf., e.g., Höhn et al., 2017), respectively.

As is well known (Rohlfs, 1937; 1977), the coexistence and linguistic contact between Italo-Romance speakers and Greek speakers (viz. Greko/Grecanico in Calabria and Griko in Salento) have led over the centuries to a considerable amount of grammatical convergence. In these areas, the result of contact has often manifested itself in processes of reanalysis of existing Romance features and patterns to adjust to the Greek model (Ledgeway, 2013; Ledgeway et al., 2018). Such convergence phenomena include, for example, case (dative and genitive), the use of determiners, verb movement and complementation (finite vs infinitival) (on the recession of the infinitive as a syntactic category, see Rohlfs, 1958; Loporcaro, 1995; 2013: 155–156;). For example, southern Calabrian (exemplified with Sinopolese in (13a)) shows a convergence towards the syntax of Griko (13b) as it allows finite complement clauses such as in (13a)

on the model of Italo-Greek (13b), while the expected Italo-Romance feature would be infinitival complementation (13c) (data from Rohlfs, 1972: 320, 327).

(13) a. Sinopolese (Reggio *Vogghiu*   *mi*   *dormu.*
          Calabria)   want.PRS.1.SG IRR   sleep.PRS.1.SG
      b. Griko (Castrignano *etèlo*   *na*   *plòso.*
          dei Greci)   want.PRS.1.SG IRR   sleep.PRS.1.SG
      c. Italian   *Voglio*   *dormire.*
                             want.PRS.1.SG sleep.INF
                             'I want to sleep.'

It is noteworthy that finite complementation under subject co-reference is one of the features considered key to the Balkan sprachbund as it occurs in Greek, Macedonian, Bulgarian, Serbian (especially the Torlak dialects), Romani, Albanian (especially Tosk), Daco-Romanian, Aromanian and Eastern Romance (cf. feature *o* and (10) in Section 2.1). Just as in the case of the Balkan languages, for example Romanian (Maiden, 2016: 121–122), in Sinopolese, too, we observe a tendency towards an increased use of finite complement clauses. In both southern Calabrian and Greko, infinitival complementation has been maintained, often alongside competing finite *mi* / *na* clauses, in conjunction with a class of restructuring predicates such as, e.g., the modal verb *potiri* 'can' in Mosorrofa Calabrese (14) (data from Loporcaro, 1995: 342) and *kùo* 'hear' in Greko (15) (data from Ledgeway, 2013; cf. also Baldissera, 2013; 2015 for Griko).[12]

(14) southern Calabrese (Mosorrofa)
      non   *pozzu*   *caminari*
      NEG   can.1.SG   walk
      'I can't walk.'

(15) Greko
      Egò   *tus=* *àcua*   *platèttsi* / *na*   *platèttsusi.*
      I   them=   hear.PST.1.SG talk.INF /   that   talk.SBJV.3.PL
      'I heard them talking.'

---

12 With respect to Griko, Baldissera (2015: 278) observes that properties, such as the retention of the infinitive, which is not shared by Standard Modern Greek, "can be found in Medieval Greek and can be explained as a result of reinforcement of the conservative tendency by contact with the neighboring Romance varieties."

Another key 'Balkan' feature, the falling together of genitive and dative cases (cf. feature *h.* in Section 2.1), has a parallel in Palizzese. Again as the result of the contact between Italo-Romance and Italo-Greek, we find a genitive structure that apparently calques the Greek dative-genitive syncretism; in reality, it is rather "a hybrid structure in which the indirect object is referenced in part through dative marking on the verbal head [*nci*] and in part through genitive marking on the nominal dependent [*da*]" (Ledgeway, 2013: 193) as in (16) (data from Squillaci, 2017: 6–7).

(16) a. Palizzese *nci=desi*   *u*   *regalu* *da*   *figghiola.*
                   to.her=give.PST.1.SG DEF gift   DEF.GEN girl

      b. Greko   *tis=edoka*   *to*   *kaloma* *ti*   *miccedda.*
                   to.her=give.PST.1.SG   DEF   gift   DEF.GEN girl

      c. Italian   *ho dato*   *dato* *il*   *regalo* *alla*   *bambina.*
                   I.have given   given DEF gift   to.DEF   girl
                   'I've given the gift to the girl.'

Still another key Balkan feature, the diffusion of analytically realized comparatives from originally synthetic structures (feature *j* in Section 2.1), is paralleled by the change occurred in Molise Croatian (17a) under the influence of Romance (17b) (Breu, 1996: 26; 2009), as opposed to synthetic comparative formation in Standard Croatian (17c).

(17) a. Molise Croatian   b. Italian   c. Croatian
          veče   *lip*   *più*   *bello*   *ljepši*
          more   pretty   more   pretty   pretty.COMP
          'prettier'   'prettier'   'prettier'

The data thus presented and the case studies by Breu (this special issue), Ledgeway, Schifano and Silvestri (this special issue), and Ralli (this special issue) show that the encounter between Romance and Balkan non-Romance languages outside the geographic boundaries of the Balkans has given rise to the same types of change as those originated in the Balkan sprachbund.

Summing up, we have shown that the languages outside the geographic boundaries of the Balkans, viz. Greek varieties, Albanian, and Croatian, in contact with Italo-Romance, have undergone the same types of change as the sprachbund languages themselves. Both groups can be characterized in terms of Aikhenvald's (2006: 5) metaphor of 'layered' languages, that is, languages whose "inherited 'core' is discernible underneath the subsequent 'layers' of innovative influence from outside". While it is obvious that all languages are mixed to some extent, Aikhenvald stresses that the significance of this layering varies. In the specific case of the Balkan languages, the degree of "diffusional

cumulation" (Swadesh, 1951) is so extraordinary that these languages display several layers of lexical material and grammatical features as a result of multiple processes of change and "mutual reinforcement" (Lindstedt, 2000), including borrowing, contact-induced grammaticalization, and secondary reanalysis-driven processes of contact-induced change, due to largely pairwise contact over long stretches of time.

4   Overview of this Special Issue

This special issue of the *Journal of Language Contact* is opened by a 'caveat paper'. In *Establishing contact. Slavonic influence on Romanian morphology*, Martin Maiden warns the (contact) linguists that they must fully exploit the full range of available comparative evidence in order to be able to exclude the possibility that apparently contact-induced effects are, in fact, explicable by internal factors. Maiden shows that the influence of Slavic models attributed to certain paradigmatic patterns of root allomorphy in the Romanian verbal system is at best indemonstrable. He makes this case by deconstructing claims that certain aspects of the distribution of root allomorphy in verb inflectional paradigms were induced by contact with Bulgarian. A more economical explanation – he argues – is achieved if Romanian is not singled out in a pairwise comparison with Bulgarian but rather duly analyzed against the background of what is independently known on (this aspect of) the Romance verb system (cf. Maiden, 2018).

The following paper, *Convergence by shared ancestry in Romance* by Paul Widmer, Stefan Dedio and Barbara Sonnenhauser, is also a methodological paper. Because in many cases of apparent contact-induced change the relevance of shared ancestry in the language sample and its interaction with processes such as matter and pattern borrowing (Gardani 2020a, 2020b) are difficult to specify, the authors quantify the change in similarity since the late Middle Ages in a sample of Romance and Germanic languages, with data from a selected grammatical domain, viz. the expression of reflexivity, and crucially compare their dynamics with patterns of change of similarity occurring in two contact areas, the British Isles and the Balkans. The results indicate a maintenance and gain of similarity in Romance as opposed to a loss of similarity in Germanic.

In *Contact-induced complexification in the gender system of Istro-Romanian*, Michele Loporcaro, Francesco Gardani and Alberto Giudici provide the first in-depth description of the borrowing of Croatian collective numerals into the northern branch of Istro-Romanian. They show that the introduction of a few lexical items encoding quantification has precipitated changes in the

recipient language, in a way that led to a restructuring of the morphosyntactic system, introducing (sub)gender overdifferentiation on just two agreement targets and, thereby, a complexification in this area of the grammar of northern Istro-Romanian whose degree of complexity had already increased previously, in two rounds, via the borrowing from Slavic of neuter agreement markers.

In *Eastern and Western Romance in the Balkans – The contrasting but revealing positions of the Danubian Romance languages and Judezmo*, Victor Friedman and Brian Joseph compare and contrast two Romance languages, Aromanian and Judeo-Spanish, and examine the extent to which they show the effects of Balkan-specific language contact. To this end, they review the behavior of the "usual suspects" in the two languages, including all the traits listed in a-o (Section 2.1), plus a series of properties of the sound patterns of the two languages as well as the occurrence of what Friedman and Joseph (2014; 2020) have termed "ERIC loans" (= "Essentially Rooted In Conversation"). They conclude that while Aromanian is thoroughly 'Balkan' as to its structure and lexicon, Judeo-Spanish is much less so (the latter finding converging with the results in Widmer, Dedio and Sonnenhauser) and argue that the difference between the languages as to their degree of linguistic Balkanization is due to several factors, including chronology, social circumstances, and the structure of the language at the time it entered the Balkans.

The next three papers are dedicated to Balkan exclaves. In *Italo-Albanian: Balkan inheritance and Romance influence*, Walter Breu deals with contact-induced change in Italo-Albanian and its effects on the Balkan inheritance of this minority language, focusing on the TAM systems, causative construction, and periphrastic structures. He shows that many traditional Balkan features have been weakened or lost in Italo-Albanian, whereas others have even expanded, but always in the direction of Romance models.

In *The negative imperative in Southern Calabria. Spirito greco, materia romanza again?*, Adam Ledgeway, Norma Schifano and Giuseppina Silvestri investigate imperative morphology in the two extreme southern Italian dialects of Mosorrofa, Cardeto and Gallicianò (Calabria). Capitalizing on new fieldwork data, they show that the differences in the extension of the -*ri* ending in the negative imperative correlates with differences in the duration of contact with Greko, as this was lost considerably earlier in Mosorrofa and Cardeto than in Gallicianò.

In *Contrasting Romance and Turkish as donor languages: Evidence from borrowing verbs in Modern Greek Dialects*, Angela Ralli studies loan verb accommodation techniques in a language contact situation involving Greek as recipient and Romance and Turkish as source languages. By drawing evidence from the

spontaneous speech of speakers of several Greek varieties, she shows that typological (in)compatibility between the source (semi-analytical Romance, agglutinative Turkish) and the recipient (fusional Greek) plays a major role in the process of loan verb integration.

Acknowledgements

Abbreviations

The abbreviations used in this paper are based on Lehmann (2004) and the Leipzig Glossing Rules (available at https://www.eva.mpg.de/lingua/resources/glossing-rules.php). In addition, admirative is abbreviated as ADMV.

References

Aikhenvald, Alexandra Y. 2006. Grammars in contact. A cross-linguistic perspective. In Alexandra Y. Aikhenvald and Robert M. W. Dixon (eds.), *Grammars in Contact*: *A Cross-Linguistic Typology*, 1–66. Oxford: Oxford University Press.

Aikhenvald, Alexandra Y. 2018. Evidentiality and language contact. In Alexandra Y. Aikhenvald (ed.), *The Oxford Handbook of Evidentiality*, 148–172. Oxford: Oxford University Press.

Aikhenvald, Alexandra Y. and Robert M. W. Dixon. 2001. Introduction. In Alexandra Y. Aikhenvald and Robert M. W. Dixon (eds.), *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*, 1–26. Oxford: Oxford University Press.

Altimari, Francesco and Leonardo M. Savoia (eds.). 1994. *I dialetti italo-albanesi*: *Studi linguistici storico-culturali sulle comunità arbëreshe*. Rome: Bulzoni.

Baldissera, Valeria. 2013. *Il dialetto grico del Salento*: *Elementi balcanici e contatto linguistico*. Venice: Università Ca' Foscari Venezia. PhD dissertation.

Baldissera, Valeria. 2015. The Griko dialect of Salento: Balkan features and linguistic contact (Ca' Foscari University of Venice, 2013). *Journal of Greek Linguistics* 15(2): 271–279.

Banfi, Emanuele. 1991. *Storia linguistica del sud-est europeo*: *Crisi della Romània balcanica tra alto e basso medioevo*. Milan: Angeli.

Bartoli, Matteo. 1906. *Das Dalmatische. Altromanische Sprachreste von Veglia bis Ragusa und ihre Stellung in der apennino-balkanischen Romania*. Vienna: Hölder.

Boas, Franz. 1929. Classification of American Indian languages. *Language* 5(1): 1–7.

Breu, Walter. 1991a. Abweichungen vom phonetischen Prinzip bei der Integration von Lehnwörtern. In Klaus Hartenstein and Helmut Jachnow (eds.), *Slavistische Linguistik 1990*: *Referate des XVI. Konstanzer Slavistischen Arbeitstreffens Bochum/ Löllinghausen 19-21.9.1990*, 36–69. Munich: Sagner.

Breu, Walter. 1991b. System und Analogie bei der Integration von Lehnwörtern: Versuch einer Klassifikation. *Incontri Linguistici* 14(1): 13–28.

Breu, Walter. 1996. Überlegungen zu einer Klassifikation des grammatischen Wandels im Sprachkontakt (am Beispiel slavischer Kontaktfälle). *Sprachtypologie und Universalienforschung (STUF)* 49: 21–38.

Breu, Walter. 1997. Integrazione di prestiti e problemi della fonologia diacronica dell'albanese. In Francesco Altimari (ed.), *Atti del 2° Seminario Internazionale di Studi Albanesi, Rende – San Giacomo di Cerzeto – San Cosmo Alb., 8–10 giugno 1994*, 155–168. Rome: Herder.

Breu, Walter. 1998. Romanisches Adstrat im Moliseslavischen. *Die Welt der Slaven* 43(1): 339–354.

Breu, Walter. 2009. La comparazione nello slavomolisano. Un risultato tipico del contatto linguistico assoluto. In Carlo Consani, Carmela Perta, Francesca Guazzelli, and Paola Desideri (eds.), *Alloglossie e comunità alloglotte nell'Italia contemporanea*: *Teorie, applicazioni e descrizioni, prospettive; atti del XLI congresso internazionale di studi della Società di Linguistica Italiana (SLI); Pescara, 27 – 29 settembre 2007*, 9–39. Rome: Bulzoni.

Calic, Marie-Janine. 2019. *The Great Cauldron: A History of Southeastern Europe*. Cambridge, Mass.: Harvard University Press.

Campbell, Lyle. 2006. Areal linguistics: A closer scrutiny. In Yaron Matras, April McMahon and Nigel Vincent (eds.), *Linguistic Areas*: *Convergence in Historical and Typological Perspective*, 1–31. New York: Palgrave Macmillan.

Campbell, Lyle. 2017. Why is it so hard to define a linguistic area? In Raymond Hickey (ed.), *The Cambridge Handbook of Areal Linguistics*, 19–39. Cambridge: Cambridge University Press.

Capidan, Theodor. 1925. *Meglenoromânii. I, Istoria şi graiul lor*. Bucharest: Cultura Naţională.

Capidan, Theodor. 1932. *Aromânii, dialectul Aromân: Studiu lingvistic*. Bucharest: Monitorul Oficial şi Imprimeriile Statului, Imprimeria Naţională.

Dahl, Östen. 2001. Principles of areal typology. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, and Wolfgang Raible (eds.), *Language Typology and Language Universals, volume 2*, 1456–1470. Berlin: De Gruyter.

Friedman, Victor A. 1994. Surprise! Surprise! Arumanian has had an admirative! *Indiana Slavic Studies* 7: 79–89.

Friedman, Victor A. 1999. Evidentiality in the Balkan languages. In Uwe Hinrichs (ed.), *Handbuch der Südosteuropa-Linguistik*, 519–544. Wiesbaden: Harrassowitz.

Friedman, Victor A. 2006. Balkans as a linguistic area. In Keith Brown (ed.), *Encyclopedia of Language and Linguistics*, 2nd edn., 657–672. Oxford: Elsevier.

Friedman, Victor A. 2011. The Balkan languages and Balkan linguistics. *Annual Review of Anthropology* 40(1): 275–291.

Friedman, Victor A. and Brian D. Joseph. 2014. Lessons from Judezmo about the Balkan Sprachbund and contact linguistics. *International Journal of the Sociology of Language* 226: 3–23.

Friedman, Victor A. and Brian D. Joseph. 2017. Reassessing Sprachbunds: A view from the Balkans. In Raymond Hickey (ed.), *The Cambridge Handbook of Areal Linguistics*, 55–87. Cambridge: Cambridge University Press.

Friedman, Victor A. and Brian D. Joseph. 2021a. Eastern and Western Romance in the Balkans: The contrasting but revealing positions of the Danubian Romance languages and Judezmo. *Journal of Language Contact* 14(1): 127–146.

Friedman, Victor A. and Brian D. Joseph. 2022. *The Balkan Languages*. Cambridge: Cambridge University Press.

Gardani, Francesco. 2016. Allogenous exaptation. In Muriel Norde and Freek Van de Velde (eds.), *Exaptation and Language Change*, 227–260. Amsterdam: Benjamins.

Gardani, Francesco (ed.). 2020a. *Borrowing matter and pattern in morphology*. Special issue. *Morphology* 30(4).

Gardani, Francesco. 2020b. Borrowing matter and pattern in morphology. An overview. *Morphology* 30(4): 263–282. doi: 10.1007/s11525-020-09371-5

Gardani, Francesco. forthc. Contact and borrowing. In Adam Ledgeway and Martin Maiden (eds.), *The Cambridge Handbook of Romance Linguistics*. Cambridge: Cambridge University Press.

Gast, Volker and Johan van der Auwera. 2012. What is 'contact-induced grammaticalization'? Examples from Mayan and Mixe-Zoquean languages. In Björn Wiemer, Bernhard Wälchli and Björn Hansen (eds.), *Grammatical Replication and Borrowability in Language Contact*, 381–426. Berlin: De Gruyter Mouton.

Haebler, Claus. 1965. *Grammatik der albanischen Mundart von Salamis*. Wiesbaden: Harrassowitz.

Heine, Bernd. 1994. Areal influence on grammaticalization. In Martin Pütz (ed.), *Language Contact and Language Conflict*, 55–68. Amsterdam: Benjamins.

Heine, Bernd and Tania Kuteva. 2003. On contact-induced grammaticalization. *Studies in Language* 27(3): 529–572.

Heine, Bernd and Tania Kuteva. 2005. *Language Contact and Grammatical Change*. Cambridge: Cambridge University Press.

Höhn, Georg F.K., Giuseppina Silvestri, and M. Olimpia Squillaci . 2017. Greek and Romance unagreement in Calabria. *Journal of Greek Linguistics* 17(2): 263–292.

Jakobson, Roman. 1931. Über die phonologischen Sprachbünde. *Travaux du cercle linguistique de Prague* 4, 234–240.

Jireček, Konstantin. 1901. *Die Romanen in den Städten Dalmatiens während des Mittelalters. Erster Theil*. Vienna: Carl Gerold's Sohn.

Johanson, Lars. 2000. Turkic indirectives. In Lars Johanson and Bo Utas (eds.), *Evidentials: Turkic, Iranian and Neighbouring Languages*, 61–87. Berlin: De Gruyter Mouton.

Joseph, Brian D. 1983. *The Synchrony and Diachrony of the Balkan Infinitive*: *A Study in Areal, General, and Historical Linguistics*. Cambridge: Cambridge University Press.

Joseph, Brian D. 1992. The Balkan languages. In William Bright (ed.), *International Encyclopedia of Linguistics*, 153–155. New York: Oxford University Press.

Joseph, Brian D. 2010. Language contact in the Balkans. In Raymond Hickey (ed.), *Handbook of Language Contact*, 618–633. Oxford: Blackwell.

Kopitar, Jernej. 1829. Albanische, walachische und bulgarische Sprache. *Jahrbücher der Literatur* 46: 59–106.

Kovačec, August. 1971. *Descrierea istroromânei actuale*. Bucharest: Editura Academiei Republicii Socialíste România.

Leake, William M. 1814. *Researches in Greece*. London: John Booth.

Ledgeway, Adam. 2013. Greek disguised as Romance? The case of southern Italy. In Mark Janse et al. (eds.), *Proceedings of the 5th International Conference on Greek Dialects and Proceedings of the 5th International Conference on Greek Dialects and Linguistic Theory*, 184–228. Patras: Laboratory of Modern Greek Dialects, University of Patras.

Ledgeway, Adam. 2017. The Romanian definite article in a comparative Romance perspective. In Adina Dragomirescu, Alexandru Nicolae, Camelia Stan, and Rodica Zafiu (eds.), *Sintaxa ca mod de a fi: Omagiu Gabrielei Pană Dindelegan, la aniversare*, 231–247. Bucharest: Editura Universităţii din Bucureşti.

Ledgeway, Adam, Norma Schifano, and Giuseppina Silvestri. 2018. Il contatto tra il greco e le varietà romanze nella Calabria meridionale. *Lingue Antiche e Moderne* 7: 95–133.

Lehmann, Christian. 2004. Interlinear morphemic glossing. In Geert E. Booij, Christian Lehmann and Joachim Mugdan (in collaboration with Wolfgang Kesselheim) (eds.), *Morphology. An International Handbook on Inflection and Word-Formation, volume 2*, 1834–1857. Berlin: Walter De Gruyter.

Lindstedt, Jouko. 2000. Linguistic Balkanization: Contact-induced change by mutual reinforcement. In Dicky Gilbers, John Nerbonne and Jos Schaeken (eds.), *Languages in Contact*, 231–246. Amsterdam: Rodopi.

Lindstedt, Jouko. 2014. Balkan Slavic and Balkan Romance: From congruence to convergence. In Juliane Besters-Dilger, Cynthia Dermarkar, Stefan Pfänder, and

Achim Rabus (eds.), *Congruence in Contact-Induced Language Change*: *Language Families, Typological Resemblance, and Perceived Similarity*, 168–183. Berlin: De Gruyter.

Loporcaro, Michele. 1995. Recessività sintattica dell'infinito e vitalità morfologica della desinenza infinitivale in alcuni dialetti del Reggino. In Roberto Ajello and Saverio Sani (eds.), *Scritti linguistici e filologici in onore di Tristano Bolelli*, 331–358. Pisa: Pacini.

Loporcaro, Michele. 2013. *Profilo linguistico dei dialetti italiani*. 2nd edition. Rome-Bari: Laterza.

Loporcaro, Michele, Francesco Gardani, and Alberto Giudici. 2021. Contact-induced complexification in the gender system of Istro-Romanian. *Journal of Language Contact* 14(1): 72–126.

Maiden, Martin. 2016. Romanian, Istro-Romanian, Megleno-Romanian, and Aromanian. In Adam Ledgeway and Martin Maiden (eds.), *The Oxford Guide to the Romance Languages*, 91–125. Oxford: Oxford University Press.

Maiden, Martin. 2018. *The Romance Verb: Morphomic Structure and Diachrony*. New York: Oxford University Press.

Maiden, Martin. 2020. Dalmatian (Vegliote). In Michele Loporcaro and Francesco Gardani (eds.), *The Oxford Encyclopedia of Romance Linguistics*. New York: Oxford University Press. doi:10.1093/acrefore/9780199384655.013.726.

Miklosich, Franz. 1861. *Die slavischen Elemente im Rumunischen*. Vienna: Kaiserliche-königliche Hof- und Staatsdruckerei.

Reed, Jane M., Boris Kryštufek, and Warren J. Eastwood. 2004. The physical geography of the Balkans and nomenclature of place names. In Huw I. Griffiths, Boris Kryštufek and Jane M. Reed (eds.), *Balkan Biodiversity. Pattern and Process in the European Hotspot*, 9–22. Dordrecht: Springer.

Rohlfs, Gerhard. 1937. Mundarten und Griechentum des Cilento. *Zeitschrift für Romanische Philologie* 57(2–4): 421–461.

Rohlfs, Gerhard. 1958. La perdita dell'infinito nelle lingue balcaniche e nell'Italia meridionale. In *Omagiu lui Jorgu Jordan: cu prilejul împlinirii a 70 de ani*, 733–744. Bucharest: Editura Academiei Republicii Populare Romîne.

Rohlfs, Gerhard. 1972. *Studi e ricerche su lingua e dialetti d'Italia*. Florence: Sansoni.

Rohlfs, Gerhard. 1977. *Grammatica storica dei dialetti italogreci* (*Calabria, Salento*), 2nd edition. Munich: C.H. Beck.

Sandfeld, Kristian. 1902. Der Schwund des Infinitivs im Rumänischen und den Balkansprachen. *Jahresbericht des Instituts für rumänische Sprache zu Leipzig* 9: 75–131.

Sandfeld, Kristian. 1926. *Balkanfilologien, en Oversigt over dens Resultater og Problemer*. Copenhagen: Luno.

Sandfeld, Kristian. 1930. *Linguistique balkanique: Problèmes et résultats*. Paris: Champion.

Schaller, Helmut W. 1975. *Die Balkansprachen*: *Eine Einführung in die Balkanphilologie*. Heidelberg: Winter.

Schuchardt, Hugo. 1884. *Dem Herrn Franz von Miklosich zum 20. Nov. 1883. Slawo-Deutsches und Slawo-Italienisches*. Graz: Leuschner and Lubensky.

Seliščev, Afanasij. 1925. Des traits linguistiques communs aux langues balkaniques: Un balkanisme ancien en bulgare. *Revue des Études Slaves* 5(1): 38–57.

Sims, Andrea D. and Brian D. Joseph. 2019. Morphology versus syntax in the Balkan verbal complex. In Iliyana Krapova and Brian Joseph (eds.), *Balkan Syntax and* (*Universal*) *Principles of Grammar*, 99–150. Berlin: De Gruyter Mouton.

Skok, Petar. 1915. *Pojave vulgarno-latinskoga jezika na natpisima rimske provincije Dalmacije*. Zagreb: Knižara Jugoslavenske Akademije Lav. Hav. Hartmana.

Sobolev, Andrej N. 2008. On some Aromanian grammatical patterns in Balkan Slavonic dialects. In Biljana Sikimić and Tijana Ašić (eds.), *The Romance Balkans: Collection of Papers Presented at the International Conference The Romance Balkans, 4–6 November 2006*, 113–121. Belgrade: Serbian Academy of Sciences and Arts, Institute for Balkan Studies.

Squillaci, Maria O. 2017. *When Greek Meets Romance*: *A Morphosyntactic Investigation of Language Contact in Aspromonte*. Cambridge: University of Cambridge. PhD dissertation.

Stolz, Thomas. 2002. No Sprachbund beyond this line! On the age-old discussion of how to define a linguistic area. In Paolo Ramat and Thomas Stolz (eds.), *Mediterranean Languages: Papers from the MEDTYP Workshop, Tirrenia, June 2000*, 259–281. Bochum: Brockmeyer.

Stolz, Thomas. 2006. All or nothing. In Yaron Matras, April McMahon and Nigel Vincent (eds.), *Linguistic Areas*: *Convergence in Historical and Typological Perspective*, 32–50. New York: Palgrave Macmillan.

Swadesh, Morris. 1951. Diffusional cumulation and archaic residue as historical explanations. *Southern Journal of Anthropology* 7(1): 1–21.

Thomason, Sarah G. 2001. *Language Contact*: *An Introduction*. Edinburgh and Washington, DC: Edinburgh University Press and Georgetown University Press.

Toman, Jindřich. 1995. *The Magic of a Common Language*: *Jacobson, Mathesius, Trubetzkoy, and the Prague Linguistic Circle*. Cambridge, Mass.: The MIT Press.

Tomić, Olga M. 2006. *Balkan Sprachbund Morpho-Syntactic Features*. Dordrecht: Springer.

Tosco, Mauro. 2000. Is there an "Ethiopian language area"? *Anthropological Linguistics* 42(3): 329–365.

Trubetzkoy, Nicolai S. 1923. Vavilonskaja bashnja i smeshenie jazykov. *Evrazijskij vremennik* 3: 107–124.

Trubetzkoy, Nicolai S. 1928. Proposition 16. In *Actes du Premier Congrès International de Linguistes*: *À La Haye, du 10 – 15 avril 1928*, 17–18. Leiden: A.W. Sijthoff.

Tsitsipis, Lukas D. 1998. *A Linguistic Anthropology of Praxis and Language Shift*: *Arvanítika (Albanian) and Greek in Contact*. Oxford: Clarendon Press.

Vezenkov, Alexander. 2017. Entangled geographies of the Balkans: The boundaries of the region and the limits of the discipline. In Roumen Dontchev Daskalov, Diana Mishkova, Tchavdar Marinov, and Alexander Vezenko (eds.), *Entangled Histories of the Balkans – Volume 4: Concepts, Approaches, and (Self-)Representations*, 115–256. Leiden: Brill.

Zegrean, Iulia-Georgiana. 2012. *Balkan Romance*: *Aspects on the Syntax of Istro-Romanian*. Venice: Università Ca' Foscari Venezia. PhD dissertation.

Establishing Contact: Slavonic Influence on Romanian Morphology?

Martin Maiden
Statutory Professor of the Romance Languages Research Centre for Romance
Linguistics, Faculty of Linguistics, Philology, and Phonetics and Faculty of
Medieval and Modern Languages, University of Oxford, Oxford, UK
martin.maiden@mod-langs.ox.ac.uk

Abstract

It is not disputed that Slavonic languages have influenced the inflexional morphology of Romanian and its closely related Daco-Romance varieties. For example, Romanian vocatives in -o, Istro-Romanian perfective verb-roots, and probably the Megleno-Romanian first and second person singular endings -um and -iʃ, are all attributable to Slavonic. These cases generally involve loans of 'morpheme'-like entities, phonological strings associated with a particular grammatical meaning. However, it has recently been suggested (e.g., by Elson, 2017) that certain Romanian paradigmatic patterns of root allomorphy in the verb, notably those involving the effects of palatalization, are influenced by Slavonic models. Some of these patterns appear to be of a qualitatively different kind from run-of-the-mill 'morphemic' loans, in that they are autonomously morphological, and cannot be associated within any coherent grammatical meaning. The borrowing of such purely morphological patterns under conditions of language contact has not hitherto been attested in the literature on language contact, and the evidence for such cases in Romanian deserves careful scrutiny. Unfortunately, the arguments provided for these putative borrowings can be shown to be rest on seriously flawed assumptions. Examination of those arguments serves to focus our attention on the kind of criteria that need generally to be met if the effects of language contact in morphology (or any other domain) are to be plausibly demonstrated. In particular, I shall emphasize the need for appeals to language contact carefully to exploit the full range of available comparative evidence, and to establish rigorous criteria to exclude the possibility that apparent contact effects are explicable by factors internal to the history of the recipient language.

Keywords

Romanian – Bulgarian – paradigms – morphomes – diachrony – comparative

1   **An Alleged Bulgarian Influence on Daco-Romance Verb Morphology**

Romanian verb morphology shows some deviations from the historically pre-dicted development, and these are attributed by Elson (2017) to the effect of early contact with 'middle Bulgarian'.[1] While Romanian verbs with root-final dentals show an expected and regular effect of palatalization by proto-Romance yod in the first person singular present indicative, they fail to do so in the third person plural present indicative. Table (1) demonstrates this, where Latin unstressed front vowels (I or E) before a vowel yield yod ([j]) in proto-Romance, which in turn produces palatalizing and affricating effects on preceding dental conso-nants. This 'yod effect' (hereafter, YE) is predicted for the first person singular present indicative and for the third person plural present indicative. While it is duly observed in the former case, it is surprisingly absent (as indicated by '!' in Table 1) in the latter, in old and modern dialectal Romanian.[2]

The other unexpected deviation from regular sound change which Elson discusses is the fact that proto-Romance root-final velars wholly fail to show expected YE in the first person singular present indicative and in third per-son plural present indicative (Table 2). The phonetically regular outcome is an affricate [ts], [ʤ] (cf. SOCIUM > soţ 'husband'; *ˈfakja > faţă 'face'; ABSUNGIAM > osânză 'lard'), but what we systematically find in the verb is an apparently unmodified velar [k], [g]; again, the aberrant forms are indicated with '!'.

Elson (2017: 889f.) believes that such facts show Romanian to have:

> reorganized the *pattern of allomorphic variation* attested in the present system of verbs with radical-final dental, which, in Bulgarian, opposed the first person singular to the other forms, and velar, which opposed the first person singular and third person plural to the remaining forms […] In oth-er words, it adopted an organizational, or systemic, attribute of Bulgarian

---

1 For the status of the glottonym '(middle) Bulgarian', I refer readers to Elson's study, and particularly Elson (2017: 848 n11; 868 n40). I use here the term 'Slavonic' to refer to the entire language family, rather than solely to Old Church Slavonic.

2 In fact, the yod effect is not limited to the present tense, occurring also in the third person forms of the subjunctive. I return to this point later, but here follow Elson in focusing on the present.

TABLE 1   Yod effects in Romanian present tense verb morphology

Latin		
1SG.PRS.IND	3SG.PRS.IND	3PL.PRS.IND
AUDIO	AUDIT	AUDIUNT
SENTIO	SENTIT	SENTIUNT
proto-Romance		
1SG.PRS.IND	3SG.PRS.IND	3PL.PRS.IND
'audjo	*'aude*	*'audjun*
'sɛntjo	*'sɛnte*	*'sɛntjun*
(old + modern dialectal) Romanian³		
1SG.PRS	3SG.PRS	3PL.PRS
auz 'hear'	*aude*	*aud*!
simţ 'feel'	*simte*	*simt*!

TABLE 2   Apparent lack of yod effect on velars in the Romanian verb

Latin		
1SG.PRS	3SG.PRS	3PL.PRS
FACIO	FACIT	FACIUNT
FUGIO	FUGIT	FUGIUNT
'proto-Romance'		
1SG.PRS.IND	3SG.PRS.IND	3PL.PRS.IND
'fakjo	*'fake*	*'fakjun*
'fugjo	*'fuge*	*'fugjun*
Romanian		
1SG.PRS	3SG.PRS	3PL.PRS
fac! 'do'	*face*	*fac*!
fug! 'flee'	*fuge*	*fug*!

> conjunction. [...] We may therefore see, in support of Bulgarian influence in the Romanian present system, a *general* tendency for the importation of Bulgarian *organization* into Romanian verbal morphology [...].

This is a claim with intriguing implications for the typology of contact effects. The grounds for making it are as follows: Elson (2017: 879) discerns as the

---

3 Suffice it to say here that orthographic *z* (= [z]) continues an earlier affricate [ʣ]; the letter *ţ* stands for [ts]. For details of the phonological history of these forms, see e.g., Sala (1976: 120–36).

TABLE 3   Middle Bulgarian models alleged to have influenced old Romanian present tense
alternations (č = [ʧ])

(a)	middle Bulgarian	old Romanian	(b) middle Bulgarian	old Romanian
1s	*vižᵈ-a* (< *vid-y-ǫ < *vid-i-ǫ) 'see'	*vădzu* 'see'	*reka* 'say'	*fa*[k]*u* 'do'
2s	*vidiš*	*vedzi*	*rečeš*	*fa*[ʧ]*i*
3s	*vidi*	*veade*	*reče*	*fa*[ʧ]*e*
1p	*vidim*	*vedem*	*rečem*	*fa*[ʧ]*em*
2p	*vidite*	*vedeți*	*rečete*	*fa*[ʧ]*eți*
3p	*videt*	*vădu*	*rekat*	*fa*[k]*u*

model for the two Romanian patterns the 'middle Bulgarian' present indicative conjugations of *vid-i* 'see' and of *reč-e* 'say' (Table 3). The first person singular form of the middle Bulgarian verb 'see' shows the effects of 'palatalization' of a root-final dental, leaving that consonant intact in the rest of the present paradigm, while in the verb 'say', an effect known as 'softening' of velar consonants has applied throughout the present tense, save in the first person singular and third person plural,[4] where the conditioning environment is historically lacking.

According to Elson (2017: 869), the mechanisms behind the Romanian facts are as follows:[5]

we may assume that the Bulgarian present indicative distribution of paradigmatic *y vis-à-vis radical-final dental versus velar was imported into

---

4 The emergence of the alternant *dz* also in the second person singular is a historically separate phenomenon and certainly later than the yod effect.

5 There is a serious chronological problem with Elson's assumption that the phenomenon occurred 'preceding or during the Romance yotation', in that the Romance palatalization by yod was operative as early as the second century (Väänänen, 1963: 54–56), and indeed is a pan-Romance phenomenon, while it is by no means obvious that it was still productive by the time the Slavs came into contact with the Romance speakers. Elson himself recognizes as much (2017: 869 n40), yet then takes the illegitimate step of trying to salvage his hypothesis by suggesting, without independent motivation, that the date of contact between the two linguistic groups must have been earlier than is generally assumed. Fortunately, he also more plausibly suggests that his analysis could be recast in terms of the existing *reflexes* of palatalization – and this is surely the only safe way to approach the facts.

Romanian preceding or during the Romance yotation, and perhaps softening as well, yielding the distribution of reflexes of yotation we find in the contemporary language [...].

He further writes (Elson, 2017: 870) that

all verbs with radical-final dental have a reflex [of the effect of yod] in in the first person singular, and only in that form, while *all* verbs with radical-final velar [...] lack a reflex in both the first person singular and third person plural. This consistency [...] follows from the Bulgarian distribution, which was instantiated *regularly*: intra-paradigmatic **y* occurred in the first person singular of all verbs with radical-final dental, but in none with radical-final velar.

We shall see that, in reality, the evidence for 'Bulgarian' influence is very weak. First, however, we need to consider the theoretical significance of Elson's hypothesis.

2   Theoretical Significance of the 'Bulgarian' Hypothesis

2.1   *Typology of Slavonic Influences on Romanian*
It is uncontroversial that the prolonged contact in the Middle Ages (perhaps from the sixth century until the twelfth) between Romanian and Slavonic determined significant changes in the former (see, e.g., Rosetti, 1986: 263–320; Sala, 2013: 214–226). There has been obvious and extensive lexical borrowing, which penetrated quite basic semantic domains, such as the names of body parts (e.g., *glezna* 'ankle', *obraz* 'cheek'), including 'semantic calquing', as in the words *picior* and *mână* which, although undoubtedly of Latin origin (from PETIOLUS 'stalk' and MANUS 'hand'), have acquired a typically Slavonic pattern of reference in that they denote, respectively, the entirety of the upper and of the lower limb, without differentiating 'foot' from 'leg' and 'hand' from 'arm'. In the domain of morphology,[6] there is also considerable borrowing (cf. Rosetti, 1986: 278–282; Petrucci, 1999: 90–135), especially involving derivational affixes. Just a few examples are: the iterative prefix *răs-* (e.g., *a răsciti* 'to read and re-read'), the adjectival or agentive suffix *-nic* (e.g., *obraznic* 'cheeky',

---

6   Syntactic influences from Slavonic are widely assumed, if controversial (see, e.g., Dragomirescu, 2015).

zilnic 'daily'), feminine diminutive or agentive *-iță* (e.g., *fetița* 'little girl', *actriță* 'actress'), feminine ethnic *-că* (e.g., *româncă* 'Romanian woman').

Slavonic influence on Romanian inflexional morphology is less evident (cf. Sala, 2013: 216). An incontestable case involves the ending *-o* which is available to form the vocative of words ending in *-ă* (the overwhelming majority of them feminine), such as *fată* 'girl', vocative *fato.* A classic example first brought to prominence by Weinreich (1968), involves the Megleno-Romanian desinences 1SG -um, 2SG -iʃ, which are apparently borrowed from Macedonian. These are described as 'undisputably' Slavonic by Elson (2017: 886), although their status as a Slavonic loan has in fact been seriously disputed (Friedman, 2009).[7] An eye-catching source of innovation in inflexional morphology is the creation, under Croatian linguistic influence, of morphological aspect distinctions throughout the paradigm of the verb in Istro-Romanian dialects, by means of borrowing of affixes or indeed of entire lexemes in order to furnish a full set of distinctively perfective forms in opposition to the imperfective (see, e.g., Kovačec, 1971: 123–130; Sala, 2013: 222f.; Maiden, 2016a: 111).[8]

What all the generally acknowledged and indisputable examples of Slavonic influence on Daco-Romance morphology share is their 'concreteness': they usually involve readily segmentable pieces of morphological structure (affixes or inflexional desinences) associated with clearly defined derivational meanings or grammatical functions, and in this respect they are rather like traditional lexical borrowings. In traditional terms, they might be viewed as different from the latter only in the respect that they are 'bound morphemes'. The Istro-Romanian case mentioned above involves the importation of a variety of devices as markers of paradigmatic opposition, but they are 'concrete' in the sense that they can be generally identified as markers of a particular value for aspect. The examples given in Elson (2017) are of a kind significantly different from all the foregoing, a type which he calls 'realizational'. Elson (2017: 888) writes:

> When innovation in morphological systems is realizational [...] there are no new forms, but only changes in the phonemic composition of existing ones. We therefore, perhaps understandably, find a strong tendency among linguists to interpret such compositional changes as the result of non-phonetic innovation in the form of *analogy*, i.e., as the result of system-internal innovation with no regard to the defensibility of its in-

---

7 See also Kossmann (2015: 259f.).
8 See Hurren (1969) for a more detailed account of the various aspectual values marked in the Istro-Romanian verb.

vocation (i.e., whether there was an existing pattern which might have
served as the basis for the changes in phonemic composition). It is in
such situations that we must give serious consideration to contact […]

What is at issue is a matter of imposition on Romanian of Bulgarian patterns
of paradigmatic distribution, instead of the expected and phonologically reg-
ular patterns.[9] One of these patterns (Table 3(b)) pertains to an opposition
between the alternants [tʃ] and [k], whose paradigmatic distribution is defined
(for Bulgarian) over a set of cells which is irreducible to any common semantic
or functional denominator, and whose content actually involves opposite val-
ues (singular vs plural, and third person vs first person), while being arbitrar-
ily confined to the present tense (other tenses do not show this alternation).
No coherent 'meaning' can be assigned to these alternants, and their paradig-
matic domain is synchronically arbitrary. The same is true, albeit rather less
obviously, for the type of alternation exemplified in Table 3(a). The opposition
involved here involves the combination of values 'first person' and 'singular',
which might be regarded as expressing EGO, but the fact that it is confined just
to the present tense (other first person singular forms do not show the alter-
nant in the relevant verbs) again confers on it an idiosyncratic paradigmatic
distribution. Both these 'relizational' occurrences probably qualify as 'morpho-
mic' (cf. Aronoff, 1994),[10] in that they are defined over 'irreducibly heteroge-
neous feature combinations' Maiden (2018: 20).[11] They cannot, that is to say,
be assigned any coherent function and they are, in that sense, 'autonomously
morphological', having no synchronic motivation outside the morphology
itself. Maiden (2018: 22) also specifies a particular, typological, criterion for the
identification of a morphomic structure:

> An important criterion for morphomic status is *local uniqueness*. […]
> morphomic structures are virtually always the local, fortuitous, cumula-
> tive effects of other (often phonological) changes. Precisely because they

9   Elson's contrast between 'realizational' and the 'incremental' (addition of new forms) uses
    two well-established terms in morphological theory in a quite different way from what has
    come normally to be understood by them (see, e.g., Stump, 2001: 2).

10 I do not mean to suggest that 'relizational', in Elson's sense, and 'morphomic', in mine, are
    synonymous. My interest here is specifically in certain of the phenomena identified by him
    as 'relizational' which seem to me also to be morphomic.

11 'Potentially' in that, as Maiden goes on to argue, the psychological reality of a putative
    morphomic structure can only be guaranteed when speakers implement some change
    which replicates the pattern. In the present case, the very fact of borrowing might count as
    such a change.

tend to reflect the combined effects of more than one extramorpholog-
ically motivated change, it is highly unlikely that the same pattern will
recur in any other language. If one does find the same pattern in a dif-
ferent language with a different history, and if one can rule out shared
inheritance (or language contact: see Chapter 7), then we may suspect a
shared extramorphological motivation for the phenomenon, even if it is
not entirely clear to us what that motivation might be. [...] An unambigu-
ously morphomic structure is unlikely ever to exist independently in two,
let alone more, unrelated [...] languages.

Despite the parenthetical allusion in the foregoing quotation to 'language
contact' as a possible source of the appearance of a morphomic structures in
languages that are unrelated,[12] the case that I actually allude to in the quota-
tion above is not a matter of direct borrowing of a morphomic structure, but
of borrowing from Romance into Germanic of a syntactic construction which
subsequently gives rise to a morphomic distribution in both language-families,
effectively as a result of grammaticalization (see Maiden, 2018: 252). What is
involved is the emergence of the 'past participle', as morphomically distrib-
uted over both passive and perfective constructions. In contrast, and to the
very best of my knowledge, nothing quite like what Elson claims, implying
(in my terms) morphomically distributed alternant patterns, is attested any-
where else in the literature on morphological borrowing.[13] For example, the
wide-ranging overviews of borrowing in inflexional morphology provided by
Kossmann (2015) or Gardani (2018) mention nothing similar.[14] It is for this rea-
son that Elson's claims acquire especial theoretical interest. Can morphomic
patterns–inherently arbitrary and synchronically 'nonsensical' as they are–re-
ally be transferred under conditions of language contact?

   I need to anticipate a disappointing conclusion: we shall see that there is
simply no good reason to believe that the phenomena presented in Section
1 owe anything to contact with 'middle Bulgarian', despite Elson's claims.
Consequently, nothing will be added here to our understanding of morphomic
structures in language contact. Rather, the process of testing the correctness
of the notion that these might be effects of contact will focus our minds on
some essential desiderata for establishing that any linguistic phenomenon is

---

12   In the case of Romance and Slavonic, the contact languages are related, but only distantly.
13   Although see Gardani et al. (2015: 13) for a case of apparent borrowing of inflexion-class
      marking.
14   Gardani clearly distinguishes 'abstract' patterning in borrowing, but what is principally
      involved (Gardani, 2018: 4f.; 10–12) is the ordering of constituents in compounds and is, in a
      sense, 'syntactic'.

an effect of contact. We may begin by setting out some common sense require-
ments for the construction of a plausible argument for contact (Section 2.2).

2.2 *Conditions on Plausible Arguments for Contact Effects*

Certain fairly obvious conditions need to be fulfilled if we are to present a
cogent and plausible argument for some linguistic phenomenon X in language
R being an effect of contact with language B.[15]

i.   X must demonstrably exist in B at the time of contact with R.
ii.   The manifestation of X in language R should be the *same* as its mani-
        festation in language B: if X in language R is identical in fine detail to
        something in language B, then contact may plausibly be assumed; if the
        manifestation of X in language R is only 'rather like' some phenomenon
        in language B, the case for X as a contact effect is weakened.
iii.   X should be unique to B and R: if X also appears in sister varieties of R
        (let us call them I and G), which have never been in contact with B, then
        the case for X as due to contact with B is very seriously weakened, and it
        becomes likely that X is an independent development in language R and
        its family, and only coincidentally resembles what we find in B.

Of course, none of the above are *necessary* conditions for X to be a result of
contact between R and B.[16] They are, however, the kind of conditions that must
be met if we are to articulate a *plausible case* for a contact effect.[17] If they are
not met, then the claim that X is an effect of contact becomes mere specu-
lation, of little or no scientific value. Elson's arguments as given in Section 1,
theoretically suggestive as they are, ultimately underscore the importance of
meeting such conditions – because they simply fail to satisfy (iii) and fall some
way short of satisfying (ii). We see why in what follows.

3 The Romanian Phenomena are not Unique to Romanian among Romance Languages

Elson (2017) is clearly aware of the need to satisfy what I have called condi-
tion (iii) above. His argument is indeed that the sister varieties of Romanian,

---

15   The labels 'B' and 'R' are intended to be arbitrary and thus cross-linguistically valid, but
      they are also obviously inspired in our case by 'Romanian' and (middle) 'Bulgarian'. And
      'I' and 'G' below can also particularly be interpreted as standing for Italo-Romance and
      Gallo-Romance, as will become apparent.
16   See also Elson (2017: 890f.).
17   Compare also the criteria proposed by Ascoli (1881–1882) for demonstrating substrate
      influences.

belonging to a group of languages descended from what he calls 'Proto-East-Romance', do not show the phenomena at issue, that they are therefore unique to Romanian within that group of languages, and that it is therefore legitimate to seek the origins of these phenomena in some circumstance peculiar to Romanian. That circumstance is identified as contact with middle Bulgarian, in which language he discerns morphological patterns very similar to what we find in Romanian.

Now, whether there ever was such a thing as 'Proto-East-Romance' is highly dubious, and the label 'East-Romance' is probably more a geographical expression than an early node in the branching of the Romance languages (cf. Malkiel, 1991; Bossong, 2016: 71). Elson's use of the label is also unnecessarily restrictive: the occurrence of the relevant phenomena in *any* Romance language other than Romanian would seriously weaken the argument, given that no other Romance language was in contact with 'middle Bulgarian'.[18] However that may be, the definition which Elson provides of 'Proto-East-Romance' ignores vast quantities of relevant historical-comparative data, and thereby wholly distorts the picture. 'Proto-East-Romance' is defined by him as 'the form of Romance recoverable through comparative reconstruction limited to Italian and Romanian' (Elson, 2017: 846 n5). In fact, to do any kind of responsible 'comparative reconstruction' we always need to explore and exploit the full range of comparative data available. Whatever 'Proto-East-Romance' might be, it must surely comprise at least the *entire Italo-Romance linguistic domain*, of which 'Italian' is merely a fragment, one face in an extremely diverse crowd of linguistic varieties, and specifically a form of Florentine Tuscan.[19] To treat 'Italian' as somehow constituting the complement, within 'Proto-East-Romance', of Romanian,[20] is an elementary error almost guaranteed to be fatal. Yet Elson (2017: 869) sets up a simple dichotomy between the 'Romanian distribution and its deviation from the expected distribution, *attested in Italian*' [my emphasis]. The consequences are predictably infelicitous.[21]

---

18   One would imagine that Dalmatian (cf. Maiden, 2016b) would also count as belonging to 'Proto-East-Romance', and it has clearly been in contact from an early date with Slavonic, but the the available data do not allow us to observe possible parallels with Romanian in the relevant respects.

19   This is an oft-unheeded truth constantly repeated in manuals of Romance or Italian linguistics (e.g., Harris, 1988: 18f.; Maiden, 1995: 3–5).

20   In the relevant respects, what we see for 'Romanian' does happen to be valid for all Romanian's sister Daco-Romance varieties, however.

21   The only non-Italian Italo-Romance form which Elson mentions is *facciono* (Elson, 2017: 836), but it is presented as the 'Italian' reflex of FACIUNT 'they make' (although the Italian form is *fanno*). I am unaware that anything like it has ever been attested in Italian

TABLE 4   Resemblance between first person singular and third person plural present indicative forms in old Italian

1SG.PRS	3PL.PRS		1SG.PRS	3PL.PRS
CREDO	CREDUNT	>	*credo*	*credono*
MITTO	MITTUNT	>	*metto*	*mettono*
UENIO	UENIUNT	>	*vegno*	*vegnono*

The 'expected distribution, attested in Italian', projected to be the norm for 'Proto-East-Romance', has, in Elson's analysis, a rather complex, and specifically morphological, origin. The appearance, allegedly normal for 'Proto-East-Romance', of the palatal alternant in the third person plural present indicative (as well as in the present subjunctive and first person plural present indicative), is explicable, Elson believes (2017: 865 n36), from 'the obvious segmental relationship between the first person singular and the third person plural in the ĕ and ɪ conjugations' in Italian. Examples of the relevant kind are as in Table 4. On this basis, original second conjugation verbs were then allegedly remodelled (implicitly, at some early stage of the alleged branch of Romance) so that their third person plural present indicative form was also made to contain the form of the first person singular present indicative, in the manner shown in Table 5.

Now what is allegedly an 'obvious segmental relationship' just vanishes if one takes proper account of the comparative Italo-Romance data. Indeed, a mere glance at a major linguistic atlas of Italy, such as the AIS, instantly makes 'obvious' a very different state of affairs indeed. The presence of the palatal[22] alternant in the third person plural is characteristic not just of 'Italian' (and of Tuscan), but of most of the dialects of the Rome-Ancona corridor (Lazio, Umbria, parts of the Marche). Yet in these latter dialects, unlike 'Italian', the third person plural *does not, and never did, subsume the form of the first person singular.* Here (a) the third person plural originally ended in -u vs first person

---

or medieval Florentine. It is observable in some medieval texts from Umbria (principally the *Statuto del Comune e del Popolo di Perugia del 1342 in volgare*), and–very marginally–also from Tuscany (data from *OVI*).

22   In some cases the presence of a velar alternant overlies an earlier palatal. See, e.g., Maiden (1995: 137).

TABLE 5   Alleged analogical mechanism for the introduction of palatalized root-allomorphs
           into the old Italian third person plural present indicative

1SG.PRS	3PL.PRS		1SG.PRS	3PL.PRS
UIDEO	UIDENT	>	*veggio*	? = *veggiono*
TACEO	TACENT	>	*taccio*	? = *tacciono*
PLACEO	PLACENT	>	*piaccio*	? = *piacciono*
UALEO	UALENT	>	*vaglio*	? = *vagliono*

TABLE 6   Palatalized alternants (or their reflexes) in central Italian first person singular and
           third person plural present indicative

Norcia			Rieti		
1SG.PRS	3PL.PRS		1SG.PRS	3PL.PRS	
ˈvɛŋgo	ˈvjɛŋgu	'come'	ˈbɛŋgo	ˈbeŋgu	'come'
ˈvɔjjo	ˈvwojju	'want'	ˈbɔʎo	ˈboʎu	'want'
ˈsattʃo	ˈsattʃu	'know'	ˈpɔttsə	ˈpottsu	'can'
ˈvejjo	ˈvijju	'see'	ˈmɔro	ˈmoru	'die'
ˈbeo	ˈbiu	'drink'	ˈedo	ˈidu	'see'
ˈkɔjjo	ˈkwojju	'gather'			

Ascrea		
1SG.PRS	3PL.PRS	
ˈwɛŋgo	ˈweŋgu	'come'
ˈtɛŋgo	ˈteŋgu	'hold'
ˈpɔttso	ˈpottsu	'can'
ˈwɔʎo	ˈoʎʎu	'want'

singular in -o, and, (b) -u caused metaphonic raising or diphthongization of
the stressed vowel in the third person plural (cf. Merlo, 1909; Loporcaro and
Paciaroni, 2016: 236). Typical examples are given by *AIS* maps 1661, 1695, 1696,
1699, for points 576 (Norcia, in Umbria) and 624 (Rieti, in Lazio), and from
Fanti (1939) for the dialect of Ascrea (Table 6).

Such examples can be multiplied *ad libitum* for dialects of this area, and
they show quite clearly that the presence of the palatal alternant in the third
person plural present indicative cannot be an effect of a general pattern of
identity between that form and the first person singular, as Elson claims, since

no such identity ever existed in the relevant dialects. Final -u as a third person plural marker is a direct reflex of Latin -UNT, and its presence is detectable over vast areas of the central and southern Italo-Romance domain; indeed it may once have been general in Italo-Romance. As for metaphony, it is a phenomenon detectable practically throughout Italo-Romance (see, e.g., Maiden, 1991; 2016c), and clearly one of great antiquity. The allegedly 'obvious' pattern discernible in Italian is a local feature of Tuscan varieties, and cannot plausibly be projected onto 'Proto-East-Romance'.

That the specific explanation of the 'Italian' pattern provided by Elson cannot be correct does not, of course, necessarily disprove the notion that 'Italian' might yet somehow represent the 'Proto-East-Romance' norm, from which Romanian allegedly deviates. However, and once again, a cursory glance at the comparative dialectological data beyond Italian (or Tuscan) is enough to reveal a truly fatal problem for the claim that the particular distributional pattern of Romanian is a deviation from the norm. Elson's analysis is simply back-to-front: it is Italian that is 'abnormal', while the pattern found in Romanian is actually normal not just for 'East' Romance, but for *the Romance languages generally*.

The Romance languages virtually never show YE palatalization in the root-final consonant of the third person plural present indicative (regardless of whether the ending continues Latin -ENT or -UNT), except in Tuscany, Lazio, Umbria, and the Marche, where an (originally) palatalized alternant regularly appears (or appeared historically) before the reflex of -UNT. In Italo-Romance the continuant of -UNT may be discernible as an ending conserving a back vowel (-o(n)-, -u(n)-, etc.), and/or in a metaphonically raised or diphthongized vowel showing the effects of an original *-u(n). Using these criteria, while we find the palatalized alternant or its reflex in the first person singular present indicative, there is no sign of it in the third person plural present indicative in, for example, *AIS* maps 1691, 1693, 1694, 1695 (data in Table 7a-e) or Vignoli (1911; 1920; 1925) for the Lazio dialects of Amaseno, Veroli, and Castro dei Volsci (data in Table 7). The parallels with the distribution found in Romanian are also given for the relevant verbs.

TABLE 7   Yod effect palatalization in the first person plural present indicative, but not in
            the third person plural present indicative, in central and southern Italo-Romance
            dialects

(a) Verb 'to come'		
	1SG.PRS	3PL.PRS
Scanno, Abruzzo	ˈvjeŋgə	ˈvjeːnənə
Trevico, Campania	veŋk	ˈvjennə
Vernole, Puglia	ˈɛɲu	ˈɛːnune
San Chirico Raparo, Basilicata	ˈveŋgu	ˈvjɛːninu
Mistretta, Sicily	viˈeɲu	viˈeːnu

cf. old Romanian	viiu	vinu
(b) Verb 'to want'		
	1SG.PRS	3PL.PRS
Ruvo, Puglia	ˈvwɔɟɟə	ˈvwɔlənə
Vernole, Puglia	ˈɔʎu	ˈɔːlune
Catenanuova, Sicily	ˈvwoɟɟu	ˈvwolinu

cf. old Romanian	voiu	voru
(c) Verb 'to know'		
	1SG.PRS	3PL.PRS
Ruvo	ˈsattʃə	ˈsapənə
Carovigno	ˈsattʃu	ˈsapunu
Vernole	ˈsattʃu	ˈsapune
(d) Verb 'to do'		
	1SG.PRS	3PL.PRS
Carovigno	ˈfattsu	ˈfakunu
Vernole	ˈfattsu	ˈfakune

cf. old Romanian	(facu)	facu
(e) Verb 'to see'		
	1SG.PRS	3PL.PRS
Serracapriola, Puglia	ˈvajə	ˈvedənə
Trevico, Campania	veχ	ˈvirənə
Avetrana, Puglia	ˈɛʃu	ˈɛtinu
Acri, Calabria	ˈviju	ˈviðuði
Catenanuova, Sicily	ˈviju	ˈviðinu

cf. old Romanian	văzu	vădu
(f) Verb 'to please'		
	1SG.PRS	3PL.PRS

TABLE 7   Yod effect palatalization in the first person plural present indicative, but not in
the third person plural present indicative, in central and southern Italo-Romance
dialects (*cont.*)

	1SG.PRS	3PL.PRS
Amaseno, Lazio	ˈpjatʃə	ˈpjakənə
Veroli, Lazio	ˈpjatʃo	ˈpjatʃənə / ˈpjakunə[23]
Castro dei Volsci, Lazio	ˈpjatʃə	ˈpjakunə / ˈpjatʃənə
cf. old Romanian	(*placu*)	*placu*

The situation is overwhelmingly the same for the Romance varieties of
northern Italy and for the Gallo-Romance[24] and Ibero-Romance domains. The
AIS data-sets mentioned above, for example, show that palatalized alternants
do not generally occur in the present indicative third person plural, whether
that form continues Latin -ENT or -UNT.

In sum, judicious assessment of easily available comparative evidence
makes it plain that the pattern of alternation purportedly peculiar to
Romanian within the alleged 'Proto-East-Romance' family is not in the least
limited to Romanian: it is actually the general Romance pattern. What really
needs explaining is the real 'deviant'–Italian (and neighbouring central Italo-
Romance dialects): for a partial explanation of the Italian facts, see Maiden
(2020). The fact that there is nothing special about the Romanian pattern fun-
damentally undermines Elson's assumption that this pattern must be explica-
ble by something distinctive of Romanian, namely contact with Slavonic. Of
course it remains conceivable that Bulgarian influence somehow 'reinforced'
the Romance pattern in Romanian, even if it did not cause it, but the insur-
mountable problem is that we simply cannot tell.

---

23   Cf. also old Neapolitan *piacuno* (Ledgeway, 2009: 376f.).
24   See, e.g., Anglade (1921: 289, 333, 345, 347f., 351f.) for old Occitan. *AIS* map 1695 shows
possible evidence for this type in the reflex of UENIUNT 'they come' in the Franco-
Provençal of Brusson and Ronco Canavese, but there are no other examples. There is
also some occasional evidence from the Romansh dialects of the lower Engadine for
the iotacized alternant in third person plurals (see Decurtins, 1958: 133, 145, 186). Two
otherwise completely isolated exceptions from the *AIS* showing apparent YE palatalization
in the third person plural present indicative are Acri (Calabria), with 1SG.PRS.IND ˈvieɲu
'come' ~ 3PL.PRS.IND ˈvjeɲinu, and Spinazzola (Puglia), with 1SG.PRS.IND ˈvajʃə 'see' ~ 3PL.
PRS.IND ˈvɛʃənə.

Finally, there is a suggestive fragment of philological evidence. Mihăescu (1960: 142, 244) finds *facunt*, for expected *faciunt*, in an inscription from Aquincum (Pannonia Inferior). Unfortunately, we lack the corresponding first person singular present indicative or the present subjunctive forms–so we cannot tell for sure whether *facunt* was part of exactly the alternation *pattern* found in Romanian[25]–but if Mihăescu is correct in identifying it as the precursor of Romanian *fac(u)*, then the hypothesis that the latter reflects Slavonic influence becomes untenable, given that no datable inscription from Aquincum is later than AD 377 (Mihăescu, 1960: 27), and that the town is known to have been abandoned by the beginning of the fifth century.

4 The Romanian Alternation Patterns are Less Like those of Middle Bulgarian than they Seem

There remains a further problem, reflecting my criterion (ii), above, in that the Romanian data actually seem unlike the Bulgarian data in a respect that Elson effectively passes over.[26] Exactly the same paradigmatic distributional pattern resulting from iotacization is to be found in old Romanian verbs with original root final -n and -l, as well as those in root-final dentals: e.g., *viu* (< *'venjo) 'I come' (3PL *vinu*), *saiu* (< *'saljo) 'I jump' (3PL *saru*). Despite Elson's claim to the contrary, radical-final -n, at least, occurs in Romanian anything but infrequently. He does acknowledge this iotacization of nasals and laterals, among other consonants, (2017: 851 n18), only to dismiss it from consideration, claiming that the details 'do not contribute to [his] discussion'. One

---

25   Expected prevocalic I is absent in a number of other words in inscriptions from this area (Mihăescu, 1960: Section 67), so we cannot be certain that *facunt* is a case of *particular* absence of the expected outcome in the third person plural present indicative. See also Ionescu-Ruxăndoiu (2018: 182).

26   They are also different in respect of their paradigmatic distribution, in that the Romanian alternants at issue also occur in the third person forms of the present subjunctive and in the gerund, categories for which no direct counterpart can be found in middle Bulgarian. These facts are not, however, necessarily incompatible with the hypothesis of middle Bulgarian influence, given the robust and recurrent patterns of identity between the relevant present-tense alternants and the subjunctive and gerund which independently exist in Romanian (as Elson, 2017: 872f. acknowledges, at least as far as the subjunctive is concerned), which could have served as templates for the paradigmatic distribution of the 'Bulgarian' alternation type if it originated in the present tense. On this, the reader is further referred to Elson (2017: 871–878), who appears however unaware of different analyses of these phenomena by Maiden (1996a; 2011a,b; 2013a); see also Maiden (2018: 99f.).

reason given is that 'in the case of nasals and liquids, palatalization occurs in conjunction with other changes which eliminated its reflex'. But if this outcome, phonologically very different from what we see in middle Bulgarian, nonetheless displays the paradigmatic distribution at issue, then we have a serious potential counterexample to Elson's claim of Bulgarian influence on the paradigmatic outcome in Romanian, and the facts should therefore have been carefully examined. As it happens, those facts actually could be reconciled with the 'Bulgarian' hypothesis. The historically underlying alternants in Romanian are palatal consonants [ɲ] and [ʎ] (see, e.g., Sala, 1976: 228–230), and these might, in principle, have been associated by speakers with the middle Bulgarian palatalized nasals and liquids which also occurred in the relevant morphological environments (cf. Leskien, 1922: 52 for old Bulgarian). However that may be, here we see, once again, the need for due diligence in exploring of all the available comparative-historical data.

5   The Type *fa*[k] *fa*[ʧ]*e fa*[k] *fa*[k]*ă* can be Internally Motivated

Elson (2017: 888) criticizes 'a strong tendency among linguists to interpret such compositional changes as the result of non-phonetic innovation in the form of *analogy*, i.e., as the result of system-internal innovation with no regard to the defensibility of its invocation (i.e., whether there was an existing pattern which might have served as the basis for the changes in phonemic composition)'. In the absence of such an existing pattern, he says, we may legitimately appeal to contact as an explanation. This is a mistake: the onus remains on the proponent of a contact-based explanation convincingly to rule out the possibility of an internal motivation, and in the case of the alternation type in (Table 8) that is not so easy to do.

The alternation-type in Table 8 is merely *similar to* Bulgarian, while it is *exactly like*, in phonological content and paradigmatic distribution,[27] an existing, native, Daco-Romanian pattern. This latter is entirely attributable to the effects of regular Romance palatalization and affrication of velars before front vowels. Moreover, the pattern seen in Table 8 is a typically Romance development, with counterparts in other Romance languages. Consider, first, the Romanian verbs *a zice* 'say' and *a merge* 'go', where the alternation [k] ~ [ʧ], and [g] ~ [ʤ] are regular results of sound change (Table 9).

---

27   At least in respect of the present indicative and the subjunctive. There are some differences in other tenses but it is, after all, the present which is the focus of Elson's attention and the locus of alleged middle Bulgarian influence.

TABLE 8   Velar palatal alternations in the present tense of the Romanian verb

1SG.PRS.IND	3SG.PRS.IND	3PL.PRS.IND	3SG/PL.SBJV
fa[k] 'do'	*fa*[ʧ]*e*	*fa*[k]	*fa*[k]*ă*
fu[g] 'flee'	*fu*[ʤ]*e*	*fu*[g]	*fu*[g]*ă*

TABLE 9   Regular morphological consequences of Latin palatalization of velars before front
            vowels in the Romanian present and subjunctive

Latin					Romanian			
1SG.PRS	3SG.PRS	3PL.PRS	3SBJV		1SG.PRS	3SG.PRS	3PL.PRS	3SBJV
DICO	DICIT	DICUNT	DICA(N)T	> *zi*[k]	*zi*[ʧ]*e*	*zi*[k]	*zi*[k]*ă*
MERGO	MERGIT	MERGUNT	MERGA(N)T	> *mer*[g]	*mer*[ʤ]*e*	*mer*[g]	*mear*[g]*ă*

Taking into account the special behaviour of the third person plural present, as discussed in Section 5, this pattern for palatalization of velars before front vowels is exactly the same as that historically predictable for YE. Both types of verb have in common the alternations 3SG [ʧ] ~ 3PL [k] and 3SG [ʤ] ~ 3PL [g], as shown in Table (10). It is therefore not true to say, as Elson (2017: 868) does, that there exists for such a development 'no motivation internal to Romanian as a function of the structural details of its conjugation'. The relevant motivation simply does exist, independently of any possible 'Bulgarian' influence. Nor does it seem valid to assert that 'we cannot claim that verbs of the E conjugation and the Ĕ/I complex, in which a reflex of yotation is expected but absent (e.g., *tăc-ea*), were analogized to those of the Ĕ conjugation in which a reflex is not expected (e.g., *zic-e*) because this would leave us without motivation for parallel evolution among verbs with radical-final dental (e.g., *ved-ea*, etc. were not analogized to *cred-e* etc.).' Elson's formulation is obscure, but he appears to mean *lack* of parallel evolution in verbs such as *vedea*, because verbs with the root-final dentals do not show the development alleged to have affected verbs with root-final velars. Yet this argument would only apply by assuming that the relevant changes predated the emergence of the palatal~velar alternations mentioned above, which are unparalleled in verbs with root-final dentals.

Even if there is no need to appeal to putative 'middle Bulgarian' influence in this case, might one not still invoke it to explain the *direction* of the analogical

change?[28] Why do we not get 1SG.PRS **faț ~ 3SG.PRS face and 1SG.PRS **ziț ~ 3SG.PRS zice rather than the actually occurring fac ~ face and zic ~ zice– that is, why do we not get a situation in which the historically expected changes effected by yod occur in reflexes of FACIO, and are then analogically extended to verbs such as DICO? Could this be because this latter pattern does not have a counterpart in Bulgarian, while the opposite and actually occurring one does occur there? Such an explanation might have some traction (although it would be impossible to prove), were it not for the fact that what we observe in Romanian is also observed independently in other Romance languages while, to the very best of my knowledge, the putative opposite development considered above is found absolutely nowhere in Romance. For more extensive illustration of these facts see, e.g., (Maiden, 2018: 93–122). Consider, for example, the velar of Spanish 1SG.PRS.IND hago ~ 1/3SG.PRS.SBJV haga vs the historically underlying yod whose effects are still seen in Portuguese faço ~ faça (< Lat. FACIO FACIAT), Italian 1SG.PRS.IND fuggo ~ 3PL.PRS.IND fuggono ~ 1/3SG.PRS.SBJV fugga with the velar, for older fuggio ~ fuggiono ~ fuggia showing the historically regular alternation (FUGIO FUGIUNT FUGIAT). The Romanian substitution of velar for affricate alternants in such cases conforms to a widespread Romance development. Yes, it rather resembles Bulgarian, but how can we know that the resemblance is anything other than coincidental?

6   The 'Cluster' Criterion

We may add to the three mentioned in Section 2.2 a fourth criterion of plausibility in arguing for contact effects. The verisimilitude of the claim that some phenomenon X in language R is due to contact with language B increases if there are in R several other phenomena of a similar kind that incontrovertibly come from B; that is, if X is part of a 'cluster' of similar phenomena. The 'Bulgarian hypothesis' regarding YE could be seen as plausible in a general way because it is not the only alleged Bulgarian 'organizational' influence on Romanian (see, e.g., Elson, 1994; 1999), and this is why Elson argues (2017: 890) that '[w]e may therefore see, in support of Bulgarian influence in the Romanian present system, a general tendency for the importation of Bulgarian organization into Romanian verbal morphology […]'. But what is the status of the other cases which constitute this 'general tendency'? I limit myself to morphological

---

28   Note Thomason's observation that 'a growing body of evidence suggests that multiple causation – often a combination of an external and one or more internal causes – is responsible for a sizable number of changes' (Thomason, 2010: 32).

TABLE 10   Morphological parallels between the effect of yod and the palatalization of velars
before front vowels in Romanian

Latin					Romanian			
1SG.PRS	3SG.PRS	3PL.PRS	3SBJV		1SG.PRS	3SG.PRS	3PL.PRS	3SBJV
DICO	DICIT	DICUNT	DICA(N)T	>	*zi*[k]	*zi*[tʃ]*e*	*zi*[k]	*zi*[k]*ă*
MERGO	MERGIT	MERGUNT	MERGA(N)T	>	*mer*[g]	*mer*[dʒ]*e*	*mer*[g]	*mear*[g]*ă*
FACIO	FACIT	FACIUNT	FACIA(N)T	>	(*fa*[k])	*fa*[tʃ]*e*	*fa*[k]	*fa*[k]*ă*
FUGIO	FUGIT	FUGIUNT	FUGIA(N)T	>	(*fu*[g])	*fu*[dʒ]*e*	*fu*[g]	*fu*[g]*ă*

phenomena in Romanian which can be viewed as 'morphomic', in that they involve characteristics of the abstract organization of inflexional paradigms to which no coherent function or meaning can be ascribed.[29]

Elson (2017: 871) argues that another reflexion of Bulgarian influence in Romanian is that here, unlike (nearly) all other Romance languages, root-final velars in first conjugation verbs undergo phonologically regular palatalization before inflexional front vowels. Thus, while Romanian, Italian, and Spanish all behave in the same way in showing the expected alternations arising from palatalization in non-first conjugation verbs (e.g., 3SG.PRS.IND Ro. *zi*[tʃ]*e*, It. *di*[tʃ]*e*, Sp. *di*[θ]*e* 's/he says' vs 3SG.PRS.SBJV Ro. *zi*[k]*ă*, It. *di*[k]*a*, Sp. *di*[ɣ]*a*), in the first conjugation only Romanian shows the type of alternation expected on historical phonological grounds, while other Romance languages do not (e.g., 3SG.PRS.IND Ro. *joa*[k]*ă*, It. *gio*[k]*a*, Sp. *jue*[ɣ]*a* 's/he plays' vs 3SG.PRS.SBJV Ro. *joa*[tʃ]*e*, It. *gio*[k]*i*, Sp. *jue*[ɣ]*e*). For Elson, these facts are clear evidence of Bulgarian influence, because Bulgarian, unlike the sister Romance languages of Romanian, imposes no conjugational restriction on palatalization. Yet in fact, Romanian merely behaves in the normal and historically predictable way, and there is nothing to be 'explained'. At most one might say that this differentiated behaviour on the part of Romanian is yet another linguistic manifestation of the early isolation of its speakers from the main body of Romance languages

---

29   For discussion of another possible candidate for Slavonic morphological influence on the organization of Romanian morphology, namely the truncation of the infinitive such that the 'short' form has the value of a verb, and the 'long' form that of a noun, see, e.g., Petrucci (1999: 128f.). The truncation itself is a widespread Romance phenomenon, and cannot be attributed to any parallel development in Bulgarian. Petrucci suggests that the Romanian pattern was mapped onto the meaning of the Bulgarian distinction between 'long' and 'short', but here we are dealing with concrete expression of distinctions of grammatical meaning, rather the kind of more abstract organizational phenomenon with which this study is concerned.

– an isolation in which the Slavic incursions, rather than the Slavonic languages themselves, did of course play a role. What actually needs to be explained is why the other Romance languages do impose a conjugational restriction on palatalization (see, on this, Maiden, 2018: 277–283), not why Romanian does not. In fact, this behaviour of the Romanian first conjugation is not the only respect in which Romanian differs from other Romance languages in the sense of a greater phonological propensity to palatalize velars. This is not the place to explore the issues, but for discussions see Skok (1926), Merlo (2014), or Maiden (2019: 110f.).

Another case of arguably 'morphomic' influence of Bulgarian on Romanian is addressed in Elson (1999). He observes that Romanian retains from Latin a morphologically distinct subjunctive form only in the third person, the subjunctive forms of all other persons having been replaced (in nearly all verbs) by those of the present indicative (Table 11). Elson attributes this fact to structural influence from Bulgarian. But Bulgarian does not have a morphological subjunctive, and so he actually locates the similarity with the Romanian *subjunctive* in the Bulgarian *imperative* (and more precisely in the first conjugation imperative). Note that the subjunctive marking in Romanian involves a kind of 'reversal' of the inflexional ending such that the subjunctive ending of first conjugation verbs (such as *a zice* 'to say') is identical to the third person singular ending of the present of *a cânta*, and vice versa. Given the complexity of Elson's analysis, it is easiest to quote from it verbatim (Elson, 1999: 147).

> The characteristics of the imperative – a single form unmarked morphologically for person/number, and canonically identical to forms of the present but opposed to them by a realization of the nonterminal suffix characteristic of the present of another conjugation – are exactly those characteristics of the Romanian subjunctive with respect to the indicative. Thus the relationship of *veda* [an exemplar of the Bulgarian first conjugation, MM] to its imperative in Bulgarian is identical to that of all verbs to their subjunctive in Romanian [...]. When, to these formal identities, we add that, in each language, one of the paradigms in question (i.e., the present) is indicative, and the other (i.e., the subjunctive in Romanian and the imperative in Bulgarian) is modal, and that, in each language, the indicative and modal paradigms are not only desinentially identical to each other, comprising a nonterminal, vocalic suffix which may be followed by a terminal consonantal one, but that the present and imperative in each are opposed to past paradigms in lacking an overt marker of tense, which the past paradigms possess, *the possibility of innovation*

TABLE 11   Limitation of distinctive present subjunctive morphology to the third person in Romanian

Latin						
	1SG	2SG	3SG	1PL	2PL	3PL
IND	CANTO 'sing'	CANTAS	CANTAT	CANTAMUS	CANTATIS	CANTANT
SBJV	CANTEM	CANTES	CANTET	CANTEMUS	CANTETIS	CANTENT
IND	DICO 'say'	DICIS	DICIT	DICIMUS	DICITIS	DICUNT
SBJV	DICAM	DICAS	DICAT	DICAMUS	DICATIS	DICANT
Romanian						
	1SG	2SG	3SG	1PL	2PL	3PL
IND	cânt	cânți	**cântă**	cântăm	cântați	**cântă**
SBJV	cânt	cânți	**cânte**	cântăm	cântați	**cânte**
IND	zic	zici	**zice**	zicem	ziceți	**zic**
SBJV	zic	zici	**zică**	zicem	ziceți	**zică**

internal to Romanian as an explanation for the loss of the personal forms in its subjunctive ceases to be a serious alternative. There can be little doubt that we are justified in claiming that Romanian calqued the formal pattern of relationship between the present indicative and the imperative of first conjugation verbs in Bulgarian by adopting the single characteristic which distinguished it from Bulgarian in this regard, i.e., a reduced modal paradigm comprising only a third person form. It did this by abandoning personal forms of the plural subjunctive. [My emphasis]

Elson's analysis contains an obvious *petitio principii*, since the very thing which needs to be demonstrated, namely Bulgarian influence, is tacitly assumed. The assumption seems to be that if no internal motivation for some phenomenon is detectable, then it is legitimate to invoke language contact. The task then becomes, in this case, one of finding *something* in Bulgarian that–in the absence of anything that really resembles the Romanian situation–looks at least rather like the Romanian facts and might just about explain them, given sufficient *ad hoc* adjustments. Yet the onus on anyone appealing to language contact in order to explain some linguistic change is first to exclude beyond reasonable doubt the possibility of an internal development, and this Elson does not do. The presence of forms identical to the indicative in three of the four Romanian paradigm-cells at issue is eminently consistent with much more widespread Romance developments. There is in fact an extensive Romance tendency for indicative forms to replace present subjunctive forms, particularly in the first and second persons plural, and this is systematically

observable in many Italo-Romance varieties,[30] as demonstrated in detail by, for example, Maiden (2010: 133–135; 2012: 33–35, 45–47). Romanian second person singular subjunctives are overwhelmingly identical[31] to the corresponding present indicative forms by virtue of ending in -*i* (e.g., 2SG.PRS. IND *vezi* 'see' = 2SG.SBJV *vezi*; 2SG.PRS.IND *dormi* 'sleep' = 2SG.SBJV *dormi*; cf. Latin UIDES – UIDEAS, DORMIS – DORMIAS). Elson interprets this as evidence of replacement of the subjunctive form by that of the present indicative, but one could equally say that Romanian has generalized -*i* as second person singular marker in both moods and all tenses (only certain imperatives escape this generalization). Elson (1999: 145) is aware of this possibility, but argues against it on the grounds that 'Italian'–this language again being somehow elevated to the undeserved status of proxy for all eastern Romance–has a second person singular present subjunctive ending in -*a* (e.g., 2SG.PRS.SBJV *veda* 'see'; *dorma* 'sleep'), not in -*i*. But here he is anachronistically projecting onto an entire branch of the Romance languages a relatively recent innovation of Italian. Historically, Italian itself with much of the rest of Italo-Romance had a second person singular present subjunctive ending -*i* (or -*e*), not -*a* (see, e.g., Rohlfs, 1968: 296f., 299–301; Maiden, 1995: 129; 1996b: 161). Once again, the drawbacks of too limited a historical and comparative perspective are all too evident.

Only the behaviour of the first person singular, in the above case, remains as not obviously explicable in terms of these internal (Romance and Romanian) developments, but in a system where both second person forms, and one first person form, are already identical to the indicative, it is not surprising that the first person singular may follow suit. If the third person is more resistant, this fact can probably be attributed to the very high frequency of occurrence of the third person (yet note that in some Daco-Romance varieties such as Istro-Romanian even the third person subjunctive is replaced by indicative forms: cf. Kovačec, 1971: 150). If appeal to internal developments may not deliver a completely satisfying account of the morphology of the Romanian subjunctive, it very nearly does so – and this is enough to make any appeal to Bulgarian influence look very shaky. In any case, the entire argument from the Bulgarian imperative–in fact solely from the Bulgarian *first conjugation* imperative–is implausible, perhaps most of all because one might expect the Bulgarian imperative to map functionally onto its obvious counterpart, namely the Romanian imperative. Since imperative and subjunctive morphology are clearly distinct in Romanian, it is hard to see why structures specific to imperatives would be

---

30   Elson (1999: 144) acknowledges this fact, yet dismisses its significance.

31   In only very few verbs is there a subjunctive root allomorph distinct from the indicative.

taken to apply instead to subjunctives. Appeal to Bulgarian influence in this case does not in fact seem anything like a 'serious alternative', and it is certainly not such if we take into account my criterion (ii) in Section 2.2, namely that cogent claims for contact effects will involve structures in both of the languages in contact which are maximally similar. To say that there is 'little doubt' of Bulgarian influence is wrong.[32]

7   Conclusion

I am very open to the proposition that 'realizational', or more specifically, 'morphomic', patterns of allomorphy might be borrowed from one language to another. If the Romanian data could have been made to show this, a significant and original addition would have been made to the theory of morphological change in diachrony. But if theoretically significant claims are to be made about the role of contact in morphological change, and are to have scientific validity, suitably rigorous conditions must be met in order to guarantee that the phenomena at issue are not independent of contact. Alas, the relevant phenomena as discussed in Elson (2017) turn out not to be in the least unique to Romanian among the Romance languages, and 'Bulgarian' influence remains therefore at best undemonstrable. Romanian morphology unquestionably is, in some respects, 'Slavonic', but it remains to be convincingly shown that it is in any way 'Slavonic' at the level of *purely morphological* paradigmatic organization of the kind in which I am interested.

---

32   I have not mentioned Elson's argument (Elson, 1994) that Romanian has calqued from Bulgarian 'a characteristic organizational feature of many Bulgarian verbs: morphological relationship between the preterite and the participle (i.e., the occurrence in each of a morpheme expressing the grammatical meaning common to them)' (Elson, 1994: 27). It needs to be said that the tendency analogically to ascribe the same stem allomorph to both the preterite and the past participle is far from unique to Romanian among the Romance languages (see, e.g., Meyer-Lübke, 1895: 372, 413, 419, 426; Nyrop, 1960: 127, 77f., 87, 140; Rohlfs, 1968: 369f., 373–375). Nonetheless, this tendency does seem to me far more extensive and systematic in Romanian–where nearly all preterites and past participles have the same stem–than elsewhere in Romance. Could Bulgarian influence be at work here? It might be (perhaps 'reinforcing' a general Romance tendency?). However, from the point of view of the particular kind of phenomenon I am interested in here, namely potential transfer of 'morphomic' phenomena by contact, this case does not qualify. The distribution of the relevant stem allomorph in Romanian actually *is* morphomic (for reasons set out at length in Maiden (2013b) – for example that the supine is also affected), but as I understand it the original relevant pattern in Bulgarian would not qualify as 'morphomic' in my sense, since it expressed a 'common grammatical meaning' (namely perfectivity) present in both the preterite and the participle.

Some other general methodological points also emerge from the foregoing. One is that absence of evidence should not be taken for evidence of absence. There is a frustratingly large amount of the former in Romanian–given that the first thousand years or so of the history of the language are veiled in obscurity–and it is rash to appeal too readily to contact as an explanation of phenomena whose internal context and motivation we cannot immediately discern (cf. Elson, 1999: 146; 2017: 868, on this point).[33] We also urgently need to form a clear idea of the historical circumstances of any linguistic contact–something which in the case of Romanian is again difficult. The contact effects in Romanian morphology identified by Elson are of such structural complexity and abstractness that I at first assumed that his argument must be that these effects emerged in the minds of native bilinguals or, possibly, in the minds of adult Bulgarian speakers acquiring Romanian. Surprisingly, however, Elson (1999: 151) argues that the changes he describes in Romanian subjunctive morphology must have been made by Romanian-speaking non-native 'borrowers' of Bulgarian forms, who had an apparently imperfect knowledge of Bulgarian paradigmatic morphology. This idea is, to say the least, puzzling: why would non-Bulgarian-speaking Romanian speakers set about remodelling their native (and differently structured) inflexional morphology on the basis of partial similarities with the structural idiosyncrasies of a language they did not really know? The fundamental point here is that if we are to assess possible Slavonic influences on Romanian morphology, we first need a clearer understanding than we presently have of the circumstances and dynamics of the contact (cf. also Elson, 2017: 891).

I have repeatedly insisted in this study that for a cogent demonstration of assumed effects of language contact we need a properly detailed comparative and historical knowledge of all the languages in play. This exposes me to the justifiable charge that I myself have an inadequate grasp of the history of Bulgarian and other Slavonic languages. I have depended here almost entirely on Elson's account of the historical Bulgarian facts, which I have been happy to take on trust. I have no doubt, however, that my analysis would have benefited from a proper training in the Slavonic data. In reality, very few linguists working on contact can ever hope to have equally deep comparative-historical knowledge of languages belonging to different language-families (and whatever the defects of Elson's arguments in this case, his ability to span Slavonic and Romance remains unusual and admirable). I suggest, therefore, that for work on contact between languages belonging to different families to be truly

---

33   On the other hand we should not assume, either, that the apparent availability of an internal motivation automatically rules out an effect of contact, where a robust case for the latter could be constructed.

of value it will not only have to satisfy the general criteria of plausibility I have set out here, but it will also be the result of close *collaboration* between scholars with complementary expertise in the relevant languages.

References

AIS = Jaberg, Karl and Jakob Jud. 1928-1940. *Sprach- und Sachatlas Italiens und der Südschweiz*. Zofingen: Ringier.

Anglade, Joseph. 1921. *Grammaire de l'ancien provençal*. Paris: Klincksieck.

Aronoff, Mark. 1994. *Morphology By Itself*. Cambridge, MA: MIT Press.

Ascoli, Graziadio. 1881-1882. Lettere glottologiche. Prima lettera. *Rivista di Filologia e d'Istruzione Classica* 10: 1–71.

Bossong, Georg. 2016. Classifications. In Adam Ledgeway and Martin Maiden (eds.), *The Oxford Guide to the Romance Languages*, 63–72. Oxford: Oxford University Press.

Decurtins, Alexi. 1958. *Zur Morphologie der unregelmässigen Verben im Bündnerromanischen*. Bern: Francke.

Dragomirescu, Adina. 2015. Are there Slavonic features in the syntax of Romanian? Two case studies. *Diacronia* 1: 1–13.

Elson, Mark. 1994. Slavic influence on the organization of Romanian conjugation. *Indiana Slavic Studies* 7: 27–35.

Elson, Mark. 1999. Slavic influence in the evolution of the Romanian subjunctive. *Zeitschrift für Balkanologie* 35: 139–152.

Elson, Mark. 2017. On the history of radical-final palatalization in Romanian conjugation: The interaction of phonetic phenomena with non-phonetic. *Zeitschrift für romanische Philologie* 133: 844–892.

Fanti, Renata. 1939. Note fonetiche e morfologiche sul dialetto di Ascrea. *L'Italia dialettale* 15: 101–133.

Friedman, Victor. 2009. The diffusion of Macedonian inflections into Megleno-Romanian: A reconsideration of the evidence. In Steven Franks , Vrinda Chidambaram , and Brian Joseph (eds.), *A Linguist's Linguist. Studies in South Slavic Linguistics in Honor of E. Wayles Browne*, 223–233. Bloomington: Slavica.

Gardani, Francesco. 2018. On morphological borrowing. *Language and Linguistics Compass 12*. DOI: 10.1111/lnc3.12302.

Gardani, Francesco, Peter Arkadiev, and Nino Amiridze. 2015. Borrowed morphology: An overview. In Francesco Gardani , Peter Arkadiev , and Nino Amiridze (eds.), *Borrowed Morphology*, 1–23. Berlin: De Gruyter Mouton.

Harris, Martin. 1988. The Romance languages. In Martin Harris and Nigel Vincent (eds.), *The Romance Languages*, 1–25. London: Routledge.

Hurren, Antony. 1969. Verbal aspect and archi-aspect in Istro-Romanian. *La Linguistique* 2: 59–90.

Ionescu-Ruxăndoiu, Liliana. 2018. In Marius Sala and Liliana Ionescu Ruxăndoiu (eds.), *Istoria limbii române* I, 170–188. Bucharest: Univers Enciclopedic Gold.

Kossmann, Maarten. 2015. Contact-induced change. In Matthew Baerman (ed.), *The Oxford Handbook of Inflection*, 251–271. Oxford: Oxford University Press.

Kovačec, August. 1971. *Descrierea istroromânei actuale*. Bucharest: Editura Academiei.

Ledgeway, Adam. 2009. *Grammatica diacronica del napoletano*. Tübingen: Niemeyer.

Leskien, Adam. 1922. *Handbuch der altbulgarischen* (*altekirchenslavischen*) *Sprache*. Heidelberg: Winter.

Loporcaro, Michele and Tania Paciaroni. 2016. The dialects of central Italy. In Adam Ledgeway and Martin Maiden (eds.), *The Oxford Guide to the Romance Languages*, 228–245. Oxford: Oxford University Press.

Maiden, Martin. 1991. *Interactive Morphology. Metaphony in Italy*. London: Routledge.

Maiden, Martin. 1995. *A Linguistic History of Italian*. London: Longman.

Maiden, Martin. 1996a. The Romance gerund and system-dependent naturalness in morphology. *Transactions of the Philological Society* 94: 167–201.

Maiden, Martin. 1996b. On the Romance inflectional endings *-i* and *-e*. *Romance Philology* 50: 147–182.

Maiden, Martin. 2010. Riflessioni comparative e storiche sulla sorte del congiuntivo presente nelle varietà italoromanze. In Giovanni Ruffino and Mari D'Agostino (eds.), *Storia della lingua italiana e dialettologia*, 129–149. Palermo: Centro di studi filologici e linguistici siciliani.

Maiden, Martin. 2011a. Morphological persistence. In Martin Maiden , John Charles Smith , and Adam Ledgeway (eds.), *The Cambridge History of the Romance Languages*. I. *Structures*, 155–215, 699–706. Cambridge: Cambridge University Press.

Maiden, Martin. 2011b. Morphophonological innovation. In Martin Maiden , John Charles Smith , and Adam Ledgeway (eds.), *The Cambridge History of the Romance Languages*. I. *Structures*, 216–267, 706–713. Cambridge: Cambridge University Press.

Maiden, Martin. 2012. A paradox? The morphological history of the Romance present subjunctive. In Sascha Gaglia and Marc-Olivier Hinzelin (eds.), *Inflection and Word Formation in Romance Languages*, 27–54. Amsterdam: John Benjamins.

Maiden, Martin. 2013a. Semi-autonomous' morphology? A problem in the history of the Italian (and Romanian) verb. In Silvio Cruschina , Martin Maiden , and John Charles Smith (eds.), *The Boundaries of Pure Morphology: Diachronic and Synchronic Perspectives*, 24–44. Oxford: Oxford University Press.

Maiden, Martin. 2013b. The Latin 'third stem' and its Romance descendants. *Diachronica* 30: 492–530.

Maiden, Martin. 2016a. Romanian, Istro-Romanian, Aromanian, Megleno-Romanian. In Adam Ledgeway and Martin Maiden (eds.), *The Oxford Guide to the Romance Languages*, 91–125. Oxford: Oxford University Press.

Maiden, Martin. 2016b. Dalmatian. In Adam Ledgeway and Martin Maiden (eds.), *The Oxford Guide to the Romance Languages*, 126–138. Oxford: Oxford University Press.

Maiden, Martin. 2016c. Italo-Romance metaphony and the Tuscan diphthongs. *Transactions of the Philological Society* 114: 198–232.

Maiden, Martin. 2018. *The Romance Verb. Morphomic Structure in Diachrony*. Oxford: Oxford University Press.

Maiden, Martin. 2019. Alberto Varvaro e la storia «interna» delle lingue romanze. In Laura Minervini (ed.), *Filologia letteratura e linguistica di Alberto Varvaro: atti delle giornate di studio*, 101–116. Padua: Antenore.

Maiden, Martin. 2020. Un problema trascurato di morfologia storica: la terza persona plurale del presente indicativo nell'italoromanzo. In Patrizia del Puente et al. (eds.), *Tra etimologia romanza e dialettologia. Studi in onore di Franco Fanciullo*, 245–259. Alessandria: Edizioni dell'Orso.

Malkiel, Yakov. 1991. 'Western Romance' versus 'Eastern Romance': The terms, the images, the underlying concepts. *Romanische Forschungen* 103: 141–156.

Merlo, Clemente. 1909. Gli italiani *amano*, *dicono* e gli odierni dialetti umbro-romaneschi. *Studii romanzi* 6: 69–83.

Merlo, Roberto. 2014. Un problema trascurato di fonetica storica romena: la 'palatalizzazione delle velari' tra eredità latina e interferenze slave. *Dacoromania* 19: 165–197.

Meyer-Lübke, Wilhelm. 1895. *Grammaire des langues romanes* II. *Morphologie.* Paris: Welter.

Mihăescu, Haralambie. 1960. *Limba latină în provinciile dunărene ale imperiului roman*. Bucharest: Editura Academiei Republicii Populare Romîne.

Nyrop, Kristoffer. 1960. *Grammaire historique de la langue française* II. *Morphologie*. Copenhagen: Gyldendal.

OVI = Opera del vocabolario italiano. Florence: CNR. http://www.ovi.cnr.it/.

Petrucci, Peter. 1999. *Slavic Features in the History of Romanian*. Munich: Lincom.

Rohlfs, Gerhard. 1968. *Grammatica storica della lingua italiana e dei suoi dialetti: morfologia.* Turin: Einaudi.

Rosetti, Alexandru. 1986. *Istoria limbii române*. Bucharest: Editura științifică și enciclopedică.

Sala, Marius. 1976. *Contributions à la phonétique historique du roumain*. Paris: Klincksieck.

Sala, Marius. 2013. Contact and borrowing. In Martin Maiden, John Charles Smith, and Adam Ledgeway (eds.), *The Cambridge History of the Romance Languages* I. *Structures*, 187–236. Cambridge: Cambridge University Press.

Skok, Petar. 1926. Zur Chronologie der Palatalisierung von *c g qu* vor *e i y i* im Balkanlatein. *Zeitschrift für romanische Philologie* 46: 385–410.

Stump, Gregory. 2001. *Inflectional Morphology. A Theory of Paradigm Structure*. Cambridge: Cambridge University Press.

Thomason, Sarah. 2010. Contact explanations in linguistics. In Raymond Hickey (ed.), *The Handbook of Language Contact*, 31–47. Oxford: Blackwell.

Väänänen, Veikko. 1963. *Introduction au latin vulgaire*. Paris: Klincksieck.

Vignoli, Carlo. 1911. Il vernacolo di Castro dei Volsci. *Studj romanzi* 7: 116–296.

Vignoli, Carlo. 1920. *Vernacolo e canti di Amaseno*. Rome: Società filologica romana.

Vignoli, Carlo. 1925. *Il vernacolo di Veroli in provincia di Roma*. Rome: Società filologica romana.

Weinreich, Uriel. 1968. *Languages in Contact.* The Hague: Mouton.

BRILL

Convergence by Shared Ancestry in Romance

Paul Widmer
Professor of Comparative Indo-European Linguistics, University of Zurich,
Zurich, Switzerland
paul.widmer@uzh.ch

Stefan Dedio
Research Assistant, (Indo-European Studies Group), Department of
Comparative Language Science, University of Zurich, Zurich, Switzerland
stefan.dedio@uzh.ch

Barbara Sonnenhauser
Professor of Slavonic linguistics, University of Zurich, Zurich, Switzerland
barbara.sonnenhauser@uzh.ch

Abstract

In many cases of apparent contact-induced change the contribution of genealogical correlation in the language sample and its interaction with processes such as matter and pattern replication are difficult to specify. In order to get a better sense of the relevance of shared ancestry, we quantify the change in similarity since the late Middle Ages in a sample of Romance and Germanic languages with data from a selected grammatical domain (expression of reflexivity). We compare their dynamics to patterns of change of similarity in two contact zones in Europe, namely the British Isles (Dedio et al., 2019) and the Balkans. Concerning the genealogical signal, the results indicate a maintenance and gain of similarity in Romance as opposed to a loss of similarity in Germanic. This hints at the importance of the inherited states, the time since the split from the common ancestor, and subsequent developments. We presume that these factors are likely to be at the origin of the maintenance and increase in similarity observed for the sampled Romance varieties. While this result cannot be generalized beyond the specific case study presented here, the basic approach will contribute to a better understanding of how contact, genealogy and culture interact in shaping the dynamics of linguistic similarity.

Keywords

convergence – divergence – Romance – Germanic – Balkan area – reflexive constructions

1   Introduction

Languages may be similar for various reasons, most prominently because similarity between languages tends to increase through language contact, because shared ancestry constrains evolutionary pathways such that potential divergences over time are slowed down, or because shared cultural practices of speaker communities foster gain or maintenance of similarity of specific linguistic properties (cf. the role of Latin in medieval Europe or the role of Pāli in the Buddhist sphere). Since these factors never occur in isolation it is obvious that change and maintenance of similarity are conditioned by an interplay of such drivers with a multitude of interacting socio-economic, spatio-temporal, and linguistic factors. Assessing and comparing evolutionary dynamics within and across groups of genealogically, spatially, and culturally related languages is therefore not a trivial undertaking, and telling apart the contribution of genealogical, spatial, and cultural drivers even less so.

The Balkans provide a major case in point for these challenges. At least since Trubetzkoy's introduction of the notion of 'sprachbund' (Trubetzkoy, 1928), the languages situated in the geographical area of the Balkan peninsula have been regarded as constituting a prototypical linguistic convergence area. This assumption has remained more or less unquestioned and in turn has led to the teleological interpretation of changes observed for exemplary features to result from language contact (see Introduction to this volume). Upon closer inspection, however, the Balkans constitute a prime example for the methodological challenges related to the identification of linguistic areas, in particular identifying the contributions of universal tendencies vs. inheritance vs. contact, assessing the dynamics of historical development, and avoiding the danger of circularity resulting from focusing on highly salient, emblematic examples. Since linguistic areas are assumed to be characterized by a trend of convergence that is significantly stronger than what could be expected from family-internal developments, the problem of how to estimate the diachronic dynamics (i.e., degree and amount) of convergence within and across (sub-) families necessarily gains center stage. In assessing the dynamics of convergence, it does not suffice to focus on features that have been changing towards increasingly similar patterns (see Introduction to this volume). Instead, further

information needs to be considered, in particular 1) information on features that resist change, 2) information on the directions of change, i.e., gain or loss of a feature, whereby both may have led to a state of indeterminacy, e.g., an attestation of various options. Obviously, this task is hard to accomplish by qualitatively oriented research.

The present paper aims to approach the above-mentioned challenges by applying an empirical method for assessing and comparing the dynamics of change of similarity across languages with variable genealogical relatedness, thus complementing qualitative research. In doing so, it takes a slight detour by focusing not on the Balkans in the first place, but on some Romance varieties of Europe and their development within their family and in the context of the Balkan languages. The analysis proposed in this paper is carried out on the basis of a set of morphosyntactic features involved in constructions expressing reflexivity. This construction is not commonly discussed in the context of the Balkan area, such as to avoid the above-mentioned danger of teleological interpretations of the data.

In addressing the interaction between genealogical, spatial, and cultural factors in language change, Romance varieties provide a good test case (see Joseph, 1999 for a similar point). They form a well-attested genealogical grouping with a thoroughly investigated internal history, stretching across various linguistic and geographical spaces from West to East in southern Europe, interacting with each other and quite a few other linguistic varieties across very diverse cultural settings (e.g., Catholic, Protestant, Orthodox, Jewish contexts, and/or under Ottoman and Arabic rule, etc.) since the early Middle Ages.

Because of the broad spatial distribution that offers opportunities for manifold contact events we expect the influence from various contact scenarios to manifest itself as a lineage-internal divergence irrespective of cultural and genealogical bonds, whereas a strong genealogical signal is expected to keep the divergence to a minimum or bring about a gain of similarity within the Romance varieties irrespective of space.

To get a better understanding of the properties of the genealogical signal in Romance, we propose to compare the change in similarity observed in a sample of Romance varieties to the changes in another lineage, namely Germanic, and to the changes in various configurations (see Dedio et al., 2019), i.e., geographical spaces with their own particular history and socio-cultural and socio-linguistic properties. To evaluate the dynamics of change, we apply the methods introduced by Dedio et al. (2019), which are based on the changes in similarity between linguistic varieties over time in predefined groups of languages, and make use of their set of morphosyntactic and morphological variables used to encode reflexivity, adding Balkan and Romance varieties to

their sample. In order to assess the contribution of Latin, the common cultural point of reference, to the change in similarity in Romance we add Classical and Medieval Latin, which remain, and are kept, unchanged during the period of interest. For practical reasons, the timespans between ancestor and successor varieties show more variation in our sample than in Dedio et al.'s; this may have an effect on the results we are not able to control for at the moment.

The article is structured as follows: We will introduce the data and methods in Section 2. The results concerning the dynamics of change observed for the different samples will be presented in Section 3 and discussed in Section 4, focusing in particular on the Romance and Balkan samples. Section 5 provides a conclusion and a more general embedding of the insights gained for assessing the role of contact for linguistic diversity and change.

2   Data and Methods

Teasing apart areal and genealogical signals in quantitative studies of linguistic diversity and change is a rather new enterprise that only started gathering pace in the last decade. The reliability of the methods proposed so far (e.g., Freckleton and Jetz, 2009; Nelson-Sathi et al., 2010; Willems et al., 2016; Kelly and Nicholls, 2017; Murawaki and Yamauchi, 2018) has not been thoroughly established yet, and although some of them seem promising, they are not suitable for this study, as they operate on a macro-level. We thus apply the method developed in Dedio et al. (2019) for identifying convergence within a group of languages and expand their data collection with data relevant for our purposes. Supplementing their sample of Gallo-Romance (French, Normand, Jèrriais) by Italian and Spanish and Balkan-Romance varieties (Romanian, Aromanian, Judezmo, cf. Table 1), we explore the dynamics of change in the Romance data against the backdrop of their sample from the British Isles, a sample of languages from the Balkans,[1] and another lineage, viz. Germanic. The method applied in Dedio et al. (2019) uses predefined areal groups, i.e., configurations, that are based on geographical, historical and sociological information to infer signals of convergence within these groups contrasted with the developments outside these predefined areas. As the languages of the Romance subphylum of Indo-European cannot in any way be interpreted as belonging to a single coherent area, we depart from their notion of configuration and apply the

---

[1] Note that the documents serving as data basis for the ancestor varieties of Macedonian, Bulgarian, Serbian and Torlak all constitute mixtures of Church Slavonic and vernacular features of the respective modern varieties each. This is the closest one can get to the older stages of these languages.

TABLE 1　Language pairings added to the collection in Dedio et al. (2019)

ancestor variety	approximate date CE	modern variety
Old Castilian	1250	Judezmo
Old Castilian	1250	Spanish
Old Albanian	1500	Modern Albanian
Old Italian	1250	Modern Italian
Middle Greek	1250	Modern Greek
Old Romanian	1500	Aromanian
Old Romanian	1500	Modern Romanian
Early modern Macedonian	1550	Macedonian
Early modern Bulgarian	1600	Bulgarian
Serbian Church Slavonic	1250	Torlak
Serbian Church Slavonic	1250	Serbo-Croatian

method also to groups of languages defined by shared ancestry (i.e., Romance and Germanic).

As Dedio et al. (2019), we focus on the expression of reflexivity, i.e., constructions in which the A and P arguments share the same referent (excluding constructions with exclusively reciprocal or passive semantics, and constructions that lack overt marking). The expression of these constructions encompasses a set of morphological, morphophonological, and morphosyntactic devices that can be captured in terms of binary variables as defined in the following list (for a full description, see Dedio et al., 2019; abbreviations in parentheses refer to the variables in Table 2). Note that by taking the construction as a starting point, our perspective differs from approaches that identify reflexivity as one possible function of particular devices, such as pronouns (see Cennamo, 2014 for Italian, Fehrmann et al., 2010 for Slavic).

Positional dependency (DEP.): Is the position of the reflexive marker directly dependent on the position of the verb? I.e., is there a rule that the marker must be placed relative to the verb? This includes basic rules like 'the reflexive marker is the innermost marker left of the verbal root' or more complex ones like 'with regular inflected verbs, the marker is in slot 3 of the verbal template, but in slot 1 with infinitives.'

Stress: Can the reflexive marker establish its own stress domain? We have opted for splitting stress and phonological interaction into two values as stress domains tend to be larger than other domains of phonological and prosodic interaction (Bickel et al., 2009: 72) and this distinction helps capturing variation in our data.

TABLE 2 States of the variables in each language pairing (represented by the modern language variety in the first column). The value on the left of each column represents the state in the older stage, the value on the right state at the end of the investigated period; for Latin, both values of each variable are identical.

	dep.	stress	interac.	allom.	num.	pers.	case	pre	post	rec.	pass.	exp.	syncr.
French	F T	T F	T T	F F	T T	T T	T F	T T	T T	T T	T T	T T	T
Jerriais	F T	T F	T T	F F	T T	T T	T F	T T	T T	T T	T T	T T	B
Normand	F T	T F	T T	F F	T T	T T	NA	T T	T T	T T	T T	T T	B
Picard	T T	F F	T T	F F	T T	T T	F F	T T	T T	T T	T T	T T	B
Italian	B T	B B	B T	F F	T T	T T	F F	T T	T T	T F	T T	B T	T
Spanish	T T	B F	T F	F F	T T	T T	F T	T T	T T	T F	T T	T T	T
Romanian	F T	B T	B B	F B	T T	T T	T B	T B	T B	T T	T B	B T	T
Aromanian	F T	B F	B T	F B	T T	T T	T B	T B	F B	T T	T B	B F	T
Judezmo	T T	B F	T T	F T	T T	T T	T F	T T	T T	T F	T T	T F	F
Classical Latin	B B	T T	B B	T T	T T	T T	B B	B B	T T	F B	B B	T T	B
Medieval Latin	B B	B B	B B	T T	T T	T T	B B	B B	T T	B B	T T	T T	T
Greek	B B	B B	B B	B B	T T	T T	B B	B B	T T	B B	B B	T B	F
Albanian	B T	B B	B T	B T	B B	B B	B B	T B	T T	B B	B B	B B	B
Macedonian	B T	T F	F F	F F	F F	F F	T B	T T	T T	T B	T B	T T	T
Bulgarian	T T	F F	F F	F F	F F	F F	T T	T T	T T	T T	T T	T T	T
Torlak	B T	B F	F F	F F	F F	F B	T B	T T	T T	T B	T B	T F	T
Serbo-Croatian	B F	B B	B F	F F	F F	F F	B T	T T	T T	B B	B B	F F	T

Interaction (INTERAC.): Does the reflexive marker interact phonologically with surrounding linguistic items (i.e., vowel harmony, liaison, mutations, etc.)? This includes the whole range of phonological integration like syllabification patterns, vowel harmony phenomena, liaison, etc., but excludes stress (see above).

Allomorphy (ALLOM.): Does the reflexive marker display phonologically, morphologically, syntactically or lexically conditioned allomorphy if all relevant features (i.e., number, person, case, TAM of the verb) remain the same?

Inflection (NUM., PERS., CASE): Is the reflexive marker specified for a) person, b) number, and/or c) case?

Positioning (PRE, POST): Where is the marker positioned with respect to the verbal root (PRE or POST)? The marker may be positioned to the left of the verbal root or to its right or, as in some cases, both.

Equivalence set (REC., PASS.): Does the marker form an equivalence set with passive or reciprocal? I.e., is the reflexive marker also used to express passives or reciprocals? Both functional overlaps are widely attested in the languages of the world and are present in our sample (e.g., the Old Norse "medio-passive" is used to form reflexives, reciprocals, and anti-passives). As with reflexive marking in general, we do not distinguish between "normal" or "unmarked" ways to express these two functions, but also include marginal strategies.

Expandable (EXP.): Can the reflexive construction be expanded with an intensifier or a similar formant for stress, clarification or similar ends? For most languages in our sample this is identical to the reflexive – intensifier distinction variable of SAE, but we wanted this variable to have a broader scope in the event one of the languages without this distinction developed the ability to use an additional reflexive/intensifier (e.g., ****I myself hurt myself*).

Third person number syncretism (SYNCR.): Does the reflexive marker distinguish number values in the third person? With this variable, we try to capture a common variation in our data, e.g., zero differentiation with Modern High German *sich*, full differentiation like in Modern Standard English, or is there one marker that is used for singular and plural, while there is another one that encodes plural only like in Fering, where the singular P-pronoun can also be used in plural constructions.

Any of the above variables can take the states TRUE, FALSE, and, as languages tend to make use of more than one construction to express reflexivity, BOTH. By comparing the values of ancestor and successor varieties such as Old Norman French and Normand, we get an impression of the direction in which a linguistic lineage develops (Section 3.1). The similarities between the feature values of contemporaneous varieties serve as the basis for assessing the trend and magnitude of convergences between individual languages or larger areal or phylogenetic groups (see Section 3.2).

For each variable we first explore the direction of change between ancestor and successor language and subsequently, following the method introduced in Dedio et al. (2019), we analyze the data as follows (refer to the supplementary information for technical details):

1. We compute the pairwise similarity of the languages in our sample (cf. Table 1) for two different points in time, namely t_1 = 1400 CE (± c. 200 years) and t_2 =c. 1950 CE, using an adapted version of the simple matching coefficient (SMC; Cheetham and Hazel, 1969).

2. Change between ancestor and successor languages is implemented by connecting the sampled languages at t_1 and t_2 to their phylogenetic ancestors. Lineages may fork into new sub-lineages, and therefore an ancestor language can have multiple successor languages at a given point in time. For example, Old Romanian is the ancestor of both Aromanian and Modern Romanian, cf. Table 1.

3. We then first compute the pairwise similarity between all languages that belong to the same t (e.g., Old Castilian and Middle Greek), and subsequently the change over time between pairings of identical ancestry, e.g., between the pairing Old Castilian/Old Romanian and their successor pairing Modern Spanish/Modern Romanian, and so on.

4. To assess how the developments relate to possible area formation processes and lineage specific trends, we split the measurements of changes in similarity into groups according to whether both languages of a pair are part of the Balkan area or both languages of a pair belong to the same lineage. The development of similarity inside of each of the resulting groups are visualized as rainbow plots.

5. In order to estimate the trend of the development in the Romance and Germanic samples, we model the change in similarity as a Bernoulli process $Y \sim$ Bernoulli (p), where p is the probability of success, i.e., in our case gain in similarity. To explore the magnitude of the process in Romance and Germanic, we model the change in similarity as transformed beta distribution $Y \sim$ Beta(α, β) in the interval $[-1,1]$ and estimate the posterior predictive distribution given the observed pairwise change within Romance. For a full description, refer to Dedio et al. (2019) and Ranacher et al. (2019).

3   Results

We first qualitatively investigate individual developments within Romance and the Balkan area before turning to a quantitative evaluation of the areal and phylogenetic developments.

3.1   *Direction of Changes*

In this section we investigate the direction of the changes per variable within each pairing. The feature specifications for each variable are given in Table 2, with the left value showing the specification for the older, the right value that for the younger stage (with the exception of Classical and Medieval Latin that serve as control factors for the influence of cultural practices and are kept constant).

Based on these specifications, the following observations concerning the change patterns can be made:

Overall sample: Two variables, namely NUM. and PERS., have not changed at all across the whole language sample, hinting at more general extra- and/or intra-linguistic trends.

Romance: All Romance varieties fully converge on, or preserve, a specification with DEP., NUM., ALLOM., PERS., PRE., REC., and PASS. With POST., all converge except for Aromanian; with EXP., the two Balkan-Romance varieties Judezmo and Aromanian display a behavior different from the other Romance varieties. Concerning STRESS, Italian and Romanian are the only varieties within Romance that fail to converge.

Balkan sample: The development in the Balkan sample is much less uniform than within Romance, there is no obvious discernible pattern.

Note that the development of these features is meaningful only in the context of the overall construction and its overall 13 features, such that NUM. and PERS. need to be included in all analyses even though they remain constant for all varieties.

Since it is difficult to obtain an overview of the more general trends displayed by the specification given in Table 2, we apply a quantitative approach that helps to discern the change in similarity for all of these features and to compare different groups of languages according to these changes. We zoom in and compute the pairwise changes in similarity for the Balkan and the Romance sample varieties, the latter including Classical and Medieval Latin. Romanian, Aromanian and Judezmo are included in both samples in order to be able to assess their (non)compliance with the overall trends in both samples. The results are shown in Fig. 1.

Upon visual inspection, Romance displays more maintenance and convergence of similarity than divergence with the notable exceptions of the Judezmo-Aromanian pairing and Classical Latin, from which most Romance varieties slightly diverge. Romanian not only diverges from both Latin samples, but also from Spanish. The former might relate to its cultural embedding in the orthodox sphere, the latter to geographical distance. The Gallo-Romance varieties all slightly converge with Medieval Latin, while all other varieties keep the same distance from Medieval Latin or diverge to a minor degree.

FIGURE 1   Pairwise change in similarity in the sample of Romance varieties including Classical
            and Medieval Latin (left panel) and the sample of varieties from the Balkans (right
            panel). Blue indicates divergence, orange convergence. The darker the color, the
            stronger the trend. White indicates no change.

Beyond their mutual divergence, Judezmo and Aromanian do not run counter to the overall dynamics of change in similarity within Romance. Within the Balkan sample they do not display an obvious increase in similarity. Rather, both converge with some Balkan varieties but diverge from others. Doing so, they fit in the overall picture in the Balkan sample, which lacks a uniform tendency in the direction of change.

In order to assess the significance of these observations, we provide kernel density estimates of the observed pairwise changes in five subsamples in Fig. 2, namely the Romance varieties (without Latin), all non-Romance varieties, all Germanic varieties, the sample of varieties from the British Isles (Dedio et al., 2019), and the sampled varieties from the Balkans. The median, maximal and minimal values, the median absolute deviation (MAD, a measure of variability) and the percentage of data points that are greater than 0 (i.e., show gain in similarity) are reported in Table 2.

The median trend towards gain of similarity is most pronounced in the sample from the British Isles (median .35, 80% of the data points > 0), followed by the Romance sample (median .18, 86% of the data points > 0), notably with the highest amount of data points > 0 and the smallest median absolute deviation (.08) across samples, hinting at a small, but robust trend towards gain in similarity. Unlike Romance, Germanic is dominated by divergence (median -.14, only 12% of the data points > 0). The Balkan sample has a quite flat distribution clustering around 0, indicating that there is ongoing, but unbiased change.

3.2   *Trend and Magnitude of the Change of Similarity in Romance*
The result of the Bernoulli model in Fig. 3 shows that the probability of gain in similarity (i.e., success) in 1000 random samples of Romance-Romance pairs
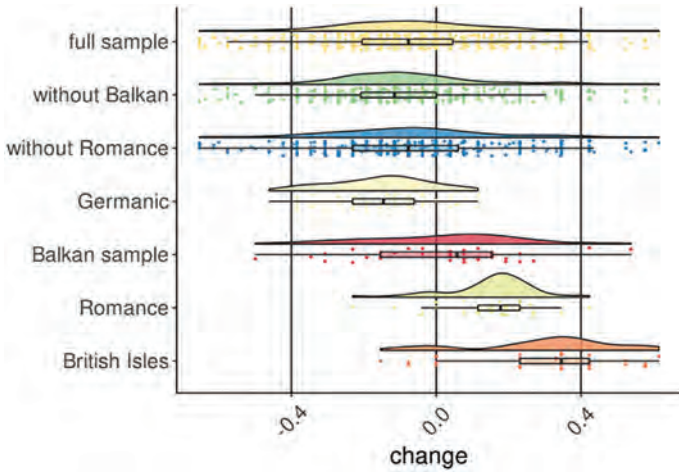
FIGURE 2   Raincloud plots and boxplots of the distribution of pairwise changes in the full sample and subsamples thereof; full sample = all languages in the sample; without Balkan = all languages except Balkan languages; without Romance = all languages except Romance languages; Germanic = only Germanic languages; Balkan sample = only Balkan languages; Romance = only Romance languages; British Isles = only languages from the British Isles.
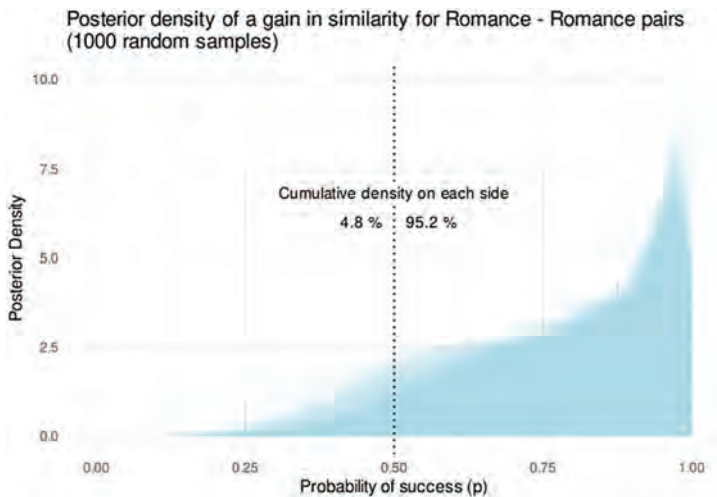

FIGURE 3   Posterior density of gain in 1000 random samples of Romance-Romance pairings. 95.2% of the samples display a gain of similarity.
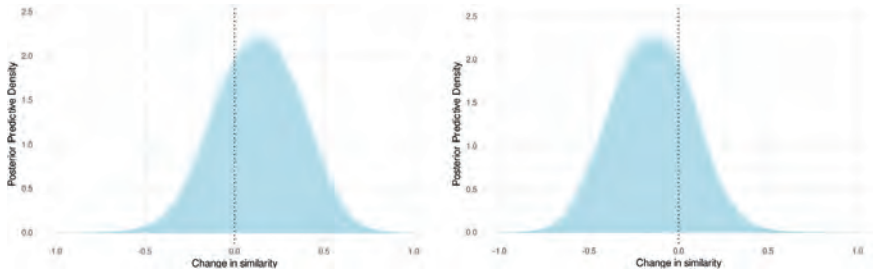
FIGURE 4   Posterior predictive distribution of changes in 1000 random samples. **Left panel:** Romance-Romance pairs (median = .11, 71% of the data > 0). **Right panel:** Germanic-Germanic pairs (median = -.15, 22% of the data > 0.)

is 95.2%. This means that the Romance languages in our sample show a trend towards convergence in their expressions of reflexivity.

The magnitude of the process is estimated with a beta model. An overlay of all posterior predictive distributions of the beta model for Romance is shown in the left panel of Fig. 4 (median = 0.11; 71% of the data > 0) and for Germanic in the right panel (median = - 0.15; 22% of the data > 0). The model predicts that in formerly unexplored Romance-Romance language pairs, there is an average gain of similarity of ca .1; for Germanic-Germanic pairs the average loss of similarity is .15. In practical terms, the posterior predictive distribution predicts that while there is a trend towards convergence, the actual average convergence in Romance amounts to a change in 1.5 variables out of 13.

4   Discussion

Towards identifying the input of the different factors contributing to changes in linguistic similarity, we examined samples of spatially and/or genealogically related language varieties between c. 800-500 ybp (ancestors) and present time (successors). We focused on Romance, a sample of languages which has a shared linguistic ancestry, stretches across contingent and non-contingent spaces, and partakes in a common cultural practice. To estimate the contributions of genealogy and contact in the Romance sample, we also contrasted a spatially defined sample of Balkan varieties with a sample from the British Isles.

4.1   *Romance Sample*
We found a small overall trend towards convergence across the varieties from the Romance lineage (see Section 3.1). The gain in similarity is smaller in

TABLE 3 Median, maximum, and minimum values of changes in all groupings, the MAD, and the percentage of data points with a positive value (i.e., gain); full sample = all languages in the sample; without Balkan = all languages except Balkan languages; without Romance = all languages except Romance languages; Germanic = only Germanic languages; Balkan sample = only Balkan languages; Romance = only Romance languages; British Isles = only languages from the British Isles.

grouping	median	maximum	minimum	MAD	data points > 0
full sample	-0.08	0.61	-0.65	0.19	30%
without Balkan	-0.11	0.61	-0.65	0.17	25%
without Romance	-0.08	0.61	-0.65	0.23	31%
Germanic	-0.14	0.11	-0.46	0.13	12%
Balkan sample	0.06	0.54	-0.5	0.20	61%
Romance	0.18	0.42	-0.23	0.08	86%
British Isles	0.35	0.61	-0.15	0.17	80%

our Romance sample than in the sample of varieties from the British Isles as reported in Dedio et al. (2019), and distinctly different from what we observe in Germanic as the other lineage-based sample. Also, the dynamics of change in Romance are quite different from the ones in our sample of non-Romance varieties, which has a rather flat distribution with more divergence than convergence.

Concerning the potential loss of similarity between non-Balkan and Balkan Romance, first sight evidence from Table 2 and Fig. 1 suggests that even though the Balkan Romance languages are not uniform in themselves, it is exactly these languages – in particular Aromanian and Judezmo – that spoil the otherwise quite uniform picture for Romance. It might thus be tempting to take this diverging behavior as attesting to the areal influence both have been exposed to. This is precisely what Friedman and Joseph (2014) suggest for Judezmo based on different data than ours, interpreting the divergences from Gallo- and Ibero-Romance as attesting to contact within the Balkan area. However, at least as concerns the part of grammar under consideration in this paper, the developments of features identified for Judezmo as compared to those identified for the Balkan sample do not reveal any similarities except for, maybe, the feature POST. This does not come as a surprise, given that Judezmo entered the Balkan area only in the 15th century, which also accounts for the considerable degree of divergence from Aromanian as the Romance variety that is most entrenched in the central Balkan area in our sample.

In a more general perspective, Romanian, Judezmo, and Aromanian, i.e., the Romance varieties that are not in direct contact with Gallo-, Italo-, and

Ibero-Romance, neither show significant convergences with the Balkan languages nor display a fundamentally aberrant behavior from the other Romance varieties (see also Joseph, 1999). This hints at lineage-internal contact not being the only relevant factor for the convergence in the Romance sample.

Moreover, the trend within Romance is stronger than in the non-Romance sample and in particular than in the subsample of varieties from the other lineage under consideration here, namely Germanic. It seems thus unlikely that these Romance developments attest to more general areal trends (e.g., those underlying larger presumed areas such as SAE). Unless the convergence in the sampled Romance varieties is mere coincidence or an artefact of our sampling choices, the findings suggests that either cultural and/or genealogical factors are at work and call for an explanation.

As for culturally induced convergence, it has been claimed that the literary national languages, which are well represented in our sample, were deliberately oriented toward Latin as prestigious role model (see Blatt, 1957; Pountain, 2010). This trend is assumed to have pertained to all evolving Romance literary languages in Europe. In our data and time span, this influence is not palpable in the case of Classical Latin, whose similarity to the Romance varieties decreases slightly or remains unchanged. Medieval Latin shows a minor gain in similarity with the Gallo-Romance varieties only, but even for Gallo-Romance this convergence is too small for it to provide conclusive evidence for a Latin influence.

As for the contribution of shared ancestry, it is notable that the Latin medio-passive conjugation was lost in the more or less simultaneous emergence of the Romance varieties in Late Antiquity and the early Middle Ages (Cennamo, 2016; Miller, 2010). The Romance varieties fell back on constructions that use pronominal elements of demonstrative or reflexive origin. During the span of time under investigation, these pronominals became ever more positionally dependent on the verb, lost the ability to establish their own stress domain, and started to increasingly interact with adjacent morphological material and assume passive functions, cf. Table 2. This shared evolutionary pathway corresponds to what is well known from grammaticalization of pronominals into markers of reflexivity (cf. Albanian *-u-* < **swe*, North Germanic *-s* < **sik*, Russian *-sja/-s'* < **sę* etc., cf. Matzinger, 2006: 120; Heine and Kuteva, 2002: 253; Vasmer, 1964: vol. 3: 823). On this cline of grammaticalization, the inherited Romance pronominals in reflexive functions lost syntactic and prosodic independence, cliticized to an already available host and became affixes eventually as in the fully grammaticalized and invariable reflexive prefix *se-* in Sursilvan Rumantsch (Spescha, 1989: 384); all of these developments are typical for grammaticalization processes (Lehmann, 2015). Concerning the hosting

site, Romance started out with a pre-/postverbal host position only (Wanner, 1987: 155) and it is in this inherited pre-/postverbal site that the Romance varieties subsequently hosted the cliticizing reflexive pronominals/markers. This inherited state also accounts for the differences in reflexive clitic placement between Romance and Slavic: unlike Romance, the Slavic varieties inherited both a Wackernagel and a pre-/postverbal host position for pronominal clitics and exploited them in different ways (Wackernagel position in BCMS, pre-/postverbal in Bulgarian, postverbal in Russian *-sja/-s'*; see Pancheva, 2005; Kuehnast, 2009; Franks, 2010).

Given the inherited state in Romance, it is not difficult to account for the emergence of passive functions in Gallo- and Italo-Romance either, since it is a well-known process of grammaticalization for reflexives to acquire passive functions as well (cf. Danish *-s*, Russian *-sja/-s'*, Heine and Kuteva, 2002: 253).

The shared ancestry of Romance with initially independent reflexive pronouns may thus represent an early stage of related (European and/or more general) grammaticalization clines (reflexive pronouns > reflexive marker; reflexive function > reflexive + passive function) that afforded systemic opportunities for converging developments. It seems that the time period covered by our investigation embraces exactly this cline of grammaticalization – which once again illustrates very clearly the relevance of the decisions made in the sampling choices and of the availability of data. Things are fundamentally different in the Balkan sample: the various lineages split long ago from their last common ancestor and there is no cross-lineage shared state at the beginning of the period under investigation.

To sum up, the observed lineage-internal convergence and maintenance of similarity in our sample of Romance varieties is most likely related to the rake-like speciation of the Romance languages. We assume that their shared inheritance from the common ancestor constrained the subsequent evolution so as to follow common paths of grammaticalization. To a yet unknown extent, this process was partly enhanced by contact induced innovation and spread (Gast and Auwera, 2012; Heine and Kuteva, 2003) mediated by mutual cultural influence, historical contingencies, and in part spatial proximity.

4.2   *Balkan Sample*
With regard to the sample of varieties from the Balkans, it is worth noting that the distribution of pairwise changes doesn't reveal a clear trend toward gain or loss of similarity, but there might be some temporal stratification. For example, no change into neither direction is found for Albanian and both Aromanian and Romanian, two Romance varieties attested long before the Ottoman rule, while with Judezmo, which came to the Balkans only during the Ottoman

times, a slight increase in similarity is observed. Overall, half of the data points center around 0, which hints at some degree of maintenance of similarity, and the dynamics add up to a rather low degree of overall change. Thereby, the Balkan sample clearly differs from the pattern observed in the sample from the British Isles (cf. Fig. 2). Note, however, that because of the contingencies of transmission, the time span covered in the Balkan sample is less uniform than in the other subsamples.

The kind of distribution detected here has been observed before and has variously led to conclude that the Balkan configuration embraces a zone of small-scale subareas of convergence with partially opposite trends across subareas (Joseph, 2010), or, alternatively, that the process of area formation on the Balkans is a phenomenon that belongs to the more remote past (Topolińska, 1995: 240). However, local convergences and Balkan-wide convergences are difficult to tell apart, and it has been argued that the invasion of the Avars and Slavs on the Balkans from the 6th c. onwards and the establishing of Ottoman rule from the mid-15th century onwards, both resulted in the dissolution of central administration and of administrative boundaries and provided a sociolinguistic situation favoring convergences and maintenance in the entire area (e.g., Lindstedt, 2000: 240; Friedman, 2011: 284). It would be premature to link the changes in our sample to any of the available theories – in fact, to some extent, our results are compatible with all of them – and after all, our sample and time span may simply be too small, or the features we looked at are not relevant for area formation in this case.

In more general terms, towards assessing the contributions and interactions of the various factors behind changes in similarity within a specific time span, for a specific language sample and a specific set of features, the state at the beginning of the period of time matters and needs to be identified for each language variety in the sample: it is an important confound and needs to be taken into account for estimating the role and degree of spatio-temporal and cultural factors and eventually getting a better sense of area formation processes.

5   Conclusion

Regarding the thirteen linguistic variables contributing to the expression of reflexivity considered in this paper, we discovered a maintenance and gain in similarity in the Romance sample in a period of time between the late Middle Ages and the present across ancestor and successor pairs. Given the manifold (family-external) contact situations provided by the wide geographical expansion of Romance, this convergence is rather unexpected. It also stands

in contrast to the development in another lineage, viz. Germanic, where we observe a decrease in similarity. However, it is strikingly similar to the development observed for the British Isles, a sample of languages of mixed genealogy, but becoming more similar because of contact.

On the basis of our data, the shared ancestry of the Romance languages emerges as the most likely source of the gain in similarity. The Romance sub-lineages branched off within a rather short span of time and inherited the same system of reflexive marking involving unbound pronominals. This shared inheritance constrained the development according to well-established pathways of grammaticalization, maintaining and even increasing the similarity, a development that runs counter to the divergence in other domains of grammar.

This case study leads us to conclude that when evaluating the impact of contact on the evolution of linguistic similarity, it is important to control for genealogical correlation between the sampled varieties and take into account the degree of relatedness in terms of time since speciation. Also, the dynamics of change in a given sample of languages depends on the state of each linguistic feature in each language at the beginning of the period of interest: completely different systems may converge due to contact, but so may very similar systems because of common pathways of grammaticalization.

As the data used in this study is a rather small set of variables from a rather small section of grammar, it is important to keep in mind that the obtained results cannot be used for generalizations without further investigation. What we hope to have provided is a proof of concept that when applied on a large set of linguistic data and a dense sample of language varieties, the approach taken here will likely enhance our understanding of the manifold interactions between drivers of gain, loss, and maintenance of linguistic similarity.

References

Bickel, Balthasar, Kristine A. Hildebrandt, and René Schiering. 2009. The distribution of phonological word domains: A probabilistic typology. In Janet Grijzenhout and Barış Kabak (eds.), *Phonological Domains: Universals and Deviations*, 47–75. Berlin: De Gruyter.

Blatt, Franz. 1957. Latin influence on European syntax. *Classica et Mediaevalia* 18: 133–178.

Cennamo, Michela. 2014. Passive and impersonal reflexives in the Italian dialects. In Paola Benincà , Adam Ledgeway , and Nigel Vincent (eds.), *Diachrony and Dialects: Grammatical Change in the Dialects of Italy*, 71–95. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780198701781.003.0004.

Cennamo, Michela. 2016. Voice. In Adam Ledgeway and Martin Maiden (eds.), *The Oxford Guide to the Romance Languages*, 967–980. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199677108.003.0060.

Cheetham, Alan H. and Joseph E. Hazel. 1969. Binary (presence-absence) similarity coefficients. *Journal of Paleontology* 43(5): 1130–1136.

Dedio, Stefan, Peter Ranacher, and Paul Widmer. 2019. Evidence for the British Isles as a linguistic area. *Language* 95(3): 498–522.

Fehrmann, Dorothee, Uwe Junghanns, and Denisa Lenertová. 2010. Two reflexive markers in Slavic. *Russian Linguistics* 34(3): 203–238. DOI: 10.1007/s11185-010-9062-7.

Franks, Steven. 2010. Clitics in Slavic. *Glossos* 10: 1–157. https://slaviccenters.duke.edu/sites/slaviccenters.duke.edu/files/media_items_files/10franks.original.pdf.

Freckleton, Robert P. and Walter Jetz. 2009. Space versus phylogeny: Disentangling phylogenetic and spatial signals in comparative data. *Proceedings of the Royal Society of London B: Biological Sciences* 276(1654): 21–30. DOI: 10.1098/rspb.2008.0905.

Friedman, Victor A. 2011. The Balkan languages and Balkan linguistics. *Annual Review of Anthropology* 40(1): 275–291.

Friedman, Victor A. and Brian D. Joseph. 2014. Lessons from Judezmo about the Balkan Sprachbund and contact linguistics. *International Journal of the Sociology of Language* (226): 3–23. DOI: 10.1515/ijsl-2013-0073.

Gast, Volker and Johan van der Auwera. 2012. What is contact-induced grammaticalization?: Examples from Mayan and Mixe-Zoquean languages. In Björn Wiemer, Bernhard Wälchli, and Björn Hansen (eds.), *Grammatical Replication and Borrowability in Language Contact*, 381–426. Berlin: De Gruyter.

Heine, Bernd and Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.

Heine, Bernd, and Tania Kuteva. 2003. On contact-induced grammaticalization. *Studies in Language* 27(3): 529–572.

Joseph, Brian D. 1999. Romanian and the Balkans: Some comparative perspectives. In Sheila Embleton, John E. Joseph, and Hans-Josef Niederehe (eds.), *The Emergence of the Modern Language Sciences. Studies on the transition from historical-comparative to structural linguistics in honour of E.F.K. Koerner, Volume 2: Methodological perspectives and applications*, 217–235. Amsterdam: John Benjamins.

Joseph, Brian D. 2010. Language contact in the Balkans. In Raymond Hickey (ed.), *The Handbook of Language Contact*, 618–633. Oxford: Wiley-Blackwell.

Kelly, Luke J. and Geoff K. Nicholls. 2017. Lateral transfer in Stochastic Dollo models. https://arxiv.org/abs/1601.07931v3 (accessed March 16, 2017).

Kuehnast, Milena. 2009. Processing clitic pronouns in Bulgarian. Evidence from normal and agrammatic comprehension. *Poznan Studies in Contemporary Linguistics* 45(4): 473–493. DOI: 10. 2478/v10010-009-0029-z.

Lehmann, Christian. 2015. *Thoughts on Grammaticalization*. 3rd ed. Berlin: Language Science Press.

Lindstedt, Jouko. 2000. Linguistic balkanization: Contact-induced change by mutual reinforcement. *Studies in Slavic and General Linguistics* 28: 231–246.

Matzinger, Joachim. 2006. *Der altalbanische Text Mbsuame e krështerë* (*Dottrina cristiana*) *des Lekë Matrënga von 1592: Eine Einführung in die albanische Sprachwissenschaft*. Dettelbach: Röll.

Miller, D. Gary. 2010. *Language Change and Linguistic Theory, Vol. 2: Morphological, Syntactic and Typological Change*. Oxford: Oxford University Press.

Murawaki, Yugo and Kenji Yamauchi. 2018. A statistical model for the joint inference of vertical stability and horizontal diffusibility of typological features. *Journal of Language Evolution* 3(1): 13–25. DOI: 1 0.1093/jole/lzx022.

Nelson-Sathi, Shijulal, Johann-Mattis List, Hans Geisler, Heiner Fangerau, Russell D. Gray, William Martin, and Tal Dagan. 2010. Networks uncover hidden lexical borrowing in Indo-European language evolution. *Proceedings of the Royal Society of London B: Biological Sciences*. DOI: 10.1098/rspb.2010.1917.

Pancheva, Roumyana. 2005. The rise and fall of second-position clitics. *Natural Language and Linguistic Theory* 23: 103–167.

Pountain, Christopher. 2010. Latin and the structure of written Romance. In Martin Maiden , John Charles Smith , and Adam Ledgeway (eds.), *The Cambridge History of the Romance Languages*, vol. 1, 606–659. Cambridge: Cambridge University Press. DOI: 10.1017/CHOL9780521800723.015.

Ranacher, Peter, Stefan Dedio, and Paul Widmer. 2019. *Evidence for Britain and Ireland as linguistic Area: Supporting online material 2: R-script*. https://drive.switch.ch/index.php/s/GptpaIz78TGCd8f/ download.

Spescha, Arnold. 1989. *Grammatica sursilvana*. Cuera: Casa Editura per Mieds d'Instrucziun.

Topolińska, Zusanna. 1995. Convergent evolution, creolization, and referentiality. *Prague Linguistic Circle Papers* 1: 239–247.

Trubetzkoy, Nikolai S. 1928. Proposition 16. *Actes du premier Congrès International de Linguistes: à La Haye du 10–15 avril 1928*, 17–18. Leiden: Sijthoff.

Vasmer, Max. 1964. *Ėtimologičeskij slovar' russkogo jazyka*. Moscow: Progress.

Wanner, Dieter. 1987. *The Development of Romance Clitic Pronouns: From Latin to Old Romance*. Berlin: Mouton de Gruyter.

Willems, Matthieu, Etienne Lord, Louise Laforest, Gilbert Labelle, François-Joseph Lapointe, Anna Maria Di Sciullo, and Vladimir Makarenkov. 2016. Using hybridization networks to retrace the evolution of Indo-European languages. *BMC Evolutionary Biology* 16(1): 180. DOI: 10.1186/s12862-016-0745-6.

BRILL

Contact-Induced Complexification in the Gender System of Istro-Romanian

Michele Loporcaro
Full Professor of Romance Linguistics, Institute of Romance Studies, Faculty of Arts and Social Sciences, University of Zurich, Zurich, Switzerland
loporcar@rom.uzh.ch

Francesco Gardani
Professor of Romance Linguistics, Institute of Romance Studies, Faculty of Arts and Social Sciences, University of Zurich, Zurich, Switzerland
francesco.gardani@uzh.ch

Alberto Giudici
Assistant in Romance Linguistics, Institute of Romance Studies, Faculty of Arts and Social Sciences, University of Zurich, Zurich, Switzerland
alberto.giudici@uzh.ch

Abstract

The paper provides the first description of the borrowing of Croatian collective numerals into Northern Istro-Romanian and explores the consequences of this borrowing for the morphosyntax of the recipient language. It argues that the collective numerals under examination, which are specified as nominative plural feminine in the Slavic model, took on a different structural specification in the Romance replica, in a way that led to a restructuring of the morphosyntactic system, introducing (sub)gender overdifferentiation on just two agreement targets and, thereby, a complexification in this area of grammar. The illustration of this change is placed against the background of the other contact-induced changes that grammatical gender has undergone in Istro-Romanian during the 20th century, which have led to the borrowing of two dedicated forms in distinct inflectional cells and the rise of two separate defective gender values, each the replica of one number value of the Slavic neuter.

Keywords

1   Introduction*

One of the probably most widely accepted claims made in the language con-
tact literature is that contact leads to the simplification of grammar. The basic
assumption behind this claim is that the mixing of linguistic systems produces
less marked structures and levels out irregularities towards "a kind of common
core-grammar" (Mühlhäusler, 1980: 28; see also Givón, 1979; Bickerton, 1981).
This simplification hypothesis, despite a few counterexamples (Thomason and
Kaufman, 1988: 29; Vanhove, 2001; Aikhenvald, 2002; 2003; Adamou, 2013; de
Groot, 2008; Melissaropoulou, 2017; Loporcaro, 2018: 51, 291), still dominates
research not only in creole languages (McWhorter, 2001; 2007), but also in
so-called regular contact-induced change (Kusters, 2003; Trudgill, 2009; 2011).

In the present paper, we describe a case of contact-induced grammatical
complexification involving Istro-Romanian, a heavily endangered Romance lan-
guage spoken by a few dozen speakers, all bilinguals with Croatian, in two areas
of Istria as well as by a few hundred speakers around the world as a heritage
language. After providing background information on the language (Section 2)
and on numeral borrowing cross-linguistically and in Slavic-Romance contact
(Section 3), we will address Istro-Romanian numerals, showing that the bor-
rowing process has concerned not only ordinal and cardinal numbers (a fact
that has long been described, Section 4), but also adnominal numerical quan-
tifiers (Section 5). Here, borrowing has made possible numerical quantification
with *pluralia tantum* nouns – with which Romance languages often resort to
alternative strategies (the "classifier solution", Corbett, 2019: 93f.; see the exam-
ples in (6)–(7) below) – and at the same time the signalling of gender/number
agreement with such head nouns on some numerical quantifiers, in a way that

---

deviates from what is found elsewhere on agreement targets in Istro-Romanian. We will argue that this resulted in morphosyntactic complexification. First, the morphosyntactic system prior to the borrowing is described in Sections 6–7, showing that contact with Croatian had already made the inherited gender system more complex, triggering the borrowing of two additional gender values in two successive steps. Then, against this background, in Section 8 we argue that the borrowing of the numerical quantifiers at issue has led to the rise of gender overdifferentiation (Corbett, 1991: 168f.) on just a few agreement targets (lower numerals). While gender overdifferentiation on lower numerals has been described before for several languages, including some Romance varieties, its rise through borrowing never has. Since this borrowing process resulted in a net increase in complexity of the gender/number agreement rules, this case study adds to the series of contact-induced changes which bring about complexification, rather than simplification, in the grammatical subsystem involved (cf. e.g., Vanhove, 2001; Aikhenvald, 2003; Adamou, 2013; Melissaropoulou, 2017; Loporcaro, 2018: 51, 291; Meakins and Wilmoth, 2020).

2   Istro-Romanian

Istro-Romanian (henceforth IR) is one of the four branches of Daco-Romance.[1] It is spoken today by a tiny number of speakers (about 100, most of them over 50 years old) in north-eastern Istria (see Map 1). It divides in two mutually comprehensible, yet clearly distinct varieties, which have been isolated from each other for centuries since the late Middle Ages and developed divergent innovations in both lexicon and grammar:[2] Northern Istro-Romanian

---

1   There is an issue about terminology here. While most authors call Istro-Romanian one of the four dialects of Romanian (see e.g., Tagliavini, 1972: 356–364), linguists from the local community (e.g., Vrzić and Doričić, 2014: 105) prefer subsuming Žejanski directly under a superordinate classificatory unit 'Eastern Romance'.

2   The list of diverging grammatical properties includes various differences in verb inflection, e.g., SIR -*rno* 1pl restrictive future employed only in conditional clauses *flårno* (Puscariu, 1926: 185) vs. NIR *aflårem* ← *aflå* 'find' (Kovačec, 1971: 143; see also Hurren, 1999: 101); the loss of the imperfect tense in NIR versus its preservation in SIR, where, however, it is restricted to the aspectual function of continuous (Hurren, 1969: 89).

(i)   a.   *in špitól am muŋkót žir svka zi*   (NIR)
        b.   *in špitól muŋkójam žir svka zi*   (SIR)
               'while at the hospital, I used to eat fruit every day'

In the lexicon too, several differences exist, often due to the different intensity of contact with different languages in SIR vs. NIR: for instance, for 'newspaper' NIR uses the Croatian loan

(henceforth NIR), spoken only in the village of Žejane (IR *Jeiăn*, Italian *Seiano*, in the municipality of Matulji, Primorje-Gorski Kotar district), and Southern Istro-Romanian (henceforth SIR), spoken in a cluster of villages lying some 22 km to the SSW as the crow flies, but at least 40 km on foot, being separated from NIR by the Učka/Monte Maggiore massif.[3] For SIR, the data cited in this

Map 1.   Istro-Romanian in Istria (after Loporcaro, 2018: 293, with modifications)
■ = Istro-Romanian; ◆ = Croatian ; ● = Italo-Romance and Croatian

---

novine (*plurale tantum*), whereas SIR has borrowed *ǧornɒle* from Italian. A detailed account of these differences is provided by Kovačec's (1998) dictionary and Filipi's (2002) atlas.

3   The villages, all included in the municipality of Kršan (in the Istria district), are those of Brdo (IR *Bărda*, It. *Berdo*), Kostrčani (IR *Costărcian,* It. *Costerciani*), Letaj (IR/It. *Letai*), Miheli (= IR/

article come from Šušnjevica when not otherwise specified, as well as from nearby Jesenovik.

All IR speakers are bilingual with Croatian (in the standard variety and the Čakavian varieties). As a result of centuries-long total language contact, the structure of IR has been massively reshaped (see Kovačec, 1963; 1966; 1968; 1971; Filipi, 2002; Sala, 2013: 218–225; Vrzić and Doričić, 2014). In the phonology, consonants with a secondary palatal articulation lost it (the contrast does not occur in Croatian), which impacted on inflectional morphology, since in Romanian palatalization induced by final -/i/ marks the plural in many nominal paradigms, whose singular/plural forms became homophonous in IR: compare Ro. *lup* 'wolf', pl. *lup^j* with IR *lup* 'wolf=wolves' (Kovačec, 1998: 108).[4] The syntax of IR is basically that of Croatian, including its relatively free word order (contrary to Romanian) as well as specific rules such as those affecting the placement of clitic auxiliaries (e.g., *vlót=am* 'I have taken', vs. Ro. *am luat*; see the examples in (14a,c) below). In the lexicon, extensive borrowing resulted in replacement even in core lexical domains: Vrzić and Doričić (2014) describe its increase over time for body parts. As a consequence, often whole IR sentences consist solely of Croatian lexical material "sans en changer autre chose que les morphemes grammaticaux" [without changing anything else but grammatical morphemes] (Kovačec, 1968: 81). Even here, Croatian has impacted, as witnessed by IR being possibly the sole Romance language in which the inherited first conjugation (Lat. LIGARE > *leyÔ* 'to tie') has become unproductive, while new verb lexemes are formed with Slavic suffixes (Kovačec, 1971: 131f.): e.g., *čiravéi* 'to have dinner', derived with the suffix *-av-* plus a non-etymological inflectional ending from the Romance base (cp. *čira* 'dinner' < Lat. CĒNAM).

IR speakers are not singled out by a specific ethnic/linguistic identity and perceive themselves as homogeneous to the Croatian environment, a circumstance which favours assimilation. In this ecological setting, generalized bilingualism and the steep increase in mobility over the past few decades triggered a language shift which is now approaching its final stage: IR nowadays does not appear to have fluent native speakers below 40 years of age and is not being passed on to the next generations.[5] During fieldwork in Istria in 2017–2018,

---

It. *Micheli*), Nova Vas (IR *Noselo/Nosela*, It. *Villanova*), Šušnjevica (IR *Suseni*, It. *Susgnevizza*), Jesenovik (IR *Sukodru*, It. *Frassineto*), and Zankovci (IR *Zancovţi*, It. *Zancovzi*) (see the list in Filipi, 2002: 31).

4 For some lexemes, the distinction was restored applying the *-ure* suffix originally restricted to neuters, as in IR *lúpure* 'wolves' competing with the unmarked plural *lup* (Kovačec, 1966: 64; see examples in context in (36a-b)).

5 Language shift is rampant in the area, so that the *Ethnologue* classification as "shifting" (EGIDS level 7: cf. https://www.ethnologue.com/cloud/ruo) is more than justified. Given this, it is obvious that higher figures are indicated in the literature as one climbs back in time.

we had a chance to interview a dozen IR speakers: the results brought to light some interesting facts that had gone unnoticed in the previous literature.

3   Borrowed Numerals and Mixed Numeral Systems in Slavic-Romance Contact and Beyond

The borrowing of numerals is a phenomenon that recurs cross-linguistically (cf. e.g., the examples gathered by Matras, 2009: 201–203, from different language families, ranging from Indonesian to Vietnamese to Tasawaq; or those provided in the articles covering several languages in Haspelmath and Tadmor, 2009), such that students of language contact have drawn generalizations concerning the borrowability of numerals. In the following scale, proposed by Matras (2007: 61) for the borrowability of word classes, the place of numerals is rather low (emphasis added):

> word classes affected by contact (Matras, 2007: 61): nouns, conjunctions > verbs > discourse markers > adjectives > interjections > adverbs > other particles, adpositions > **numerals** > pronouns > derivational affixes > inflectional affixes

While numerals are generally rather resistant to borrowing – a fact that may depend on the "assumption that all languages have some form of quantification" according to Matras (2007: 50) – there seems to be a difference in degree of borrowability between low and high numerals, in such a way that Thomason and Kaufman (1988: 74) set the borrowing of low numerals at level 3 of their scale, the one that requires, "more intense contact". In this respect, Matras observes the following:

> It appears, then, that while higher and more abstract numerals are vulnerable to borrowing due to their association with formal contexts of use, and numerals in general may become borrowing-prone through intensification

---

From the discussion in Combi (1859: 108f.) and Ascoli (1861: 48f.), it results that the overall demographic size of IRs was over 3000 about the half of the 19th century, while one century later, Tagliavini (1972: 364; first edn. 1949) and Kovačec (1971: 23) reported some 1500 speakers (NIR + SIR). More recently, Filipi (2002: 53) estimates some 90 speakers of SIR and some 80 of NIR, while Vrzić and Doričić (2014: 107) reckon 120 fluent speakers of NIR are left, all over 50 – a steep decrease, which is due partly to depopulation, partly to language shift to Croatian by the speakers still residing in the villages. The truth of the matter is that data are shaky and uncertain: in the same year, Vuletić (2014: 191 n. 9) indicates 53 NIR speakers (out of the 134 inhabitants of Žejane), based on data from the http://www.vlaski-zejanski.com/ website, provided by the first author of the previous paper (Z. Vrzić).

of economic activity in the (potentially) donor language, the proximity constraint protects 'salient' numerals, primarily those below 'ten' or 'five', but sometimes also 'ten' and even 'hundred'. With the latter two exceptions, and the exception of 'zero' whose affinity is toward the formal-abstract numerals, most attested cases add up to support an implicational hierarchy of numeral borrowing: higher > lower numerals. (MATRAS, 2009: 202)

Cross-linguistically well-known cases include e.g., Japanese, where "With a few lexical exceptions, the native system is now used only up to '10'; above '10', even those counters which prefer the native numerals must use the Chinese set" (Martin, 2004: 767). Mixed numeral systems have developed also in language contact between Romance and Slavic. For Molise Slavic, Breu (2013) describes the progression in real and apparent time of numerals borrowed from the adjacent Italo-Romance dialects, with the elderly generation using two alternative forms between '5' (*pet/čing*) and '10' (*desat/dijač*), and only loan numerals from '11' on, while the younger generation has generalized the loans from *sèj* '6' on and no longer uses native *šest* '6' etc. In our case study, the borrowing direction is the other way round, from Slavic into Romance.

4   The Impact of Language Contact on Numerals in Daco-Romanian and in IR

Daco-Romance offers interesting material in this area even outside IR. As is well known, Romanian borrowed *sută* '100' from Old Slavic *sŭto*, which has been adapted as feminine like all *o*-ending neuters among ancient loanwords from Slavic (Mihăilă, 1960; Petrovici, 1962; Buchi, 2006: 75f.; Livescu, 2008: 2648). In addition, Romanian calqued all numerals from '11' on, except inherited *mie* '1000': *unsprezece/nouăsprezece* '11/19' = OBlg. *jedinŭ/devętĭ na desęte*, *doizeci* = OBlg. *dŭva desęti* '20', etc. (cf. e.g., Schulte, 2009: 248). Istro-Romanian, which has been under contact pressure for centuries, goes much further (see Puşcariu, 1926: 153f.; Kovačec, 1966: 65f.; 1971: 117; 1998: 284; Dahmen and Kramer, 1976: 88; Frăţilă and Bărdăşan, 2010: 39; Sala, 2013: 220). As the data in (1) show, cardinal numerals from the Daco-Romance native stock are preserved up to 7 in both branches, while beyond that point, SIR replaces 8 and 9 and NIR 9 and 10 with Croatian loanwords:[6]

---

6 IR data collected in our fieldwork sessions are reported in a simplified IPA transcription: primary stress is marked as V́ (not [ˈV]) and only on non-paroxytonic words; palatal consonants are transcribed [š ž č ǧ] instead of [ʃ ʒ tʃ dʒ]. Please note that due to typographical reasons IPA [a] and [æ] appear as [*a*] and [*æ*] when italicized. Data by other scholars are given in the original orthography. We use the standard orthography for Croatian dialect data (Čakavian).

(1) Cardinal numerals in the languages under discussion: Ro(manian), Lat(in), Cro(atian)

	1 M/F/N	2 M/F/N	3 M,F/N	4	5	6
NIR	ur/ura	doi/do	trei	pɒtru	činč	šɒse
SIR	ur/urɛ	doi/do	trei	pɒtru	činč	šɒse
Ro	unu/una	doi/două	trei	patru	cinci	șase
Lat	ūnus/ūna/ ūnum	duo/duae/ duo	trēs/tria	quattuor	quīnque	sex
Cro	jedan/jedna/ jedno	dva/dvije/dva	tri	četiri	pet	šest

7	8	9	10
šɒpte	opt	dɛvet	dɛset
šɒpte	osəm	dɛvet	zɛtse
șapte	opt	nouă	zece
septem	octō	novem	decem
sedam	osam	devet	deset

From 11 on, all numerals (including '1000') are borrowed: for instance, '11' is *jedanáist* (< Čakavian *jedanajst*; cf. standard Croatian *jedanaest*), and *sto* '100' is a secondary borrowing from Slavic, which replaced the older Daco-Romance adapted loanword (*o*) *sută*. In the higher tenths, the multiples of 10 are all borrowed, while units are Romance from 1 to 4 and Slavic from 5 onwards:[7] for instance, for '25', *dvadeset ši pet* is more frequently used, according to Kovačec (1971: 117), than *dvadeset ši činč*. Climbing back in time, one can follow the spread of Slavic loans as well as other contact-induced phenomena through the sources. For example, Ugo Pellis (cf. Dahmen and Kramer, 1988: 222–224) mentions that in Žejane (at the time, as the whole of Istria, under Italian rule) the Italo-Romance (Venetan) numerals could be used as an alternative, which is no longer the case today. Otherwise, his data match those reported in Pușcariu (1926) and the later sources cited above. On the contrary, Ascoli (1861: 75), based on Combi (1859: 99–139), reports for NIR the Daco-Romance calques *ur pre zaće* '11', *doi zaće* '20', which by that time had been replaced in SIR by

---

7 The resistance of lower numerals against borrowing, observed in language after language, is probably rooted in the cognitively and genetically different substratum of numeric discrimination with small quantities (cf. e.g., the data on human infants and other animal species in Everett, 2017: 149–152; Vallortigara and Panciera, 2014: 52).

Slavic *jedennaist*, *dvaiset*,[8] nowadays the only forms in use in both branches of IR.[9]

Apart from plain object-counting, any numerical expression that is even slightly conventionalized/culturalized tends to select Slavic numerals even more: thus, 'the Three Wise Men' is *tri krɒʎi*, not **trej krɒʎi*. The same goes for the quantification of time lapses and all time indications, where Romance numerals remain in use up to 4 only, as exemplified with 'hours' in (2):

(2)   NIR
         a.   *na vedám na o ura*   'see you at 1 o'clock'
         b.   *na vedám na do/trei/patru ur*   'see you at 2/3/4 o'clock'
         c.   *na vedám na pet/*činč//*šest/*šɒse//sedəm/*šɒpte ur*   'see you at 5/6/7 o'clock'
              (cp. Rom. *la orele cinci/şase/şapte* etc.)

In these phrases, also the preposition *na* is Slavic, as is the form of the noun meaning 'hour': *ura* NIR/*uræ* SIR: this is a reborrowing from Croatian (ultimately from Romance, cf. Kovačec, 1992: 161), but is used in (2b-c) in a form that does not match IR inflection (compare the IR plural *ure* 'hours') but rather corresponds to the Čakavian genitive plural (vs. Croatian *ura*) required in this context by Slavic syntax, which is taken on board in that case form (though, of course, with no case function). The same happens with any quantification expressed in terms of other temporal units (minutes, days, months, years etc.): e.g., *pet dân, šest dân, sédâm dân, ósâm dân* '5/6/7/8 days', *pet minut(i)* '5 minutes', *sedam miseţi* '7 months', *deset let* '10 years' etc. Kovačec (1971: 218f.) shows that space and weight measures behave similarly, though here the Romance form for '5' can still be used: *pet kil/činč kíle* '5 kilograms', *pet/činč métri* '5 metres'. He also reports the following dialogue with an informant:

> Cînd, după ce am obţinut de la ISV expresiile *pet dân, šest dân, sedâm dân, ósâm dân*, l-am intrebat dacă se poate spune şi *činč zíle* etc., răspunsul a fost: *betắri rẹ zice, ali åstez ţi se rẹ ắrde* 'bătrínii ar spune, însă astăzi ai fi luắt în deridere' [When, after obtaining from ISV (= an informant from Žejane, born in 1902) the expressions *pet dân*, etc., I asked him whether one can also say *činč zíle* etc., the answer was: 'old men would say so, but today you'd be mocked (for saying it)'].

---

8   Ascoli (1861: 75) actually writes "*dvaiste*", which might be a misprint, given that the Čakavian form for '20' is *dváiset* (compare Croatian *dvádeset*). NIR *dvajset ši ur/doj/trej* '21, 22, 23' were recorded in Ugo Pellis' fieldwork in 1926–1935 (cf. Dahmen and Kramer, 1988: 224).

9   Ascoli (1861: 75) also reports NIR *nuk* '9', not confirmed by any other sources and qualified as "obscure" by Puşcariu (1926: 153).

Thus, to the competence of a NIR informant interviewed when he was about 60, in 1961–1962, *činč zile* was a ludicrous archaism. Today, according to our informants, the Romance noun form *zile* (*zi* 'day' < Lat. DIEM) can be used in the phrase 'five days', to talk about *n* days qualifying them in terms of properties, but not in order to denote a time interval of *n* days (i.e., not as a time measure): one can say e.g., *činč zile fóst=av fine* 'the five days were nice', as opposed to *fóst=am ped dən in rika* 'I spent (lit. was) five days in Rijeka'.

Sala (2013: 220), citing data from Kovačec (1966: 65f.; 1968: 99f.), Petrovici and Neiescu (1964: 191; also in 1965: 356), summarizes this situation as follows:

> The co-presence of Romance and Croatian forms from 'five' to 'eight' is by no means a matter of 'free variation'. Rather, the Croatian forms must be used in 'lexical measure phrases' (phrases expressing characteristic units of measurement, such as time, weight and distance); moreover, they must be combined with a Croatian noun, where one is available, showing Croatian noun morphology.

The same selectional restrictions on borrowed vs. native numerals now described with regard to measures hold true even when the word at issue (indicating e.g., a time lapse) is omitted, as in the exchange in (3a), where Slavic *pet* must be used even if *let* 'years' is gapped in B's answer, to be compared with (3b), where in specifying the number of chickens, rather than years, **pet* is ungramm2atical.

(3) a. A. *kəts*   *ɲn*   *ai̯*   *fost*   *la*   *soldót*
            how_much\M.PL   year(M)\PL   have.PRS.2SG   been   to   soldier
            'how many years have you been in the army?'

     B. *fóst=*   *am*   *samo*   *doi̯/*   *patru (ɲn)//*   *samo*
          been=   have.PRS.1SG   only   two.M/   four (year(M)\PL)//   only
          pet/*činč   *(let)*
          five   (years)
          'I've been there for only 2/4/5 years.'

   b. A. *kət-e*   *ɣaʎír*   *ʋri*
            how_much-F.PL   chicken(F)[PL]   have.PRS.2SG
            'how many chickens have you got?'

     B.   *samo*   *činč/*pet*
           only   five
          'only five'

All this has been duly described in the literature on IR, and serves as background information to introduce the novel data on which our study focuses.[10]

5   Calqued and Borrowed Numerals with *Pluralia Tantum* Nouns in IR

The Romance languages, not unlike Latin and many other inflecting-fusional languages, have *pluralia tantum* nouns (henceforth PT). In Latin, as seen in (4), these nouns could be determined through the numeral 'one', in a context in which the morphosyntax of number (plural number, via agreement with the head noun) was in conflict with the semantics of the numeral, denoting *one* real-world entity:

(4) *adeo*   *ut*   *un-a*   *castr-a*   *iam*
     to_the_point   that   one-NOM.N.PL camp(N)-NOM.PL   already

     fact-a   *ex*   *bin-is*   *vid-ere-nt-ur*
     become-NOM.N.PL   out_of   two-ABL.PL   seem-SBJ. IMPF-3PL-PASS
     'so that the two camps seemed to have become just a single one' (CAES., *B.C.* 1.74)

At first glance, IR behaves like Latin in this respect (examples are from NIR; most of the phrases in (5) would be identical in SIR):

(5) *ur-e*   *braɣɛš-e/mudant-e/postol-e/ocʋl-e/škʋr-e/novin-e/ɣrʋbʎ-i*
                    (or -e)/vil-e
     one-F.PL   trousers/underpants/shoe/spectacles/scissor/newspaper/
                    rake/pitchfork(F)-PL
     'a pair of trousers/underpants/shoes/glasses/scissors//a newspaper/
     rake/pitchfork'

In (5), a series of PT nouns, all feminines, select the F.PL form of the numeral *ur-e* 'one'.[11] Many of these nouns denote 'objects made up of two like parts' (Payne and Huddleston, 2002: 340; cited in Corbett, 2019: 54 n. 2) – e.g., *vil-e*

---

10   Though not in focus in the present paper, these facts are highly interesting per se, as they seem to represent a case of "parallel system borrowing" that could be added to those discussed e.g., by Kossmann (2010).

11   Note that (4)–(5) show that, while occurrence of PT nouns in a singular indefinite context is rare cross-linguistically (cf. English **a pant(s)*, **a scissor(s)*), there are languages such as Latin or NIR which are exceptional in this respect, so that this cannot be regarded as a universal property of the grammar of PT nouns (*pace* Klockmann, 2017: 29).

'pitchfork' (a traditional pitchfork had two tines) –, which is a frequent case cross-linguistically for PT, though it need not be. In fact, *γrʊbʎ-i* 'rake' also denotes an object composed of a set of ordered parts, which are however more than two, and this semantic criterion hardly applies to *novine* 'newspaper' which parallels English *news* or its Hebrew equivalent *xadašót*, etymologically and morphologically. What crucially defines the nouns in (5) is that they have only plural morphology and invariably require plural agreement. Thus, they match, as to both the inflectional and syntactic criterion, Corbett's (2019: 96) definition of PT as nouns that "have only the plural".[12]

Since historically in Daco-Romance the Latin neuter plural has merged with the feminine plural (see Section 6, (27)), selection of the feminine plural quantifier in (5) could be regarded in principle as inherited from Latin. However, both the ecology of IR and the comparative (Daco-)Romance picture suggest a different explanation. Daco-Romance does not retain the numeral agreement pattern found in Latin (4), but rather replaces it with periphrastic classifiers, as exemplified with Romanian *pereche/perechi de* 'pair/-s of' in (6a):

(6)   a.   *o pereche/două perechi de ochelari*   Romanian
            'a pair/two pairs of spectacles'
      b.   *un/doj pʊr de škʊre*   SIR
            'one pair/two pairs of scissors'

This option is also available in IR (6b) while in Romanian it is compulsory, just as it is in Italian.

(7)   a.   *un paio di forbici/occhiali*   Italian
            'a pair of scissors/spectacles'
      b.   **une/-a forbici*   **uni/-o occhiali*
            one-F.PL/-F.SG scissors   one-M.PL/-M.SG spectacles

In both Eastern Romance standard languages, plural forms of 'one' are never adnominal quantifiers but can only be indefinite pronouns/adjectives ((8a-b); an option available in IR as well, (8c)).

(8)   a.   *gli uni e gli altri / le une e le altre*   Italian
            'the ones and the others(M/F)'

---

12   Corbett takes issue with definitions of PT nouns which refer to both form and meaning (e.g., "A noun which is plural in form but singular in meaning", Trask, 1997: 172) and argues instead for a definition based on purely formal criteria (inflectional and syntactic).

    b.   *unii ochelari*   Romanian
         'some (pairs of) spectacles'/*'one pair of spectacles'

    c.   *uri kər / ure vɒč*   NIR
         'some dogs/some cows'

    d.   *unas gafas*   Spanish
         'some pairs of glasses' or '(just) one pair of glasses'

By contrast, Spanish has preserved the Latin option in (4), i.e., the pluralizability of 'one' with PT nouns, so that (8d), unlike its Italian and Romanian counterparts, is ambiguous. Plural agreement on the numeral quantifier 'one' with PT nouns is encountered occasionally in varieties which acquired it arguably via language contact. As the data in (9) indicate, Sissanese, a variety of Istrioto spoken in Sissano (South-eastern Istria, near Pula/Pola), is a case in point (see Giudici and Zanini, 2021).

(9)   *un-e*   *fórfež-e*   Istrioto dialect of Sissano (Croatia)
        one-F.PL   scissors(F)-PL
        'a pair of scissors'   (cp. Croatian *jedn-e škar-e* 'id.')

These structural facts, along with the general attrition of IR under extreme contact, suggest that it is a priori more plausible to assume that the selection of the plural form of 'one' in IR with PT nouns such as those in (5) is a contact-induced phenomenon. The data in (10) display the Slavic model, of which (5) is most likely a replica.

(10)   *jedn-e*   *hlač-e/gać-e/postol-e/naočal-e/*   Croatian
                        škar-e/novin-e/grablj-e/vil-e
        one-F.PL   trousers/underpants/shoe/spectacles/scissor/newspaper/rake/pitchfork(F)-PL
        'a pair of trousers/underpants/shoes/glasses/scissors//a newspaper/ rake/ pitchfork'

Comparison with (5) reveals that most IR nouns in the latter example are loans from Croatian, including some words from the local Čakavian dialects such as *postol-e* (cp. standard Croatian *cipel-e*), and the two lists could be made even more alike if one considers that also *braɣɛš-e*, *mudant-e*, and *ocɒl-e*, though ultimately of Latin descent, occur in Croatian dialects too and thus could be

---

13   Thanks to one anonymous reviewer for pointing this out to us.

Slavic borrowings just as well.[13] The data in (11)–(12) display the Croatian paradigm from which the numeral form in (10) is picked, with one example for each gender/number value combination (all in the nominative). The examples in (12), from Leko (2009: 25) and Corbett (2019: 78f.), illustrate agreement with PT nouns.

(11) a. *jedan*   *tanjur*   *je*   *hladan*   Croatian
        one[NOM.M.SG] plate(M)[NOM.SG] be.3SG cold[NOM.M.SG]
        'one plate is cold'

    b. *jedn-a*   *kav-a*   *je*   *hladn-a*
        one-NOM.F.SG   coffee(F)-NOM.SG be.3SG cold-NOM.F.SG
        'one coffee is cold'

    c. *jedn-o*   *piv-o*   *je*   *hladn-o*
        one-NOM.N.SG   beer(N)-NOM.SG be.3SG cold-NOM.N.SG
        'one beer is cold'

(12) a. *jedn-i*   *svat-ov-i*   *su*
     one-NOM.M.PL wedding.procession(M)-PL-NOM.PL be.3PL

     stig-l-i
     arrive-PST-M.PL
     'one wedding procession arrived.'

    b. *jedn-e*   *naočal-e*   *su*   *puk-l-e*
       one-NOM.F.PL   spectacles(F)-NOM.PL   be.3PL break-PST-F.PL
       'one pair of spectacles broke'

    c. *jedn-a*   *kol-a*   *su*   *stig-l-a*
       one-NOM.N.PL   carriage(N)-NOM.PL   be.3PL arrived-PST-N.PL
       'one carriage arrived'

Thus, the IR F.PL form *ur-e* in (5) calques Croatian *jedn-e*. The table in (13) shows the complete paradigm which Kovačec (1971: 112) gives for the indefinite article.[14]

---

14   Most forms are homophonous with those of the numeral *ur* 'one', out of which they grammaticalized. Only in the nominative/accusative case, phonetic reduction is observed, which distinguishes M.SG *ən* and F.SG *o* in (13) from the numerals *ur/ura* in (1). The neuter form *uro* – whose *-o* ending is of Slavic origin, as discussed in Section 6 – is mostly used

(13)   M   F   N   The indefinite article in IR

SG   NOM/ACC   *ən*   *o*   *uro*

        OBL   *urvǽ*   *urʎǽ*   –

PL   NOM/ACC   *uri*   *ure*   –

        OBL   *urorǽ*   *urorǽ*   –

While the occurrence of *ure* with PT nouns seen in (5) is observed in both NIR and SIR, the two branches of IR part ways as it comes to quantifying PT nouns with the numerals 'two', 'three' and 'four'. The following examples are from Žejanski (NIR).[15]

(14) a. *vl-ót=am*   *dvoje/*do/*doj*   *škɒr-e/ɣrɒbʎ-i/vil-e*   NIR
          take-PTP=have.1SG   two.X/two.F/two.M   scissors/rake/pitchfork(F)-PL
          'I took two pairs of scissors/two pitchforks/rakes'

      b. *dæ=m*   *troje/*trɛj*   *novin-e/škɒr-e*
          give.IMP=1SG   three.X/three.M=F   newspaper/scissors(F)-PL
          'give me three newspapers/pairs of scissors'

      c. *kumparót=am*   *dvoje/[??]do/*doj*   *novin-e*
          buy-PTP=have.1SG   two.X/two.F/two.M   newspaper(F)-PL
          'I bought two newspapers'

All the feminine PT nouns in (14) require a special form of '2' and '3' which, as shown in (15), is distinct from the ones occurring with ordinary count nouns.

---

pronominally, but can marginally be used as an adnominal numeral quantifier or indefinite article as well.

(i)   *lǽ-a=ts*   *ur-o*   *srebro*   SIR (Jesenovik)
        take-IMP.2SG=1SG   one.N   silver(N)
        'take one/a silver object'

15   The notation 'two.X', 'three.X' in the glosses will be explained in due course. In SIR, the ordinary feminine form is selected with such nouns, while **dvoje/*troje* are unacceptable: *do/*dvoje škɒre* 'two pairs of scissors', *ste do/*dvoje ǧornɒle* 'these two newspapers'. Quantification of such nouns can also be realized periphrastically, as shown in (6b).

(15) a. *trej//do/*doj*   *pete/urec(ʎ)i/bɒrba/ženske/məre/surár/vɒč/*
                                    metle/fɶate
          three//two.F/two.M   heel/ear/beard/woman/hand/sister/cow/
                                    broom/little girl(F)
      b. *trej//doj/*do*   *oc(ʎ)i/omir/dints/kər/kúvete/pičore/kůpure*
          three//two.M/two.F   eye/man/tooth/dog/elbow/foot/head(M)

While 'three' is invariable for gender in Daco-Romanian as well as in IR, 'two'
inflects for gender in all Romanian dialects, as illustrated for IR in (15) (from
now on, all examples will be from NIR whenever not otherwise specified).

   Hence, the form of the numeral 'two' occurring in (14) is a third distinct
form which adds to the two inherited ones. (The notation 'two.x' in the glosses
means that its categorial status is still to be defined, an issue to which we will
return in Section 8). While the form of 'one' occurring with PT nouns in (5) was
arguably calqued on Croatian, *dvoje* and *troje* in (14) are borrowed, a fact that,
to the best of our knowledge, has been mentioned so far, in the literature on IR,
only in half a line by Puşcariu (1926: 156).

> Substantive exprimând o parte sau o sumă de atâtea lucruri de acelaş fel
> găsim: *dvoi̯e* (*bi̯țvi*) < cr. *dvoje* şi (< ital.) *pắi* [...] 'pereche' [Nouns which
> refer to a part or a sum of several things of the same kind: *dvoi̯e* (*bi̯țvi*)
> 'two pairs of socks' and (from Italian) *pắi* 'pair']

Puşcariu's wording and his quoting of a periphrastic classifier in the same
context make clear that he is referring to the kind of quantification we are
interested in. Curiously, his example is drawn from a SIR oral text collected in
Šušnjevica in 1904 and printed in Puşcariu (1906: 180). Today, our SIR inform-
ants reject *dvoje* and *troje* categorically, in spite of using, of course, the homoph-
onous collective numeral forms when they speak Croatian. This may perhaps
indicate that the change whose results are evident in (14) was incipient at that
time in SIR too, where however it did not become established.

   In NIR, a borrowed form of the numeral occurs with the same nouns for '4'
as well, as seen in the series of examples in (16a), with feminine PT nouns, to
be compared with those with plain count feminines in (16b).

(16) a. *ure/   dvoje/   troje/   četvore/   *pétero/   *šestore   novine/škɒre*
          *o/   *do/   *trej/   (?)patru/   činč/   šɒse
          'one newspaper/pair of scissors//2/3/4/5/6 newspapers/pairs of scissors'

    b.   *o muʎʋra*   //   *do/trej/patru/činč/šʋse muʎér*
         'one wife'   //   'two/three/four/five/six wives'

Since *četvore* – also borrowed from Croatian – is uninflected, it will not detain us any further here. The examples in (17) show that our PT nouns consistently select feminine plural agreement on all agreement targets other than the numerals, illustrated here with demonstratives and qualifying adjectives.

(17) a. *čʋl-e*   *novin-e*   *z* *betər-e*   */*betər-i*
       DEM.DIST-F.PL newspaper(F)-PL are old-nom.F.PL /old-NOM.M.PL
       'those newspapers are old'

    b. *čʋst-e*   *postol-e* *s* *usk-e*   */*usk-i*
       DEM.PROX-F.PL shoe(F)-PL are tight-NOM.F.PL /tight-NOM.M.PL
       'these shoes are tight'

The Slavic model is exemplified in (18), where the collective numeral adjectives (also termed "numerical adjectives", see e.g., Lučić, 2015) are shown, which are selected with PT nouns of the three genders (data from Stefanović, 2014: 49; Lučić, 2015: 4f.; Kim, 2009: 114).

(18) a. *ov-i*   *dvoj-i/troj-i/četvor-i/peter-i*   Croatian
       those-NOM.M.PL   two/three/four/five-NOM.M.PL
       svat-ov-i
       wedding.procession(M)-PL-NOM.PL
       'those 2/3/4/5 groups of wedding guests'

    b. *ov-e*   *dvoj-e/troj-e/četvor-e/petor-e* *čarap-e*
       those-NOM.F.PL two/three/four/five-NOM.F.PL socks(F)-NOM.PL
       'those 2/3/4/5 pairs of socks'

    c. *ov-a*   *dvoj-a/troj-a/četvor-a/petor-a* *kol-a*
       those-NOM.N.PL two/three/four/five-NOM.N.PL carriage(N)-NOM.PL
       'those 2/3/4/5 carriages'

In (18), for simplicity, only nominative forms are listed, since it is the F.PL nominative forms (*dvoje*, *troje*) that have been borrowed into NIR: the borrowing process probably started with whole NPs headed by PT nouns of Croatian origin such as e.g., *dvoje novine* 'two newspapers'.

In the Slavic model system, those *e*-ending forms (18b) are homophonous with the non-agreeing collective numerals selecting genitive case on the noun they govern (Lučić, 2015: 5; Kim, 2009: 119).

(19)   a.   *petero*   *američk-ih*   *studenat-a*   Croatian
            five   American-GEN.M.PL   student(M)-GEN.PL
            'five American students'
       b.   *troje*   *telad-i*
            three   calf(F)-GEN.PL
            'three calves'
       c.   *dvoje*   *djec-e*
            two   child(N)-GEN.PL
            'two children'

This homophony is irrelevant structurally, because in terms of part of speech the model is (18b), not (19). However, it may have played a role enhancing the token frequency and hence the saliency of the forms *dvoje*, *troje*, *četvore*. All of the collective numerals in (18)–(19) are inherited from Common Slavic, were attested in Old Church Slavonic and are found in the modern languages, from Russian to Polish to BCS (cf. e.g., Brugmann, 1907: 22, 44f.; Kim's, 2009 comparative syntactic study; or the further references in Corbett, 2019: 93). In modern spoken BCS they seem to be falling out of use at present according to the results of Stefanović's (2011: 712) corpus study:

> l'emploi normatif des adjectifs numéraux, s'il se laisse observer ça et là, est peu vivant, relativement limité et tend à être remplacé par celui des numéraux cardinaux, avec ou sans le lexème *par* « paire » [plus précisément, les adjectifs numéraux (et les numéraux collectifs) sont concurrencés par les cardinaux correspondants pour 2, 3, 4, tandis qu'à partir de 5, ce sont presque uniquement les cardinaux qui sont utilisés]. ['the standard use of numeral adjectives, which one can observe at times, is not alive and well but rather limited and tends to be replaced by that of cardinal numerals, with or without the lexeme *par* 'pair' [more precisely, numeral adjectives (and collective numerals) are in competition with their cardinal counterparts for 2, 3, 4, whereas from 5 on, almost only cardinal numerals are used]'].

We are not aware of corpus-based studies on collective numerals in spoken Čakavian dialects, so we cannot speculate on their frequency of usage in the local contact varieties of NIR. Be that as it may, their existence in Čakavian – just as in standard Croatian (cf. e.g., Stevanović, 1989: 322 f.; Šipka, 2007: 121) –

is beyond doubt (*pace* Pranjković, 2000: 87): they are described by Ribarić (1940: 115) for Vodice, some 13 km WNW of Žejane and, as one anonymous reviewer kindly points out based on fieldnotes by Silvana Vranić, they occur in Mune Čakavian (3 km WNW of Žejane), the neighbour village's dialect used by Žejanski speakers (cf. Małecki, 1930: map. 4): e.g., *dvoje grablje/škare* 'two rakes/pairs of scissors', *dvoja kola* 'two cars', *četveroja vrata* 'four doors'. The same goes for the Čakavian variety of Orbanići, some 80 km to the WSW: *dvoji očenaši* 'two rosaries' (PT noun), *četvoreh postoli* 'of four pairs of shoes' (Kalsbeek, 1998: 178).

As said above, the first and only documentation of borrowed collective numerals in IR – the half line by Puşcariu (1926: 156) quoted above – goes back to the early 20th century, and the fact that it refers to SIR, where the change eventually aborted, may indicate that it was in the beginning at that time.

6   The Morphosyntactic System Into Which Collective Numerals Have been Borrowed

We now move on to discussing the impact that the borrowing of collective numeral forms from Croatian has had on the grammar of IR. In NIR, these forms have entered a grammatical system that distinguished two number values (singular vs. plural) and three gender values: masculine vs. feminine (inherited) vs. neuter (recently borrowed from Slavic into both IR branches), as exemplified with the paradigm of a class one adjective in (20a), compared with its Croatian counterpart in (20b) (Petrovici, 1967: 1525; Kovačec, 1971: 85).

(20) The paradigm of the adjective 'good'

	a. IR	SG	PL	b. Croatian	SG	PL
	M	*bur*	*bur-i*		M	*dobar*	*dobr-i*
	F	*bur-a* NIR/*-æ* SIR	*bur-e*		F	*dobr-a*	*dobr-e*
	N	*bur-o*	--		N	*dobr-o*	*dobr-a*

A number of studies have shown that the agreement marker *-o* occurring in IR class one adjectives (20a) was borrowed from Croatian, where it occurs in forms such as *dobr-o* in (20b). Once extracted, the affix applied to adjectives of the recipient language, including those from the inherited stock such as *bur-o*, resulting from *bur* (< Lat. BONUM) + *-o*. The introduction of this morph in IR occurred as new *o*-ending neuters such as *srebro* 'silver' entered the language without morphological adaptation, ousting the earlier strategy which produced adapted loans such as e.g., *okn-a* (NIR)/*-æ* (SIR) 'window(F)' < Sl. *okno* (Kovačec, 1998: 134; see what has been said in Section 4 on ancient loanwords

from Slavic into Daco-Romance, while commenting on borrowed *sutǎ* '100', and see Kovačec 1966: 67 on the replacement of this earlier strategy in IR with non-adapted borrowing of *o*-neuters).[16]

(21) a. *zlɒt-o=j*   *drɒɣ-o,*   *srebr-o*   N IR
          gold(N)-SG=be.PRS.3SG   expensive-N   silver(N)-SG
          nu=j   *drɒg-o*
          NEG=be.PRS.3SG   expensive-N
          'gold is expensive, silver is not'

      b. *tsəsta*   *srebr-o=j*   SIR
          DEM.PROX:N   silver(N)-SG=be.PRS.3SG
          ɒb-o/*ɒb
          white-N/white[M.SG]
          'this silver is white' (Šušnjevica, Puşcariu, 1906: 18, 39)

The *-o* ending seen in *drɒg-o* and *ɒb-o,* selected by the Slavic borrowed nouns *srebro*, *zlɒto* and the like, must have entered the language first in loan-adjectives such as *drɒgo*, to then spread to adjectives from the inherited stock such as *ɒbo* (from *ɒb* < Lat. ALBUM + *-o*), etc. Neuter *o*-forms of IR adjectives (including native ones, such as *buro* 'good-N', *groso* 'big-N') are reported as early as Puşcariu (1926: 150f.), quoting occurrences from the oral literature edited in Puşcariu (1906). Those occurrences, however, are invariably found in contexts in which no nominal controller is present, and can thus be interpreted as predicative adverbs (22a) or as instances of default agreement (22b).

(22) a. *jel'-l' odgovorescu ke jel'-s siromås si grumbo ke åru*
          'they answer to him that they are orphans and are doing badly'
      b. *ţa-i lǎhco*
          'it's easy'

Both uses are still observed today. In particular, the use of neuter agreement in default contexts, where there is no noun to trigger gender agreement, is obligatory.

---

16   Replacement of earlier adapted loans has been gradual. Thus, while Kovačec's (1998: 225) dictionary only reports *zlåto* 'gold(N)' for Žejane, Kovačec (1963: 34) says that his Žejanski informants aged 50–70 used *zlɒta=j drɒg-a* 'gold(F) is expensive-F.SG' and rejected as ungrammatical *zlɒto=j drɒg-o* 'gold(N) is expensive-N', which was instead normally used by his younger informants (aged 12–17). We have recorded *zlɒtæ* 'gold(F)', *zlɒta=j drɒgæ* 'gold(F).DEF is expensive' from an informant from Šušnjevica born in 1954.

(23) a. *ʋstez*   *je*   *ɣrumb-o*   NIR
          today   be.PRS.3SG   ugly-N
          'the weather is ugly today'

      b. *tse=j*   *de*   *nov-o?*   SIR
          what=be.PRS.3SG   of   new-N
          'what's new?'

Comparing Puşcariu's (1906) data with his own fieldnotes from the early 1960s, Kovačec (1963: 33–36; 1966: 67–70) showed that these *o*-forms had started to agree with nouns much later, and that the change was then still ongoing:

> [C]ertains substantifs en -*o* empruntés au croate (emprunts probable-ment assez récents) ont commencé à s'accorder avec les formes adjec-tivales neutres en -*o* ['some *o*-ending nouns borrowed from Croatian (probably relatively recent loans) have started to require, for agreement, *o*-ending adjective forms'] (KOVAČEC, 1966: 68).

Nowadays, as shown by ungrammatical **ʋb* in (21b), this agreement form must be categorically used with all and (almost) only the cited borrowed mass nouns. This is not just alliterative concord, given that borrowed mass nouns take the same *o*-agreement even if they do not end in -*o*, as long as they stay neuter, as shown for SIR in (24a) (the same Croatian loanword, on the contrary, has been recategorized as feminine in NIR because of its inflection class; see (24b)).

(24) a. *uʎ-a=j*   *drʋɣ-o*   SIR (Jesenovik)
          oil(N)-DEF.SG=be.PRS.3SG   expensive-N/
          *drʋɣ-a/*drʋɣ
          expensive-F.SG/expensive[M.SG]
          'oil is expensive'

      b. *uʎ-a*   *nu=j*   *drʋɣ-a/*   NIR
          oil(F)-DEF.SG   NEG=be.PRS.3SG   expensive-F.SG/
          *drʋɣ/*drʋɣ-o
          expensive[M.SG]/expensive-N
          'oil is not expensive'

Conversely, neuter *o*-agreement with other non-neuter mass nouns – either native (such as *kʋrne* 'meat(F)') or borrowed (such as *bronza* 'bronze(F)') – is generally judged ungrammatical.

(25) a. *bronz-a*   *nu=j*   *ašá drɒɣ-a/*   NIR
          bronze(F)-SG   NEG= be.PRS.3SG so   expensive-F.SG/
          *drɒɣ/*drɒɣ-o
          expensive[M.SG]/expensive-N
          'bronze is not so expensive'

      b. *kɒrn-a/sɒr-a*   *=j*   *bur-a/*
          meat(F)-SG/salt(F)-SG   = be.PRS.3SG   good-F.SG/
          *bur/*bur-o
          good[M.SG]/good-N
          'meat/salt is good'

      c. *unt-u/*   *čest-a*   *otsét*   *je*
          butter(M)-DEF.SG/   DEM.PROX\M-SG   vinegar(M)-SG   be.PRS.3SG
          bur/*bur-o/*bur-a
          good[M.SG]/-N/-F.SG
          'butter/this vinegar is good'

However, some exceptions – pointing to incipient semantic agreement – are reported by Kovačec (1966: 68):

> Quelle est la pression du neutre croate, on le voit d'après le fait que deux substantifs *vir* e *lapte*, qui en istroroumain sont masculins, s'accordent quelquefois, sous l'influence des substantifs croates correspondants de genre neutre *vino* et *mlijeko*, 'faussement' avec les formes neutres des adjectifs. ['How strong the pressure of the Croatian neuter is, is seen from the fact that the two nouns *vir* 'wine' and *lapte* 'milk', which are masculine in IR, sometimes 'wrongly' take neuter agreement on adjectives under the influence of the corresponding Croatian neuter nouns *vino* and *mlijeko*'].

The same vacillation is still observed in the competence and production of our informants:

(26)   NIR
      lɒpte-le   *nu=j*   *drɒɣ-o*   */drɒɣ*
      milk(M)-DEF.SG NEG= be.PRS.3SG expensive-N   /expensive[M.SG]
      'milk is not expensive'

As an output to the gradual spread now reviewed, the *o*-neuter has become established as a fully functional gender value. Since it is used in default contexts and with mass nouns which have been borrowed in the singular form

(with no plural), the neuter *o*-agreement in IR is number-defective, occurring in the singular only.

Since with reference to Daco-Romance the term neuter is usually employed to denote another distinct gender value, a word on the latter is in order here. Consider the (Daco-)Romanian gender system as illustrated in (27) with agreement on definite articles and a class one adjective (see Corbett, 1991: 151; Loporcaro, 2018: 92).

(27)   SINGULAR   PLURAL   Romanian

a. *pantof-ul*   *e bun*   *pantofi-i*   *sunt bun-i*

M   shoe(M)-DEF.M.SG is good[M.SG] shoe(M)-DEF.M.PL are   good-M.PL

b. *vin-ul*   *e bun*   *vinuri-le*   *sunt bun-e*

N   wine(N)-DEF.M.SG is good[M.SG] wine(N)-DEF.F.PL   are   good-F.PL

c. *băutur-a*   *e bun-ă*   *băuturi-le*   *sunt bun-e*

F   drink(F)-DEF.F.SG is good-F.SG   drink(F)-DEF.F.PL are   good-F.PL

‘the shoe/wine/drink is good’   ‘the shoes/wines/drinks are good’

Nouns such as *vin* in (27b) are traditionally termed 'neuter' in Romanian descriptive grammar, which assumes that this is a third gender, distinct from both masculine and feminine. In Loporcaro (2018: 92–109), alternative two-gender analyses of Romanian are discussed and rejected, showing that a three-gender analysis is the only one in keeping with the following widely assumed definitions, which we assume here too.

(28) a. "Genders are classes of nouns reflected in the behavior of associated words." (Hockett, 1958: 231; Corbett, 1991: 1)

b. "We should […] differentiate controller genders, the genders into which nouns are divided, from target genders, the genders which are marked on adjectives, verbs and so on." (Corbett, 1991: 151)

Under such definitions, the Romanian neuter, which is inherited from common Daco-Romance since it occurs in all of its four dialect branches (Petrovici, 1967: 1523), is a third controller gender, selecting agreement targets that are fully syncretic (with the masculine in the singular, with the feminine in the plural). As argued in Loporcaro (2018: 222), these syncretisms result from mergers. In other words, the Romanian neuter is inherited from Latin: only, it has turned from a target to a controller gender with alternating agreement.

Back to IR, while this language has acquired a new (mass) neuter via borrowing, by the time of Petrovici's (1962) study it had lost (NIR) or was in the process of losing (SIR) the inherited alternating neuter, which merges with the

masculine also in the plural. As a result, nouns like those in (29), which used to select alternating agreement (and still do in Daco-Romanian, (27b)), now select plural masculine agreement in NIR and consequently have been reassigned to the masculine.

(29)   NIR
    *češc-i/*čvšt-e*   *doj/*do*   *kúvet-e/*
    DEM.PROX\M-M.PL/DEM.PROX\F-F.PL   two:M/F   elbow(M)-PL/
    kŕp-ure
    head(M)-PL
    'these two elbows/heads'

Thus, IR shows that contact-induced pressure may result not only in the simplification of grammar, even if the latter is most often the case cross-linguistically: "language contact, especially when extensive L2 learning is involved, is a main source of complexity reduction (grammar simplification)" (Karlsson et al., 2008: viii; see also Arkadiev and Gardani, 2020). On the contrary, the rise of the *o*-neuter in IR is a case of contact-induced complexification, rather than simplification, of a language's grammar (on a par with the others discussed, with reference to gender, in Loporcaro, 2018: 51f.).

7   Intermezzo: the Slavic neuter and the Double Complexification of the NIR Gender System

In Section 6, we have seen that the *o*-neuter had already been borrowed by the time grammatical descriptions of IR became available and acquired its syntactic function of conveying agreement with neuter *o*-loans by the mid-20th century. A still later development was first reported as interference occurring in the performance of some (mostly rather attrited) Žejanski speakers by Kovačec (1963: 35).

> Nous n'avons pas rencontré de pluriels de substantifs neutres, sauf dans deux cas douteux. Une jeune fille de 21 ans, qui a vécu assez longtemps à Rijeka où elle faisait ses études, en traduisant un texte croate a employé *ašǻva pítańa* 'de telles questions' comme pluriel. L'autre exemple, si l'on applique des critères croates à l'analyse, ne pourrait que confirmer indirectement l'existence d'un embryon du pluriel neutre. Pour 'les enfants étudient' nous avons noté à Žejane *diţa se-nveţavę* et *diţa se-nveţavés* (pl. neutre ?), mais il pourrait s'agir ici seulement d'un calque du pluriel

croate dans la forme verbale [...], le substantif étant pris comme un collective de genre féminin (ce qui se rencontre quelquefois dans les dialectes čakaviens environnants: *dica se uči* à côté de *dica se uču, učiju* [...]) ['I did not come across any plurals of neuter nouns, except for two dubious cases. A 21-year-old girl, who lived for a relatively long time in Rijeka where she was studying, used, in translating a Croatian text, *ašǎva pítańa* 'such questions' as a plural. The other example, if analysed by Croatian criteria, could not but confirm indirectly the existence of an embryo of the neuter plural. For 'children study' I have recorded in Žejane *diţa se-nveţavę́* and *diţa se-nveţavés* (neuter plural?), though it could be nothing else than a calque of the Croatian plural in the verb form [...], whereby the nouns could be taken as a feminine-gender collective (which is met with at times in the neighbouring Čakavian dialects: *dica se uči* 'children learn:SG' alongside *dica se uču, učiju*' children learn/are learning:PL' [...])']

Kovačec's (1998: 69) dictionary follows this latter interpretation (singular "collective" noun with semantic plural agreement) in specifying the entry *di:tsa* 'children', as follows: "**diţa** ž (zbirno)" ['**diţa** f(eminine) (collective)']. The same grammatical specification is given for *γospoda* '(wealthy) gentlemen', reported in Kovačec (1998: 85) alongside a separate entry for *γospodɒr* and *γospodín* 'seigneur', while *di:tsa* lacks a similar M.SG counterpart altogether.

The agreement pattern described as occasional interference and/or semantic agreement by Kovačec (1963: 35) has now become established in NIR, where the cited nouns – plus *vlastela* 'noblemen', not registered in Kovačec's (1998) dictionary – select unambiguously a type of syntactic agreement that was not observed in IR prior to borrowing. In today's NIR, there is little doubt that those three lexemes must be regarded as PT nouns, for they obligatorily select plural verb agreement, as the data in (32) and (33) show. This is in keeping with their origin, as they are all homophonous with the model Croatian forms, among which *dica* is the local Čakavian dialect variant (vs. standard Croatian *djeca*) with *i* < Proto-Slavic **ě* found in the dialects of the Mune area (cf. Małecki, 1930: map. 4). In the source language, these are plurals from non-defective paradigms (*djeca*) or can occur with either plural or singular agreement (*gospoda*).

In the whole of BCS, the noun for 'children' presents an intriguing and much-discussed situation: while it serves as a plural to *dijete/dete* 'child(N)' and has plural morphosyntax (i.e. agreement), morphologically it inflects like the singular of feminine nouns ending in *-a* (such as e.g., *žena* 'woman'; cf.

Corbett, 1983: 76–81; 2000: 187f.; 2007: 39; Despić, 2017): this is seen in (19c) above, where *djec-e* is genitive plural morphosyntactically but has an *-e* ending which corresponds morphologically to the genitive singular of the feminine *a*-class: compare *žene* vs *ženā* 'woman' (GEN.SG vs. GEN.PL). By contrast, *gospoda* selects either singular or plural agreement, with a semantic difference ('gentry' vs. 'lords') discussed in Stankiewicz (1983: 157):

(30) a. *gospod-a*   *se*   *nije*   *meša-l-a*
          gentry(F)-NOM.SG   REFL   NEG:be.PRS.3SG   mingle-PST-F.SG
          s   *narodom*
          with   people(M)-INSTR.SG
          'the gentry did not mingle with the people'

      b. *gospod-a*   *se*   *jesu*
          gentry(N)-NOM.PL   REFL   be.PRS.3PL
          zavadi-l-a
          quarrel-PST-N.PL
          'the lords had a falling out'

The three nouns behave differently in the two IR branches. Our SIR informants do not accept *di:tsa* and *vlastela* as possible IR words but do use *gospoda* – in exactly the way reported in Kovačec (1998: 85) only for NIR – as a singular with collective meaning. This can be predicated of a plural NP, as shown in (31a), but when employed as a subject never takes plural verb agreement (31b).

(31) a. *jeʎ=əz*   *gospod-a*   NIR
          3M.PL=be.PRS.3PL   gentlemen
          'they are (wealthy) gentlemen'/'they are gentry'

      b. *ts-a*   *gospod-a=j/*s*   SIR
          DEM.DIST-F.SG   gentlemen(F)-SG=be.PRS.3SG/.3PL
          cʋro   *boɡʋt-a*
          very   rich-F.SG
          'those gentlemen are very rich'

Thus, in borrowing this lexeme, SIR selected one of the two options Croatian offered, viz. (31a). By contrast, NIR took the other option, (31b), as all the three above mentioned *a*-plurals, including *gospoda*, select plural agreement on verb forms, as in the model language, as shown in (32), while at the same time selecting an *a*-ending on other agreement targets which – as first remarked in Loporcaro (2018: 294f. n. 6) – is never found elsewhere in the language, where

the inherited paradigm of plural agreement targets in the relevant inflectional class(es), as seen in (20a), maximally features the binary contrast *bur-i* 'good-M.PL' vs. *bur-e* F.PL.[17]

(32) a. *č-a/*čeʎ-i*   *gospod-a=s/*j*   NIR
       DEM.DIST-N$_2$\-M.PL   gentlemen(N$_2$)-PL=be.PRS.3PL/.3SG
       jɒko   *bogɒt-a/bogɒ́ts*
       very   rich-N$_2$/rich\M.PL
       'those gentlemen are very rich'

   b. *č-a*   *gospod-a=z/*j*
       DEM.DIST-N$_2$ gentlemen(N$_2$)-PL=be.PRS.3PL/.3SG
       nægr-a/negr-i/*nægr-e
       black-N$_2$/black-M/F.PL/black-F.PL
       'those gentlemen are black'

   c. *č-a*   *pərv-a*   *di:ts-a*   *mi=z*
       DEM.DIST-N$_2$ first-N$_2$ child(N$_2$)-PL   1SG.DAT=be.PRS.3PL
       mai   *bur-a/?bur-i*
       more   good-N$_2$/good-M.PL
       'I prefer those former children'

   d. *bogɒt-a vlastel-a*   *ən*   *mostar av*   *fost grɒs-a*
       rich- N$_2$ noblemen(N$_2$)-PL   in   Mostar have.PRS.3PL   been fat-N$_2$
       'the rich noblemen in Mostar were fat'

   e. *čɒst-a*   *di:ts-a=z*   *mik-a/*mič*
       DEM.PROX-N$_2$ child(N$_2$)-PL=be.PRS.3PL   small-N$_2$/small.PL
       'these children are small'

This *a*-ending is exemplified for qualifying adjectives and demonstratives in (32), to which relative pronouns are added in (33).

(33) a. *čɒst-a*   *di:ts-a*   *kar-a/kar-ʎi*   *av*   *verít*
       DEM.PROX-N$_2$ child(N$_2$)-PL   REL-N$_2$/REL-M.PL   have.PRS.3PL   come
       ən   *sélište av*   *fost*   *visóč*
       in   village have.PRS.3PL   been   tall.PL
       'these children that came to the village are tall'

   b. *di:ts-a*   *kɒr-a/*kɒr-ʎi*   *am*   *vezút*
       child(N$_2$)-PL   REL- N$_2$/REL-M.PL   have.PRS.1SG   seen

---

17   Contrary to their Daco-Romanian counterparts, NIR *mik* 'small' and *negru* 'black' inflect differently, as the plural forms *mič* and *negri* are used for both masculine and feminine agreement (cf. Kovačec, 1998: 116, 126). However, for the latter adjective, while our informants indeed use *negri* for both genders, they also have a dedicated F.PL form *nægr-e*, which is ungrammatical with *a*-plurals as seen in (32b) but can occur e.g., in *čale do fœte=z negr-i/ nægr-e* 'those two girls(F) are black'.

'the children that I saw'

With these nouns, *a*-agreement is always acceptable while feminine plural -*e* never is (see (32b)). As far as masculine plural agreement is concerned, this is sometimes deemed fully acceptable (32b) and (33a), sometimes regarded as dubious (32c), sometimes excluded (33b). An Agreement-Hierarchy effect (Corbett, 1979), whereby M.PL agreement is the more acceptable the further away from the NP-internal attributive position, seems to be suggested by (33a), but (33b) (a judgement by the same informant) is not in line with this speculation. The crucial point for our reasoning is that verb agreement is always plural. Were this not the case, NIR would be like SIR (see (31)) or Sursilvan (see §8.1). But since these are undoubtedly plural nouns, and they select an agreement morph which never occurs with M.PL and F.PL nouns, these nouns must be specified for a distinct gender value, which is notated $N(EUTER)_2$ in the glosses, to distinguish it from the *o*-neuter seen in (20)–(23). This means that, taking the data in (32)–(33) into account, one needs to further complexify the gender system of NIR with respect to what available grammars have said so far. Our analysis is provisionally schematized in (34a).

(34) Gender/number agreement in IR

a. NIR

	SG	PL
M	*bur*	*bur-i*
F	*bur-a*	*bur-e*
N_1 (MN)	*bur-o*	–
N_2 (CN)	–	*bur-a*

b. SIR

	SG	PL
M	*bur*	*bur-i*
F	*bur-æ*	*bur-e*
N (MN)	*bur-o*	–

While the *o*-neuter$_1$ – which could be alternatively labelled M(ASS) N(EUTER) – is syntactically productive, as seen from the fact that it has taken on the default function, the *a*-neuter$_2$ – or, alternatively, C(OLLECTIVE) N(EUTER) – is not: rather, with just three borrowed nouns assigned to it, it is a vanishingly small gender value which, however, must be recognized as such. In particular, by Corbett's (2012: 84) criteria, one cannot call it an inquorate gender, since inquorate gender values are those "which comprise a small number of nouns, and whose agreements can be readily specified as an unusual combination of forms available for agreement with nouns with the normal gender values". The relevant cases reviewed in Corbett (1991: 170–175) are all instances of controller genders with no dedicated agreement targets, which – if the numbers are very small (one or two lexemes) – may be treated alternatively as lexical exceptions. Neither of these alternatives is available for the NIR neuter$_2$, since its agreement marker -*a* is a dedicated one, as no other word in the language selects it

in the plural. Thus, it matches the requirement put by Corbett (2012: 84, fn. 12): "If such nouns have their own unique agreement forms, rather than taking a combination of forms which are otherwise available, the agreement class must be recognized as a gender value, even if few nouns are involved."

As to the origin of this *a*-ending, it is clear – as Kovačec (1963: 35) remarks – that it is ultimately due to Croatian influence: if the two developments in SIR and NIR are independent from each other, in the latter this -*a* may have been extracted from phrases such as *bogata gospoda* 'rich gentlemen/lords', *draga dica* 'dear children' (for the mechanisms of direct vs. analogical borrowing of inflectional morphemes, cf. Gusmani, 1979; Gardani, 2008; 2012; 2018; 2020; Seifart, 2015).

The *a*-collectives assigned to the neuter$_2$ all share the property of not being determinable through numerical quantifiers (35a), a situation which is encountered sometimes, across languages, with collective nouns (cf. Loporcaro, 2018: 73f., for discussion of a parallel from Romansh; as (35b) shows, other quantifiers are not barred, and they regularly agree in -*a*).[18]

(35) a. **ura*   // **do/*   **doj/*   **dvoje/*   **dvoja*   *di:ts-a*
          one:*N$_2$//   two.F/   two.M/   two:FC/   two:*N$_2$ child(N$_2$)-PL
          (intended 'one child/one group of children//two children')

      b. *núškɒr-a*   *di:ts-a*
          some-N$_2$ child(N$_2$)-PL
          'some children'

For SIR, the schema in (34b) is not complete, since it displays the three target genders but omits the inherited alternating neuter (= AN in the gloss in (36a)), which has persisted longer in this IR branch (see discussion on (29)).

(36) a. *%do*   *kúvat-e/žɒ̌žet-e/lúpure*   SIR
          two.F   elbow(AN)-PL/finger(AN)-PL/wolf(AN)-PL

      b. *doj*   *kuvats/žɒ̌žets/lup*
          two.M   elbow(M)\PL/finger(M)\PL/wolf(M)
          'two elbows/fingers/wolves'

---

18   Among the numeral forms in (35a), *do, doj, dvoje* are used in other contexts in NIR, and therefore exist, as the reader by now knows, whereas **ura* and **dvoja*, to the best of our knowledge, do not (which is signalled by the asterisk included in the glosses "one:*N$_2$" and "two:*N$_2$". The latter forms have been built with the intention of exploring the theoretical possibility for speakers to create forms with the appropriate inflection for that feature-value combination, to be used with *a*-collectives, and, for *dvoja*, based on the homophonous NOM.N.PL form of the Croatian collective numeral (see (18c)).

As a matter of fact, at least one of our informants from Šušnjevica still variably allows F.PL, alongside innovative M.PL agreement (36b) – nowadays prevailing – with original Daco-Romance neuter nouns such as *kúvat* 'elbow' and *žρžet* 'finger', an option that, before it started to beat a retreat, had been extended even to original masculines such as *lup* 'wolf'.

To sum up, contact-induced complexification of the gender system seems to be on the rise in IR. The changes that led to the emergence of the two neuters in NIR are clearly contact-induced. Interestingly, two values of one and the same gender in the source language (neuter singular and neuter plural) have been copied at different times in the recipient language, not as part of one and the same paradigm but rather as two distinct defective gender values.

8   Borrowed Numerals with *PT* and the Further Complexification of the Gender System

Back to numerals, let us now consider the impact of the borrowing of Croatian *dvoje*, *troje* on the morphosyntactic system described in Sections 6–7. As shown in (37), this borrowing has turned a formerly binary option in the agreeing forms of the numeral 'two' into a three-way one, whereas all other agreement targets – exemplified in (37) with the demonstrative – only contrast two forms in the plural, in the paradigms usually given by grammars (see the demonstrative paradigm in Kovačec, 1971: 109, to which the NIR N_2 is added in (38)):

(37) a. *češc-i*   *doj/*do/*dvoje*   *omir/dints/kər/kúvete*   NIR
       this-M.PL   two.M/two.F/two.X   man/tooth/dog/elbow(M):PL
       'these two men/teeth/dogs/elbows'

   b. *čρst-e*   *do/*doj*   *surár/vρč/metle/fǽate*
      this-F.PL   two.F/two.M   sister/cow/broom/little_girl(F):PL
      'these two sisters/cows/brooms/little girls'

   c. *čρst-e*   *dvoje/*do/*doj*   *novin-e/postol-e/vil-e/škρr-e*
      this-F.PL   two.X/two.F/two.M   newspaper/shoe/pitchfork/scissors(F)-PL
      'these two newspapers/shoes/pitchforks/pairs of scissors'

(38) The proximal demonstrative in NIR:

		M	F	N_1	N_2
SG	NOM/ACC	*čest-a*	*čρst-a*	*čest-a*	

OBL		*čest-vǽ*	*česc-ǽ*	–
PL NOM/ACC		*česč-i*	*čɒst-e*		*čɒst-a*
OBL		*čest-orǽ*	*čest-orǽ*		–

The forms *dvoje* and *troje* are by now well integrated in the recipient system, so much so that, having been stripped away from the Croatian inflectional paradigm and having thus lost all the case/gender/number endings other than *-e*, they have developed oblique case forms by analogy with the nominal oblique endings of IR (compare e.g., *harta novinelor je raskinita* 'the paper of the journal is torn'):

(39)   a.   *vǻrx-ur-le*   *dvoje(r)lor*   *škɒr-e*
             tip(M)-PL-DEF.PL   two.X:OBL.PL   scissors(F)-PL
             'the tips of two pairs of scissors'
       b.   *roba*   *trojerlor*   *mudant-e*
             cloth(F)-SG   three.X:OBL.PL   underpant(F)-PL
             'the cloth of three pairs of underpants'

Synthetic oblique endings for nouns and pronouns were lost altogether in SIR and only preserved in NIR (Petrovici and Neiescu, 1965: 360). Among numerals, this is the case only in *ur* 'one', as shown in (13), while the others, including *do/doj* 'two', form the oblique case periphrastically preposing the case marker *a*: *a do/doj* 'two.OBL.F/M' (Kovačec, 1971: 117). Against this background, the morphological integration of *dvoje* and *troje* shown in (39) appears all the more remarkable.

    In what follows, we are going to explore the idea that also the borrowing of the numerals *dvoje* and *troje*, not unlike that of the *o*- and *a*-neuter agreement markers considered in Sections 7–8, may have increased the complexity of the recipient morphosyntactic system. We argue namely that this borrowing resulted in introducing gender overdifferentiation into the paradigms of the two agreement targets at issue. In other words, we propose that the three-way contrast seen in (37) has to be treated as one of (sub)gender.

8.1   *Comparative Evidence: Gender Overdifferentiation on Lower Numerals in Romance*

The typological presupposition of what we have been saying so far is that "[c]ardinal numbers sometimes show agreement; typically, this is restricted to lower numerals" (Corbett, 2006: 42). Lower numerals, cross-linguistically, have also not rarely been reported to host gender overdifferentiation, defined as follows:

    for targets to be considered overdifferentiated, a specific gender agree-

ment distinction must be restricted to a particular word-class, and even within this word-class it must be restricted to certain lexical items. (COR-BETT, 1991: 169)

Corbett (1991: 168f.) cites examples of overdifferentiation on the numerals 'two', 'three' and 'four' in Kolami-Naiki and Parji-Ollari, two central Dravidian languages in which only those numerals display dedicated agreement forms for female human nouns, in addition to those occurring on all other agreement targets, which contrast only male human vs. other. In Romance, a comparable state of affairs is observed in Romansh, as shown with examples from Sursilvan and Engadinian in (40) (data from Candinas, 1982: 110f.; Spescha, 1989: 312f., on Sursilvan; Ganzoni, 1977: 56f., on Upper, and Ganzoni, 1983: 56f. on Lower Engadinian):

(40) a. Srs.   b. Eng.   gloss
        du-s/trei-s mattatschs   *du-os/trai-s mats*   'two-M/three-M boys(M)'
        du-as/trei-s mattatschas   *du-os/trai-s mattas* 'two-F/three-F girls(F)'
        du-a/trai-a pêra   *du-a/trei pèra*   'two-N/three:N pairs(?)'

In addition to masculine and feminine, generally contrasted on plural agreement targets from all relevant classes, the numerals 'two' and 'three' feature a distinct form ending in *-a* – a diachronic successor of the Latin neuter plural agreement morph *-a* – which nowadays only occurs within complex numerals such as Eng. *duatchient* '200', *traiamilli* '3000' and the periphrastic quantifiers in (40).[19] However, until not long ago these forms could modify *a*-collectives like *bratsch-a* 'arms(F)-SG' even in their literal meaning.

(41)   *Tgi che ha duas combas e dua bratscha duei gie buca selubir da simular e mulestar il miedi* (Candinas, 2009: 91)
        'that who has two legs and two arms should not dare to simulate and disturb the doctor'

   The author of the novel from which the passage is drawn, Theo Candinas, was born in Surrein-Sumvitg, Surselva, in 1929; for younger speakers, *dua bratscha*, if at all acceptable, can only denote a measure, meaning 'two ells' (see Kämpf, 2015).

---

19   While *duo* was the Classical Latin neuter form, an analogical variant *dua*, with the nominative/accusative ending reshaped on the model of nominal inflection, is also attested: see *ThLL,* 5(1): 2241f.

Exactly the same overdifferentiation on the numerals 'two' and 'three' occurred in medieval Northern Italo-Romance (in Veneto, Lombardy, Emilia and Liguria: see Loporcaro and Tomasin, 2016) where these were the only plural agreement targets to feature a three-way gender distinction:

(42) Gender overdifferentiation on 'two' and 'three' in medieval Northern Italo-Romance

	M	F	N	
PL	*dui, tri*	*doe, tre(i)*	*doa, trea*	'two, three'
	li	*le*	–	DEF ART (= all other AGR targets)

Here too, the *a*-forms could not modify normal count nouns but were restricted to use within periphrastic quantifiers ('two/three pairs of *x*') and complex numerals: (page numbers are given in brackets):

(43) 14th century Venetian (*Tristano corsiniano*, edn. Tagliani, 2011)
   M   *li dui fradelli* (73) 'the two:M brothers(M)'; *li altri dui conpagnon* (89) 'the other two:M companions(M)'; *delli dui servi* (99) 'of the two:M servants(M)';
   F   *tute doe le palme* (108) 'both [lit. 'all:F two:F'] palms(F.SG)'; *a doe mane* (108) 'with both [lit. 'two:F'] hands(F)'; *doe çornade* (144) 'two:F days(F)';
   N   *doa para d'arme* (124) 'two:N pairs of arms/armours'; *plu de doa milia* (145) 'more than two:N thousand'; *doa tanti* (127) 'twice [lit. 'two:N'] as many'.

Of course, the data discussed in (40)–(43) differ from those from NIR in several respects. On the one hand, diachronically, overdifferentiated forms are inherited in Romansh, as they were in medieval Northern Italo-Romance, being a leftover of the Latin three-gender system which elsewhere shrunk to a binary contrast; in NIR, on the contrary, they arose from language contact. Synchronically, on the other hand, those seen in (40)–(43) are plain three-way contrasts, whereas in NIR the situation is, also in this respect, more complex.

However, there is also a striking similarity. While the data in (41) still bear witness to the original plurality of the *a*-noun forms selecting *dua* and *trei/ traia*, such noun forms in modern Romansh belong to number-defective paradigms with a form/meaning mismatch: Sursilvan *bratscha* denotes two entities but is morphosyntactically singular, a mirror image with respect to NIR PT nouns selecting *dvoje/troje* such as *novine* 'newspaper'.

8.2   *Contact-Induced Gender Overdifferentiation for Lower Numerals in*
         NIR

The scheme in (44) displays the usual situation for gender/number marking on (non-overdifferentiated) agreement targets exemplified with the paradigm of *ur* 'one, some':

(44)

	M	F	N_1	N_2	
SG	*ur*	*ur-a*	*ur-o*		'one'
PL	*ur-i*	*ur-e*		–	'one, some'

In addition to the contrasts seen in (44) – two number and three gender values (no plural **ur-a* occurs, as the neuter$_2$ never occurs with numeral quantifiers, see (35a) and fn. 18) – the schema in (45) adds complexity in the form of a layering in the feminine (here, also the N_1 does not occur, since the agreement targets at issue are plural while the N_1 only occurs in the singular):

(45)

	M		F	N_1	N_2	
			collective			
PL	*doj*	*do*	*dvoje*		–	'two'
	trej		*troje*		–	'three'

We know independently (see (17) and (37c)) that PT nouns which select *dvoje/troje* are feminine and plural, and that they share this feature specification with ordinary count feminines that select inherited *do* 'two.F' instead (37b). Thus, they all share the same gender/number specification, so that our hypothesis is that overdifferentiation in lower numerals signals what has come to be a sub-gender contrast in NIR. In (45), the subgender signalled by selection of *dvoje*, *troje* is labelled 'collective' in a merely conventional way: while this alludes to the origin of the borrowed agreeing numerals, it does not imply retention of the original semantics of collective numerals, a point to be dealt with in Section 8.3.

Synchronically, we argue that borrowed *dvoje* and *troje* are now distinct word forms in one and the same paradigm together with the inherited forms of the numerals 'two' and 'three' (the non-greyed-out cells in (45)): in other words, though differing in origin, native *doj/do* and borrowed *dvoje* have become part of one and the same numeral lexeme, and the same goes for native *trej* and borrowed *troje*.

At this point, a series of questions arise, whose discussion will require considering additional comparative evidence from Romance and beyond: a) firstly, and crucially, the question whether this idea is on the right track, considering that no such morphosyntactic analysis has been proposed yet, to the best of our

knowledge, for the many languages in which PT nouns select special numerals; b) secondly, the issue whether, in case overdifferentiation is assumed, this is best analysed in terms of (sub)gender, or whether it should rather be treated in terms of some other morphosyntactic feature; c) thirdly and finally, whether – assuming the (sub)gender analysis is correct – the gender-asymmetry seen in (45) is justified, or whether such overdifferentiation should rather be assumed also for the masculine. We will start by discussing the last issue in Section 8.3, since the data introduced there will pave the way for addressing the fundamental issue (a) in Section 8.4, where quantification with Latin PT nouns will be drafted in as a useful comparison. Finally, in Section 8.5 we will show that the analysis in terms of (sub)gender is preferable over conceivable alternatives appealing to other morphosyntactic features.

8.3   *Lack of Overdifferentiation in the Masculine and the Semantics of Dvoji and Dvoje*

Kovačec's (1998) dictionary contains a handful of masculine nouns, whose lexical entries are given in the plural and may consequently stand as candidates for PT status. These all reported in (46):

(46)   Masculine plural entries in Kovačec's (1998) dictionary
[legend: B = Brdo (SIR), J = SIR (all villages), S = Šušnjevica (SIR), Ž = Žejane (NIR), It[alian], Ven[etan]:
armi (J) 'weapons' (It. *armi*), *bizi* (S) 'peas' (Ven. *bizi*), *boşe -le* (Ž) 'testicles',[20] *cârmel' -i* (J Ž) 'pebbles of sleep dust' (Cr. *krmelji*), *cvadri* (B) 'sacred images' (It. *quadri*)

As is readily apparent, most of them are not used in Žejanski but only occur in SIR, so that only NIR *boşe* and *cârmel'* are potentially relevant to our question. We have tested them, asking our informants whether they could be quantified with M.PL *dvoji* (see the possible Croatian source *dvoji* in (18a)), with the following results:

(47)   *doj/*dvoji boše*
'two testicles'

(48)   a.   *doj kərmeʎ*
'two pebbles of sleep dust'
         b.   *%dvoji kərmeʎ*

---

20   Note that the lexeme *boşe, -le* (Ž) has a plural entry in Kovačec's (1998: 40) dictionary. However, the author also cites sg. *ən boš* 'a testicle'.

'two pebbles of sleep dust'

Most speakers reject the sentences with *dvoji* in (47) and (48b) as ungrammatical. For two of our informants, however, *dvoji kərmeʎ* (48b) is acceptable, although only if the objects belong to two different sets of pebbles of sleep dust, e.g., one/two from one eye, one/two from the other. The remaining speakers reject it outright. The crucial point for us is the fact that *dvoji* with these nouns is not selected categorically as the only grammatical form of the numeral 'two', contrary to what is observed with the feminine PT nouns in (14) and (16), nor do any other masculine PT nouns seem to exist for which this would be the case. This guarantees that (45) is correct in not positing any sub-gender contrasts for the masculine: in other words, the contrast between F *do* and *dvoje* in (15) vs. (14) is relevant to the morphosyntax, while the difference between M *doj* and *dvoji* (for the NIR speakers who deem the latter form grammatical, in (48b)) never is.

Indeed, also the *do ≠ dvoje* contrast in the feminine may convey, with non-PT nouns, a purely semantic contrast not relevant to the morphosyntax like the one seen in (49). In fact, while feminine PT nouns select the numerals *dvoje, troje* categorically, the latter are not restricted to quantification of PT nouns, but can also quantify countable plurals, exemplified with *ženska* 'woman' and *šalitsa* 'cup' in (49b) and (50b):

(49) a. *trej ženske* 'three women' (unmarked)
     b. *troje ženske* 'three women' (= three different [sets of] women)

(50) a. *vezút=ɒm do/trej šalitse* 'I've seen two/three cups'
                                    (unmarked)
     b. *vezút=ɒm dvoje/troje šalitse* 'I've seen two/three cups'
                                    (of different kinds)

When this happens, these expressions, contrary to those with cardinal numerals in (49a), (50a), indicate that what is being referred to is either two/three sets (for some speakers) or two/three items only if picked from distinctly different sets (for others):[21] for speakers of the former group, *dvoje šalitse* means 'two sets

---

21 For instance, (49b) may indicate – for speakers of the former group – that the three women at issue are instances of different types e.g., in that they come from the set of red-haired, black-haired, and blond women.

of cups', independently of the number of items in each group. The same usage of collective numerals is observed with masculine count nouns too:

(51) a. *doj kɒʎ*   'two horses' (unmarked)
      b. *dvoji kɒʎ*   'two different [groups of] horses'

(52) a. *trej kǝr*   'three dogs' (unmarked)
      b. *troji kǝr*   'three different [groups of] dogs'

In the light of this, (48b) does not seem to instance the kind of morphosyntactically obligatory use the borrowed numerals have been put to in NIR, described in (14), (16) and (37c). Rather, it seems to be interpretable as a manifestation of the same optional collective use found in the source language: the possibility to convey such 'group' meaning is part of the rich semantics of Slavic collective numerals (see Leko, 2009: 76–81 and Stefanović, 2011 for BCS).

8.4   *A Flashback: Collective Numerals and PT Nouns in Latin*

The last observation gives us the opportunity for a brief comparative discussion: in fact, the occurrence of collective numerals, both semantically contrastive and morphosyntactically selected (at least apparently, e.g., with PT nouns) is not limited to Slavic but occurs in other branches of Indo-European, including Latin (see the comparative study by Brugmann, 1907: 49), as well as in other language families: Ojeda (1997: 161–166) reviews relevant data from Finnish, Mongolian and Greenlandic.

For Latin, we have mentioned in (4) the occurrence of the plural form of the numeral *ūnus* 'one' with PT nouns. For numerals from '2' on, alongside cardinal numerals, Latin had inherited from PIE a series of collective numerals: *bīnī* '2', *trinī* '3', *quaternī* '4', *quīnī* '5', etc. Latin grammars report that these are selected with PT nouns, and this usage is widely documented in Latin texts.

(53) *Inter*   *bin-a*   *castr-a*   *Pompei*   *atque Caesaris*
      between   two-   camp(N)-ACC. Pompey's   and   Caesar's
                   ACC.N.PL PL

      unum   *flumen*   *tantum intererat*   *Apsus*   (CAES., *B.G.* 3.19)
      one   river   only   inter-   Apsus
                                         vene:IPF-3SG

      'the two camps, Pompey's and Caesar's, were separated only by the river Apsus'

That this selection may have been obligatory seems to be suggested by passages by ancient grammarians such as the following.

(54) non dicimus biga una, quadrigae duae, nuptiae tres, sed pro eo unae
      bigae, binae quadrigae, trinae nuptiae (VARR., *L.L.* 10.3.30)
      'we do not say **biga una*, **quadrigae duae*, **nuptiae tres* [with cardinal numerals] but instead say *unae bigae* "one two-horse team", *binae quadrigae* "two pair of four-horse teams", *trinae nuptiae* "three sets of nuptials" [with distributive numerals]' (translation after Taylor, 1974: 96)

In mentioning the selection of *unae, binae, trinae* in (54), Varro voices grammatical prescription, recurring in the grammars from the antiquity. Slightly different statements are met with in Flavius Caper, 2nd century AD (Keil, 1856–1880: 7.108.7f.): "*binas tabulas dicimus, non duas*" 'we say *binas tabulas* 'two writing tablets', not *duas*'; or Priscian, the most influential grammarian of Late antiquity (see Keil, 1856–1880: 3.414.25). But whether Varro's and his fellow grammarians' "*non dicimus*" can be taken as grammaticality judgements is dubious, in view of the fact that cardinal numerals are also attested with the same PT nouns (55), and even reported in the context of a metalinguistic observation by another grammarian, as is the case in Servius' commentary on Vergil in (56).

(55) a. *castra duo praesidiaria Barzalo et Claudias peterentur, sese ductante*
          'that they should go under his guidance to the two garrison camps
          of Barzalo and Claudiae' (AMMIANUS MARCELLINUS, *Res gestae*
          18.7.10, translation after Rolfe, 1982: 1.455)
      b. *exinde duabus admotis quadrigis in currus earum distentum inligat*
          Mettium
          'thereupon, two chariots were brought up, and he [Tullus
          Hostilius] orders that Mettius be stretched out between them' (LIV.
          1.28.10)

(56) *frenaque bina poetice, nam 'duo' debuit dicere: 'bina' enim secundum*
      Ciceronem non dicuntur nisi de his quae sunt numeri tantum pluralis.
      nam Cicero per epistolam culpat filium, dicens male eum dixisse 'direxi
      litteras duas', cum 'litterae', quotiens epistolam significant, numeri tan-
      tum pluralis sint. contra 'epistolas binas' non dicimus, sed 'duas' (SERV.
      A. 8.168)

> '*frenaque bina* 'two brakes' is used in poetry, indeed he should have
> said *duo*: in fact *bina* according to Cicero can be used exclusively with
> PT nouns, since Cicero in a letter finds fault with his son, telling him
> that he made a mistake in saying *direxi litteras duas* 'I sent two let-
> ters', because *litterae*, whenever it means 'letter', is a PT noun. On the
> contrary, we do not say *epistolas binas, but rather *epistolas duas* 'two
> letters" (compare *litteras duas*/*binas 'two letters of the alphabet')

Ammianus was a native speaker of Greek, born in Antioch in 330 AD, who
learned Latin as L2 (Rolfe, 1982: 1.xx), but this was not the case for Livy nor
Cicero's son, Marcus Minor, whom his father rebuked according to Servius' pas-
sage for saying, "incorrectly", *litteras duas*.[22] Based on this evidence, Löfstedt
(1958: 101) argues that the use of collective numerals (which he labels 'distrib-
utive' following a tradition that goes back to the ancient grammarians: *disper-
titiva* 'distributives' in Priscian, *De figuris numerorum*, ed. Keil, 1858: 3.413.24)
with those nouns was determined by the semantics, and hence did not really
differ from the occurrence of the same collective numerals with count nouns to
count "Einheiten, deren jede in sich ein Mehrfaches ist" ['units, each of which
is per se a multiplicity'] (Löfstedt, 1958: 100). This latter use with count nouns is
exemplified in the following examples (discussed in Ojeda, 1997: 146f.):[23]

(57)   a.   *molas asinarias unas et trusatilis unas Hispaniensis unas* (CATO,
            agr. 10.4)

---

22   These examples have been discussed in many studies, from Brugmann (1907: 49 n. 1), who
     recognizes that the grammarians' rule did not (any longer) mirror actual usage in Classical
     Latin, to Ojeda (1997: 154).

23   This emerges from Löfstedt's (1958: 101) account of the occurrence of cardinal numbers in
     (55)–(56): "Die Verwendung von Kard. für Distr. in solchen Fällen erklärt sich wenigstens
     zum Teil dadurch, dass man das Gefühl verloren hatte, dass es sich um pluralische Einheiten
     handele; *litterae* war nicht mehr eine Gruppe von Buchstaben, sondern ein Brief, eine
     epistula." ['The use of cardinal instead of distributive [i.e., collective] numerals is at least in
     part explained by the fact that one had lost the sense that these were *plural* units: *litterae*
     was no longer a group of letters, but a letter, an *epistula*.']

24   A comparable optionality is reported by Stefanović (2011) for contemporary BCS usage, as
     mentioned in Section 5 while commenting on (18)–(19). Other Slavic branches show a rather
     intricate situation. In Russian, a few PT nouns still select collective numerals categorically:
     e.g., *dvoe časóv* 'two watches' is the only grammatical way to quantify the PT noun *časý*
     'watches' with a one-word numeral expression, while the cardinal numerals *dva* 'two.M/
     N'/*dve* 'two.F' are barred. Of course, paraphrase with a periphrastic classifier is always a
     viable alternative, which indeed seems the favourite one for several of the subjects tested by
     Nikunlassi (2000: 235–241).

‘one pair of donkey mills, one pair of hand mills, one pair of Spanish mills’

b. *habetis interim bina animalia* (APUL., *De deo Socratis* 4)
   ‘you have in the meantime two kinds of creatures’

c. *boves trinos* (CATO, *agr.* 10.1)
   ‘three yoke of oxen’

In conclusion, a difference between (4) and (53) emerges: with PT such as *castra*, the plural form of *ūnus* was mandatory, while **unum castra* would have been ungrammatical, whereas the selection of collective *bīna*, *trina* (instead of *duo, tria*) with nouns of the same kind was optional.[24]

   This comparison corroborates the conclusion that the NIR replica numerals *dvoje, troje* selected categorically with the feminine PT nouns in (14), have unique properties. Their contact source is collective numerals whose semantics is still visible in NIR in their marginal use with count nouns exemplified in (48)–(52). However, categorical selection in, say, *dvoje/troje novine* ‘two/three newspapers’ is dictated by the morphosyntax, not by the semantics. In other words, these borrowed forms have become fully integrated in the NIR lower numeral lexemes filling a morphosyntactically defined paradigm cell, as shown in (45).

8.5   *Complicating Gender or Number? Comparative Evidence from Romance and Beyond*

When analysing rather intricate systems, ascribing a given contrast to one or the other morphosyntactic feature may prove a non-trivial issue. For instance, in his discussion of PT nouns Corbett (2019: 54f.) mentions Cicipu, a Benue-Congo language spoken in northwest Nigeria, in which there is just one PT, the noun *à-húlá* ‘name’, which “has a plural form, plural agreements, and this is so whether it denotes one name or more than one”; he adds in a footnote: “McGill (2009: 253) treats this noun as belonging to an inquorate gender, but I believe it should be seen as a number problem (it lacks a singular form) rather than a gender problem.” Similar problems present themselves also in Romance, and briefly addressing some of this evidence will help consolidating our analysis of NIR.

8.5.1   A Controversial Case: Asturian *o*-Agreement as a Value of Gender or Number

A case in point from Romance is that of (Central) Asturian, where all prenominal modifiers, exemplified in (58) with the definite article, mark the usual binary contrast (as in Spanish or Italian), while other agreement targets not preceding the noun within the NP signal a three-way distinction (data from the

Central Asturian dialect of Lena; see Neira Martínez, 1955: 70–72; 1978: 260; the standardized variety of Asturian displays the same behaviour):

(58)

gender	countness	det	N	Adj	Central Asturian (Lena)
F	count	*la*	*casa*	*fria*	'DEF.F.SG cold house'
	mass		*tsiche*	*frio*	'DEF.F.SG cold milk'
M	mass	*el*	*café*	*frio*	'DEF.M.SG cold coffee'
	count		*pie*	*friu*	'DEF.M.SG cold foot'

This three-way distinction has been dubbed one of subnumber by Corbett (2000: 126), who proposes that the singular subdivides into mass and singular in a second number system:

(59)

Top system (number)	singular		plural	
Second system (subnumber)	mass	singular	plural	
(exemplified with masculine)	*fri-o*	*fri-u*	*fri-os*	'cold'

The alternative analysis proposed in Loporcaro (2018: 172–179), on the contrary, regards the binary contrast seen in (58) on definite articles and the three-way one seen on postnominal adjectives as manifestations of two concurrent gender systems, along the lines of the cross-linguistic study by Fedden and Corbett (2017).

(60) Alternative analysis (Loporcaro 2018: 172–192): two concurrent gender systems

SYSTEM 1: TWO VALUES

MASC.	*el pie/café*
	DEF.M.SG
	foot(M,MC)/coffee(M,N)
FEM.	*l-a*	*casa/tsiche*
	DEF-F.SG	house(F,FC)/ milk(F,N)

DOMAIN: [__ N]ₙₚ

SYSTEM 2: THREE VALUES

MASC.		
COUNT	*pie* foot(M,MC)	*fri-u* cold-MC
NEUTER	*tsiche/café* milk(F,N)/ coffee(M,N)	*fri-o* cold-N
FEM.		
COUNT	*casa* house(F,FC)	*fri-a* cold-FC

DOMAIN: elsewhere

In our NIR case, an analysis differing from the one put forward here in terms of (sub)gender and positing a number contrast instead, seems to be less likely, given the overall morphosyntactic structure of a Romance language like IR as well as the general properties of grammatical number. The values of the number feature are defined semantically in terms of the numerosity of real world entities.

> These values of the number feature have meanings and forms associated with them. The main part of the meaning of the singular is that it refers to *one* real world entity, while the plural refers to *more than one* distinct real world entity. [emphasis added, *M.L. et al.*] (CORBETT, 2000: 4)

In the data in (37), the number value of all the contrasting items *doj/do/dvoje* and *trej/troje* is identical in terms of real world entities: the quantified NPs *dvoje novine/vile* 'two newspapers/pitchforks', *do mɘre/ženske* 'two hands/women', and *doj dints/omir* 'two teeth/men' all denote exactly two real world entities, and the same identity goes for *trej/troje*, so that there seems to be no cogent semantic/referential reason to postulate any contrast among them, as to this category. In Romance, where the number contrast is binary (singular vs. plural), quantified phrases containing 'two' and 'three' are all equally non-singular, i.e., plural. Alternatively, such a reason could be provided by the morphosyntactic system, as is the case in languages such as Finnish.

8.5.2   A Different Case: Number Contrasts in Numerals in Finnish

Finnish shows "an unusual interaction between numerals and nouns", thoroughly discussed in Hurford (2003: 584–589; quote from p. 584). In this language, all numerals have both singular and plural forms, the latter used to indicate sets of objects (contrast (61b) with the singular forms (61a)) and also selected obligatorily with PT nouns (61c) (Hurford, 2003: 587):[25]

(61) a. *yksi*   *kenkä*   / *kaksi*   *sukaa*

---

25 In Finnish, plural numerals agree with head nouns in all cases. In the singular, this happens with *yksi* 'one', while other formally singular numerals govern a noun in the partitive singular, whenever the relevant NP receives nominative or accusative case, the only two cases occurring in (61)–(62) (Hurford, 2003: 585). In NPs which receive any of the remaining cases, case-agreement is observed.

one:NOM.SG   shoe:NOM.SG   two:NOM.SG sock:PRTV.SG
'one shoe'   / 'two socks'
b. *yhdet*   *kengät*   / *kahdet*   *sukat*
one:NOM.PL   shoe:NOM.PL   two:NOM.PL sock:NOM.PL
'one group (typically a pair) of shoes' / 'two groups (pairs) of socks'
c. *kahdet*   *sakset*   / *kahdet*   *kasvot*
two:NOM.PL   scissors:NOM.PL   two:NOM.PL face:NOM.PL
'two pairs of scissors'   / 'two faces'

This is interesting in many respects, for our discussion. One reason is that, for NPs in which plural numerals modify count nouns, Hurford (2003: 588) describes diverging judgements among his informants, in a way somewhat reminiscent of the variation in interpretation discussed in (49)–(52) while commenting on what we have labelled the semantic use of borrowed *dvoji*, *dvoje* in NIR:

(62) a.   *oppilaat*   *saivat*   *kolme*   *kirjaa*
            pupils   got   three:ACC.SG   book:PRTV.SG
            'pupils received three books'
      b.   *oppilaat*   *saivat*   *kolmet*   *kirjat*
            pupils   got   three:ACC.PL   book:ACC.PL
            'pupils received piles of three books'/'three groups of pupils
            received (some) books'

Sentence (62a), containing a singular numeral, is systematically ambiguous for all informants – just as its English translation equivalent – between a reading where the quantified NP has wide scope ("there is a set of just three books which the pupils, as a group, receive") and a distributive reading where *oppilaat* has scope over *kolme kirjaa* ("each individual pupil receives a set of three books"). Hurford's (2003: 588) informants part ways when it comes to interpreting (62b), where, the plural numeral induces different interpretations: for one informant, "each pupil receives copies of the same three books as the other pupils", while for another the reading is that "a teacher has three variously sized groups of pupils and gives each group of pupils one pile of books; we don't know how many books are in each pile, but there are exactly three piles". As Hurford (2003: 588) puts it, "What is common to the interpretations suggested by both informants is the idea of three sets (alias types, piles) of books."

This very variation shows that number contrasts in numerals, though well-entrenched in the morphology and morphosyntax of Finnish, fall in a grey zone: while the unmarked option has an unambiguous meaning, the other one (here,

the plural) is trickier. With count nouns, there is vacillation in interpretation and, in addition, Hurford (2003: 587) reports judgements by speakers who deem plural numerals awkward in this or that context. With PT nouns, by contrast, the use of plural numerals (to the exclusion of singular ones) is described as categorical and unproblematic, and this is generally the case in Finnish grammars (cf. e.g., Whitney, 1956: 173). Thus, the Finnish evidence shows that a difference in number is an option, cross-linguistically, for numeral quantification with PT vs. plain count nouns. Finnish is well equipped for this, as its numerals are "declined in the same way as nouns" (Whitney, 1956: 171). By contrast, IE languages such as Latin and the Slavic languages take this option only for the numeral 'one', and even for this, many Romance languages – with the exceptions seen in (5), (8d) and (9) – have to resort to the classifier strategy instead. On the whole, thus, the Romance languages differ from Finnish in that they do not feature a declensional paradigm of numerals in which a regular number contrast can be hosted. Consequently, the distinction introduced by *dvoje* and *troje*, contrasting respectively with *doj/do* and *trej*, is doomed to remain an isolated irregularity, which is indeed what overdifferentiation means. When it comes to labelling the morphosyntactic feature involved, gender seems the natural choice in terms of system-adequacy, given the non-availability of number (contrary to Finnish) and given comparable cases of gender overdifferentiation on lower numerals in Romance (see Sections 8.1, 8.5.3).

The occurrence of minor number values, with restricted range in the lexicon of some languages (see Corbett, 2000: 89–110), might be described as a kind of pendant to gender overdifferentiation: for instance, in Arapesh (Papua New Guinea) "pronouns and nouns typically distinguish singular and plural […]. But just the first person pronoun has singular versus dual versus plural" (Corbett, 2000: 91). Corbett's cross-linguistic review of minor numbers does not include any examples from the Romance languages.[26]

---

26 Another way of treating systems where number does not behave uniformly across word classes is the distinction of a top and a second system (as shown in (59)), which can coexist with distinctions in range. Corbett (2000: 92f., 120f.) illustrates this point with Yimas (Papua New Guinea), in which both nouns and pronouns contrast singular, dual and plural, while only personal pronouns contrast paucal in addition. The additional contrast for this minor number value defines at the same time the top number system, covering personal pronouns, while the second system covers nouns.

27 Note that *kavətso:nə* in (63b) is a count noun and occurs there in the singular, just as the word *parrottsə* 'black bread' in (64): the corresponding plural(s) would have selected the plural form of the demonstrative, viz. *kwidd-i.*

8.5.3   A Bipartite Gender Value (for One Class of Targets) in Northern
               Apulia
Thus far, we have argued that the contrast between *dvoje* and *do* must be
accounted for in the morphosyntax rather than being a matter of mere seman-
tics (Sections 8.3–8.4), and that an account in terms of (sub)gender seems pref-
erable over one in terms of number (Sections 8.1–8.2; 8.5.1–8.5.2). As a final
piece of comparative evidence in support of our analyisis, we will now show
that there are indeed comparable cases of Romance varieties in which just one
gender value is subdivided in two subgenders, contrasted on just one overdif-
ferentiated agreement target. One such variety, the Northern Apulian dialect of
Sannicandro Garganico (province of Foggia), is discussed in Loporcaro (2018:
289–291), based on data from Carosella (2005: 89) and Gioiosa (2000: 91–95).
In Sannicandrese, only one class of targets, demonstratives, is sensitive to a
[±human] contrast, and this sensitivity is restricted to the masculine (63a-b),
one of the two gender values normally contrasted in the dialect, which shows
elsewhere (on articles, adjectives, participles etc.) a plain binary contrast:[27]

(63) a. *kwidd-u*   *krəstja:nə*   Sannicandro Garganico
          DEM.DIST\M_HUM-M.SG   man(M_HUM) (province of Foggia, Italy)
          'that man/person'
      b. *kwedd-u*   *kavətso:nə*
          DEM.DIST\NONM_HUM-M.SG trousers(M_NONHUM)
          'that pair of trousers'
      c. *kwedd-a*   *kumma:ra/vakka/kami:ša*
          DEM.DIST\NONM_HUM-F.SG   godmother/cow/shirt(F)
          'that godmother/cow/shirt'

More precisely, as specified in the glosses in (63), affixal inflection encodes the
same binary masculine vs. feminine contrast found elsewhere, and it is only
the combination of affixes with the allomorphs of the demonstrative stem that
marks the subgender contrast: the allomorphs *kwidd-* (distal), *kwiss-* (interme-
diate), and *kwist-* (proximal, exemplified in (64)) occur with [masculine, singu-
lar, human] head nouns, while the complementary allomorphs *kwedd-, kwess-*,
and *kwest-* occur elsewhere, including with [masculine, singular, non-human]
head nouns – as shown in (64), Sannicandrese has a convergent system
(Corbett, 1991: 155) neutralizing gender in the plural:

(64)	SINGULAR	PLURAL	gloss

M	HUM	*u/kwist-u kumba:rə*	*i/kwist-i ǧǧú:vənə*	'DEF/DEM_PROX godfather/ youngster, -s'
	NONHUM	*u/kwest-u parrottsə*	*i/kwist-i kavaddə*	'DEF/DEM_PROX black bread/ horse, -s'
F		*la/kwest-a kami:ša*	*i/kwist-i kami:ša*	'DEF/DEM_PROX shirt, -s'

This parallel supports the analysis proposed for NIR lower numerals in (45), in that it shows that overdifferentiation within just one gender value on just one agreement target may arise anew, in a Romance variety.

9   Conclusion

The NIR case departs from the other cases of gender overdifferentiation in Romance discussed thus far, because neither in Romansh and medieval Northern Italo-Romance (40)–(43), nor in the Northern Apulian dialect mentioned in (63)–(64), was this overdifferentiation induced by contact. The two cases considered for comparison differ from each other, in turn, in that in Romansh both form and function (of e.g., Sursilvan *dus*, *duas*, and *dua*) are inherited (though the functional domain of *dua* has shrunk massively), whereas in Sannicandrese the forms are inherited but the functions have been reshuffled, since *kwist-u* vs. *kwest-u*, nowadays both masculine contrasting as [+human] vs. [–human], must be traced back to Late Latin masculine *ECCUM-ISTUM vs. neuter *ECCUM-ISTOC, i.e., to a gender contrast, not one of subgender.

In NIR, overdifferentiation in lower numerals arose via borrowing of *dvoje* and *troje* as a net increase in complexity (number of contrasts), thus adding to the not too many cases reported so far of contact-induced morphosyntactic complexification. On the whole, the NIR system has become more complex through contact in several ways, all involving borrowing from Croatian of agreement targets which had different functions in the source language. The symmetrically defective values of the two neuters (N_1 and N_2) both derive from one and the same non-defective gender value of Croatian, the neuter. The

---

28   Once the latter was established, also *ure* can be viewed as a form filling the now available collective F.PL subgender cell.

overdifferentiation on 'two' and 'three', by contrast, arose capitalizing on borrowed numeral forms which, in the source system, contrasted in lexical/semantic terms with non-collective numerals but, once borrowed, entered one and the same lexeme paradigm with the Daco-Romance inherited numerals *doj/do* and *trej* respectively. This borrowing may have started as whole Croatian NPs headed by PT nouns and consequently containing collective numeral forms came to be used in NIR discourse, much like in the case of other numerically quantified borrowed NPs considered in Section 4. Also, this borrowing process cumulated onto another, also contact-induced, distinctive property of IR, viz. the availability of the F.PL form of the numeral *ur/ura* 'one:M/F' for quantification of PT nouns, seen in (5). This was probably a calque on Slavic, shared by NIR and SIR, which however did not in itself impact on gender since the F.PL form *ure*, selected with PT nouns, contrasted with *ur/ura* in number. By contrast, as *dvoje* and *troje* became novel forms in the paradigm of the numerals 'two' and 'three', adding to inherited *doj/do* and *trej*, a contrast in number was not an option, since all these forms are uniformly plural. This resulted in the subgender contrast we have described.[28]

To sum up, the result of our analysis of NIR can be schematized as in (65), where the class one adjective *bur* 'good' illustrates the core grammatical system, originally consisting of the four inherited cells occupied by *bur, -a, -i, -e*. In addition, the paradigm of agreement targets such as *bur* has been enriched with the N_1 (*buro*), which found its way into the gender system (of both branches of IR), in spite of its scantiness in terms of controller lexemes, because of its syntactic function as the default agreement marker. At a later – and quite recent – stage, only in NIR the N_2 (*bura*) has arisen: this completes the set of agreement options available in today's NIR for all class one adjectives, articles, personal pronouns and demonstratives. In addition, the paradigms of the two numerals 'two' and 'three' show the further complexification of the gender system in this Romance variety.

(65)

	M	F		N_1	N_2	
		collective				
SG	*bur*	*bura*		*buro*		'good'
	buri	*bure*			*bura*	
PL	*doj*	*do*	*dvoje*		–	'two'
	trej		*troje*		–	'three'

As we have argued, borrowing of *dvoje* and *troje* from Croatian, now selected categorically in NIR with a handful of feminine PT nouns, has enriched the

paradigm of the two numeral lexemes at issue, but also affected the morpho-syntactic system, yielding (sub)gender overdifferentiation within the feminine. This was the rather unexpected conclusion our analysis brought us to, considering that the original purpose of our fieldwork in Istria was an inspection of the numeral system of this highly attrited, endangered language.

Abbreviations

BCS	=	Bosnian-Croatian-Serbian
Eng.	=	Engadinian
Lat.	=	Latin
(N/S)IR	=	(Northern/Southern) Istro-Romanian
PT	=	*plurale tantum/pluralia tantum*
Sl.	=	Slavic
Srs.	=	Sursilvan

Acknowledgements

References

Adamou, Evangelia. 2013. Replicating Spanish *estar* in Mexican Romani. *Linguistics* 51(6): 1075–1105.

Aikhenvald, Alexandra Y. 2002. *Language Contact in Amazonia*. New York: Oxford University Press.

Aikhenvald, Alexandra Y. 2003. Mechanisms of change in areal diffusion: New morphology and language contact. *Journal of Linguistics* 39(1): 1–29.

Arkadiev, Peter and Francesco Gardani. 2020. Introduction: Complexities in morphology. In Peter Arkadiev and Francesco Gardani (eds.), *The Complexities of Morphology*, 1–19. Oxford: Oxford University Press.

Ascoli, Graziadio Isaia. 1861. *Colonie straniere in Italia*. *Studi critici*. Vol. 1, 37–85. Milan: Politecnico [repr. Sala Bolognese: Forni, 1980].

Bickerton, Derek. 1981. *Roots of Language*. Ann Arbor: Karoma Publishers.

Breu, Walter. 2013. Zahlen im totalen Sprachkontakt: das komplexe System der Numeralia im Moliseslawischen. *Wiener Slawistischer Almanach* 72: 7–34.

Brugmann, Karl. 1907. *Die distributiven und die kollektiven Numeralia der indogermanischen Sprachen. Mit einem Anhang von Eduard Sievers: altnordisch* tvenn(i)r, þrenn(i)r, fernir. Leipzig: Teubner.

Buchi, Eva. 2006. Wieviel Wortbildung, wieviel Morphologie verträgt die etymologische Forschung? Bemerkungen zur Beschreibung rumänischer Slavismen. In Wolfgang Dahmen, Günter Holtus, Johannes Kramer, Michael Metzeltin, Wolfgang Schweickard, and Otto Winkelmann (eds.), *Lexikalischer Sprachkontakt in Südosteuropa*. Romanistisches Kolloquium XII, 73–90. Tübingen: Narr.

Candinas, Theo. 1982. *Romontsch sursilvan. Grammatica elementara per emprender igl idiom sursilvan*. Cuera: Ligia Romontscha.

Candinas, Theo. 2009. *Ein Elsässer im Ersten Weltkrieg/In schuldau d'Alsazia en l'Emprema uiara*. Frauenfeld: Reinhold Liebig Verlag.

Carosella, Maria. 2005. *Sistemi vocalici tonici nell'area garganica settentrionale fra tensioni diatopiche e dinamiche variazionali*. Rome: Nuova cultura.

Combi, Carlo. 1859. *Porta orientale*, Anno III. Fiume: Schubart – Rezza/Trieste: C. Coen.

Corbett, Greville G. 1979. The agreement hierarchy. *Journal of Linguistics* 15(2): 203–224.

Corbett, Greville G. 1983. *Hierarchies, Targets and Controllers. Agreement Patterns in Slavic*. London and Canberra: Croom Helm.

Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press.

Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.

Corbett, Greville G. 2006. *Agreement*. Cambridge: Cambridge University Press.

Corbett, Greville G. 2007. Deponency, syncretism and what lies between. In Matthew Baerman, Greville G., Dunstan Corbett Brown, and Andrew Hippisley (eds.), *Deponency and Morphological Mismatches*, 21–43. Oxford: British Academy and Oxford University Press.

Corbett, Greville G. 2012. *Features*. Cambridge: Cambridge University Press.

Corbett, Greville G. 2019. Pluralia tantum nouns and the theory of features: A typology of nouns with non-canonical number properties. *Morphology* 29(1): 51–108. DOI 10.1007/s11525-018-9336-0.

Dahmen, Wolfgang and Johannes Kramer. 1976. Observații despre vocabularul istroromânei vorbite la Jeiăn. *BANF* 1: 79–89.

Dahmen, Wolfgang and Johannes Kramer. 1988. Le inchieste istro-rumene di Ugo Pellis. Parte prima: Questioni 1-1512. *Balkan-Archiv* (Neue Folge) 13: 209–281.

de Groot, Casper. 2008. Morphological complexity as a parameter of linguistic typology: Hungarian as a contact language. In Matti Miestamo, Kaius Sinnemäki and Fred Karlsson (eds.), *Language Complexity*: *Typology, Contact, Change*, 191–215. Amsterdam and Philadelphia: John Benjamins.

Despić, Miloje. 2017. Investigations in mixed agreement: Polite plurals, hybrid nouns and coordinate structures. *Morphology* 27(3): 253–310.

Everett, Caleb. 2017. *Numbers and the Making of Us. Counting and the Course of Human Cultures*. Cambridge, Mass.: Harvard University Press.

Fedden, Sebastian and Greville G. Corbett. 2017. Gender and classifiers in concurrent systems: Refining the typology of nominal classification. *Glossa: A Journal of General Linguistics* 2(1): 34. 1–47.

Filipi, Goran. 2002. *Istrorumunjski lingvistički atlas. Atlasul Lingvistic Istroromân. Atlante Linguistico Istrorumeno*. Pula: Znanstvena udruga Mediteran.

Frățilă, Vasile and Gabriel Bărdășan (2010). *Dialectul istroromân. Straturi etimologice. Partea I*. Timişoara: Editura Universităţii de Vest.

Ganzoni, Gian Paul. 1977. *Grammatica ladina. Grammatica sistematica dal rumantsch d'Engiadin' Ota per scolars e creschieus da lingua rumauntscha e tudas-cha*. Samedan: Lia Rumantscha.

Ganzoni, Gian Paul. 1983. *Grammatica ladina. Grammatica sistematica dal rumantsch d'Engiadina Bassa per scolars e creschüts da lingua rumantscha e francesa*. Samedan: Lia Rumantscha.

Gardani, Francesco. 2008. *Borrowing of Inflectional Morphemes in Language Contact*. Frankfurt am Main: Peter Lang.

Gardani, Francesco. 2012. Plural across inflection and derivation, fusion and agglutination. In Lars Johanson and Martine I. Robbeets (eds.), *Copies versus Cognates in Bound Morphology*, 71–97. Leiden and Boston: Brill.

Gardani, Francesco. 2018. On morphological borrowing. *Language and Linguistics Compass* 12(10): 1–17.

Gardani, Francesco. 2020. Morphology and contact-induced language change. In Anthony Grant (ed.), *The Oxford Handbook of Language Contact*, 96–122. Oxford: Oxford University Press.

Gioiosa, Matteo. 2000. *Grammatica del dialetto sannicandrese*. San Nicandro Garganico: Gioiosa.

Giudici, Alberto and Chiara Zanini. 2021. A plural indefinite quantifier on the Romance-Slavic border. *Word Structure* 14: 195-225.

Givón, Talmy. 1979. Prolegomena to any sane creology. In Ian F. Hancock (ed.), *Readings in Creole Studies*, 3–35. Amsterdam and Philadelphia: John Benjamins.

Gusmani, Roberto. 1979. Sull'induzione di morfemi. In Gerhard Ernst and Arnulf Stefenelli (eds.), *Sprache und Mensch in der Romania. Heinrich Kuen zum 80. Geburtstag*, 110–116. Wiesbaden: Steiner [repr. in *Saggi sull'interferenza linguistica*, 2nd edn. Firenze: Editrice Le Lettere, 1986: 155–164].

Haspelmath, Martin and Uri Tadmor (eds.). 2009. *Loanwords in the World's Languages: A Comparative Handbook*. Berlin: De Gruyter Mouton.

Hockett, Charles F. 1958. *A Course in Modern Linguistics*. New York: Macmillan.

Hurford, Jim. 2003. The interaction between numerals and nouns. In Frans Plank (ed.), *Noun Phrase Structure in the Languages of Europe*, 561–620. Berlin and New York: Mouton De Gruyter.

Hurren, Anthony H. 1969. Verbal aspect and archi-aspect in Istro-Rumanian. *La Linguistique* 5(2): 59–90.

Hurren, Anthony H. 1999. *Istro-Rumanian: A Functionalist Phonology and Grammar*. PhD dissertation, University of Oxford.

Kalsbeek, Janneke 1998. *The Čakavian Dialect of Orbanići near Žminj in Istria*. Amsterdam and Atlanta: Rodopi.

Kämpf, Sylvina. 2015. *Der Kollektiv in der Romania – eine Untersuchung des Bündnerromanischen*. BA thesis, University of Zurich.

Karlsson, Fred, Matti Miestamo and Kaius Sinnemäki. 2008. Introduction: The problem of language complexity. In Matti Miestamo, Kaius Sinnemäki and Fred Karlsson (eds.), *Language Complexity: Typology, Contact, Change*, vii–xiv. Amsterdam and Philadelphia: John Benjamins.

Keil, Heinrich. 1856–1880. *Grammatici Latini*. 7 vols. Leipzig: Teubner [repr. Hildesheim: Olms, 2007].

Kim, Hyoung Sup. 2009. *The structure and use of collective numeral phrases in Slavic: Russian, Bosnian/Croatian/Serbian, and Polish*. PhD dissertation, University of Texas at Austin.

Klockmann, Heidi. 2017. *The Design of Semi-lexicality: Evidence from Case and Agreement in the Nominal Domain*. Utrecht: LOT.

Kossmann, Maarten. 2010. Parallel System Borrowing: Parallel morphological systems due to the borrowing of paradigms. *Diachronica* 27(3): 459–487.

Kovačec, August. 1963. Notes de lexicologie istroroumaine. Sur la disparition des mots anciens et leur remplacement par des mots croates. *Studia Romanica et Anglica Zagrabiensia* 15–16: 3–39.

Kovačec, August. 1966. Quelques influences croates dans la morphosyntaxe istroroumaine. *Studia Romanica et Anglica Zagrabiensia* 21–22: 57–75.

Kovačec, August. 1968. Observations sur les influences croates dans la grammaire istroroumaine. *La Linguistique* 1: 79–115.

Kovačec, August. 1971. *Descrierea istroromânei actuale*. Bucharest: Editura Academiei Republicii Socialiste România.

Kovačec, August. 1992. Éléments italiens du lexique istroroumain. *Linguistica* 32(2): 159–175.

Kovačec, August. 1998. *Istrorumunjsko-Hrvatski Rječnik* (*s gramatikom i tekstovima*). Pula: Znanstvena udruga Mediteran.

Kusters, Wouter. 2003. *Linguistic Complexity*: *The Influence of Social Change on Verbal Inflection*. Utrecht: LOT.

Leko, Nedžad. 2009. *The Syntax of Numerals in Bosnian*. Munich: Lincom Europa.

Livescu, Maria. 2008. Histoire interne du roumain: morphosyntaxe et syntaxe/Interne Sprachgeschichte des Rumänischen: Morphosyntax und Syntax. In Gerhard Ernst, Martin-Dietrich Gleßgen, Christian Schmitt and Wolfgang Schweickard (eds), *Romanische Sprachgeschichte/Histoire linguistique de la Romania. 3. Teilband*, 2646–2692. Berlin and New York: De Gruyter.

Löfstedt, Bengt. 1958. Zum Gebrauch der lateinischen distributiven Zahlwörter. *Eranos* 56: 71–117, 188–223.

Loporcaro, Michele. 2018. *Gender from Latin to Romance: History, Geography, Typology*. Oxford: Oxford University Press.

Loporcaro, Michele and Lorenzo Tomasin. 2016. Marcamento di genere iperdifferenziato come vestigio del neutro nell'italo-romanzo settentrionale antico. *Lingua e Stile* 51: 37–64.

Lučić, Radovan. 2015. Observations on collective numerals in Standard Croatian. *Journal of Slavic Linguistics* 23(1): 3–31.

Małecki, Mieczysław. 1930. *Przegląd słowiańskich gwar Istrji* [Overview of the Slavic Dialects of Istria]. Kraków: Polska akademja umiejętności.

Martin, Samuel Elmo. 2004. *A Reference Grammar of Japanese.* Honolulu: University of Hawai'i Press [first published by Yale UP, New Haven and London 1975; revised edn. by Charles E. Tuttle Company, Rutland, Vermont/Tokyo, Japan 1988].

Matras, Yaron. 2007. The borrowability of grammatical categories. In Yaron Matras and Jeanette Sakel (eds.), *Grammatical Borrowing in Cross-linguistic Perspective*, 31–74. Berlin and New York: De Gruyter.

Matras, Yaron. 2009. *Language Contact*. Cambridge: Cambridge University Press.

McGill, Stuart J. 2009. *Gender and Person Agreement in Cicipu Discourse*. PhD dissertation, SOAS, University of London.

McWhorter, John H. 2001. The world's simplest grammars are creole grammars. *Linguistic Typology* 5(2–3): 125–166.

McWhorter, John H. 2007. *Language interrupted. Signs of non-native acquisition in standard language grammars*. Oxford and New York: Oxford University Press.

Meakins, Felicity and Sasha Wilmoth. 2020. Overabundance resulting from language contact: Complex cell-mates in Gurindji Kriol. In Peter Arkadiev and Francesco Gardani (eds.), *The Complexities of Morphology*, 81–104. Oxford: Oxford University Press.

Melissaropoulou, Dimitra. 2017. On the role of language contact in the reorganization of grammar: A case study on two Modern Greek contact-induced varieties. *Poznan Studies in Contemporary Linguistics* 53(3): 449–485.

Mihăilă, Gheorghe. 1960. *Împrumuturi vechi sud-slave în limba romînă. Studiu lexico-semantic*. Bucharest: Editura Academiei Republicii populare romîne.

Mühlhäusler, Peter. 1980. Structural expansion and the process of creolization. In Albert Valdman and Arnold R. Highfield (eds.), *Theoretical Orientations in Creole Studies*, 19–55. New York: Academic Press.

Neira Martínez, Jesús. 1955. *El habla de Lena*. Oviedo: IDEA [repr. 2005].

Neira Martínez, Jesús. 1978. La oposición 'contínuo'/'discontínuo' en las hablas asturianas. In *Estudios ofrecidos a Emilio Alarcos Llorach*, III, 255–279. Oviedo: Universidad de Oviedo [repr. in Id., *Bables y castellano en Asturias*. Madrid: Silverio Cañada, 1982: 163–186].

Nikunlassi, Ahti O. 2000. The use of collective numerals in contemporary Russian: An empirical approach. *Wiener Slawistischer Almanach* 45: 209–246.

Ojeda, Almerindo E. 1997. A semantics for the counting numerals of Latin. *Journal of Semantics* 14: 143–171.

Payne, John and Rodney Huddleston. 2002. Nouns and noun phrases. In Rodney Huddleston and Geoffrey K. Pullum (eds.), *The Cambridge Grammar of the English Language*, 323–523. Cambridge: Cambridge University Press.

Petrovici, Emil. 1962. Phonetic evolution, substitution of sounds or morphological accommodation? (In connection with the treatment of final *-o* in the Slavonic elements of Rumanian). *Revue roumaine de linguistique* 7: 17–20.

Petrovici, Émile. 1967. Le neutre en istro-roumain. *To honor Roman Jakobson. Essays on the occasion of his seventieth birthday, 11 October 1966*, 1523–1526. The Hague and Paris: Mouton.

Petrovici, Emil and Petru Neiescu. 1964. Persistenţa insulelor lingvistice. Constatări făcute cu prilejul unor noi anchete dialectale la istroromîni, meglenoromîni şi aromîni. *Cercetări de Lingvistică* 9(2): 187–214.

Petrovici, Emil and Petru Neiescu. 1965. Persistence des îlots linguistiques. Constatations faites à l'occasion de nouvelles enquêtes dialectales chez les istro-roumains, mégléno-roumains et aroumains. *Revue roumaine de linguistique* 10(4): 351–374.

Pranjković, Ivo. 2000. *Hrvatski jezik. Udžbenik za 3. razred gimnazije*. Zagreb: Školska knjiga.

Puşcariu, Sextil (în colaborare cu d-nii M. Bartoli, A. Belulovici şi A. Byhan). 1906. *Studii istroromâne*. Vol. 1. *Texte*. Bucharest: Cultura naţionala.

Puşcariu, Sextil (în colaborare cu M. Bartoli, A. Belulovici şi A. Byhan), 1926. *Studii istroromâne*. Vol. 2. *Introducere, gramatica, caracterizarea dialectului istroromân*. Bucharest: Academia Română.

Ribarić, Josip. 1940. *Razmještaj južnoslovenskih dijalekata na poluotoku Istri* [Distribution of South Slavic Dialects on the Istrian Peninsula]. Belgrade: Srpska Kraljevska Akademija.

Rolfe, John C. (ed.). 1982. *Ammianus Marcellinus* [The Loeb Classical Collection; 1st edn. 1935]. Cambridge, Mass.: Harvard University Press and London: William Heinemann Ltd.

Sala, Marius. 2013. Contact and borrowing. In Martin Maiden, John C. Smith and Adam Ledgeway (eds.), *The Cambridge History of the Romance Languages*. Vol. 2. *Contexts*, 187–236. Cambridge: Cambridge University Press.

Schulte, Kim. 2009. Loanwords in Romanian. In Martin Haspelmath and Uri Tadmor (eds.), *Loanwords in the World's Languages*: *A Comparative Handbook*, 230–259. Berlin: De Gruyter Mouton.

Seifart, Frank. 2015. Direct and indirect affix borrowing. *Language* 91(3): 511–532.

Šipka, Danko. 2007. *A Comparative Reference Grammar of Bosnian, Croatian, Serbian*. Edited by R. David Zorc . Hyattsville, Md.: Dunwoody Press.

Spescha, Arnold. 1989. *Grammatica sursilvana*. Cuera: Casa editura per mieds d'instrucziun.

Stankiewicz, Edward. 1983. The collective and counted plurals of the Slavic nouns. In Michael S. Flier (ed.), *American contributions to the Ninth International Congress of Slavists*, Kiev, September 1983, Vol. 1. *Linguistics*, 277–292. Columbus, Ohio: Slavica Publishers [repr. in Edward Stankiewicz, *The Slavic Languages: Unity in diversity*, 173-170. Berlin: Mouton De Gruyter 1986].

Stefanović, Aleksandar. 2011. Les numéraux en serbo-croate (bosniaque, croate, monténégrin, serbe). Normes des standards et problèmes syntaxiques. *Revue des études slaves* 82(4): 709–714.

Stefanović, Aleksandar. 2014. Kvantifikacija imenica pluralia tantum. *Sarajevski filološki susreti II*, 45–67. Sarajevo: Bosansko filološko društvo.

Stevanović, Mihailo. 1989. *Savremeni srpskohrvatski jezik. Gramatički sistemi i književnojezička norma. I. Uvod. Fonetika. Morfologija* [Contemporary Serbo-Croatian Language. Grammatical Systems and Literary Norm. Vol. 1. Introduction. Phonetics. Morphology]. Belgrade: Naucna knjiga.

Tagliani, Roberto. 2011. *Il Tristano corsiniano. Edizione critica*. Rome: Scienze e Lettere.

Tagliavini, Carlo. 1972. *Le origini delle lingue neolatine*. 6th edn. Bologna: Pàtron.

Taylor, Daniel J. 1974. *Declinatio. A Study of the Linguistic Theory of Marcus Terentius Varro*. Amsterdam: John Benjamins.

ThLL = Thesaurus linguae Latinae. Editus auctoritate et consilio academiarum quinque Germanicarum Berolinensis Gottingensis Lipsiensis Monacensis Vindobonensis. Leipzig: Teubner 1900ff.

Thomason, Sarah G. and Terrence Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.

Trask, Robert L. 1997. *A Student's Dictionary of Language and Linguistics*. London: Arnold.

Trudgill, Peter. 2009. Sociolinguistic typology and complexification. In Geoffrey Sampson, David Gil and Peter Trudgill (eds.), *Language Complexity as an Evolving Variable*, 98–109. Oxford: Oxford University Press.

Trudgill, Peter. 2011. *Sociolinguistic Typology: Social Determinants of Linguistic Complexity*. Oxford: Oxford University Press.

Vallortigara, Giorgio and Nicla Panciera. 2014. *Cervelli che contano*. Milan: Adelphi.

Vanhove, Martine. 2001. Contacts de langues et complexification des systèmes: le cas du maltais. *Faits de Langues* 18: 65–74.

Vrzić, Zvjezdana and Robert Doričić. 2014. Language contact and stability of basic vocabulary: Croatian words for body parts in Vlashki/Zheyanski (Istro-Romanian). *Fluminensia* 26(2): 105–122.

Vuletić, Nikola. 2014. Les minorités linguistiques invisibles et/ou cachées de la Croatie: les communautés arbënishtë, istro-roumaine et istriote. In Ksenija Djordjević Léonard (ed.), *Les minorités invisibles: diversité et complexité (ethno)sociolinguistiques*, 182–192. Paris: Michel Houdiard Éditeur.

Whitney, Arthur H. 1956. *Teach yourself Finnish*. London: The English Universities Press Ltd.

Eastern and Western Romance in the Balkans – the Contrasting but Revealing Positions of the Danubian Romance Languages and Judezmo

Victor A. Friedman
Andrew W. Mellon Distinguished Service Professor Emeritus, Humanities and Department of Linguistics, The University of Chicago, Chicago, Illinois, USA, and Honorary Associate, Department of Languages and Linguistics, La Trobe University, Melbourne, Australia
vfriedm@uchicago.edu

Brian D. Joseph
Distinguished University Professor, Department of Linguistics, and The Kenneth E. Naylor Professor of South Slavic Linguistics, Department of Slavic and East European Languages and Cultures, The Ohio State University, Columbus, Ohio, USA
joseph.1@osu.edu

Abstract

The fate of two languages in the Balkans under conditions of language contact is discussed here. These languages, representing different branches of the Romance family, are the Ibero-Romance language Judezmo from the eastern branch and the South Danubian language Aromanian from the western branch. Both have been subject to intense contact with other languages in the Balkans but they show differential outcomes of this contact and thus differential degrees of involvement in the Balkan sprachbund. We document the similarities and differences in these outcomes, offer an explanation of their causes, and discuss the consequences they have for understanding the Balkan sprachbund.

Keywords

Aromanian – Judezmo – Balkan sprachbund – Balkanism

1   Introduction

The Balkans – the southeast European peninsula named for the mountain range running through the eastern part of that peninsula – are now, as they have been in the past, home to several languages of the Italic branch within Indo-European, including Latin in ancient times and various Romance languages more recently. While the similarities among these latter languages are interesting and useful for understanding the development of the Romance languages as a group, the differences they show offer especially important perspectives on the language contact situation in this part of the world. In what follows, we explore these differences and their import for understanding the Balkans. Various preliminaries about Romance languages and about Balkan language contact are needed to set the stage for this exploration; these are provided in Sections 2 and 3, after which we focus on two specific Romance languages in the region, Aromanian and Judezmo, also known as Judeo-Spanish or Ladino, and then present the contrasts in their development in Section 4 and discuss the consequences of these contrasts in Sections 5 and 7, with a specific case-study of the developments with the infinitive in Section 6.[1]

Our basic thesis here is that a comparison of Aromanian and Judezmo is very revealing regarding the Balkan sprachbund in that Aromanian is very "Balkan" along various parameters to be argued for, while Judezmo is less so, much less so actually. Thus, this is a "tale of two languages" and pits eastern Romance versus western Romance in the Balkans.[2]

---

1   In using the name Judezmo, we follow Bunis (2018: 185–187), who gives a thorough discussion of almost all the names this language has been known by since its inception. (Victor A. Friedman can add that he has heard the term *Spanyol* in what is now North Macedonia). Here Bunis' (2018: 189) conclusion after a discussion of the various names, their origins and meanings is worth quoting: "Nevertheless, *djudezmo* still enjoys some popular use among native speakers and is the name preferred by many Jewish-language scholars – as a unique innovation arising within the speaker community; because of its designation of the language as a 'Jewish language', sharing terminological parallels with some other Jewish languages (e.g., Yiddish); and as a memorial to major Judezmo-speaking communities, such as those of Salonika, Bitola (Monastir), and Rhodes, many of whose everyday members called their language *djudezmo* until they were annihilated in the Holocaust."

2   Victor A. Friedman wishes to thank the John Simon Guggenheim Memorial Foundation and the Fulbright-Hays Program whose fellowships enabled him to conduct research among the last remaining speakers of Judezmo as well as many Aromanian speakers in what is now North Macedonia in 2008–2009. The material adduced here for Aromanian and Judezmo is all from published sources as noted in the list of references, and thus mostly treats 20th century Judezmo, but the text as a whole is also informed by Victor A. Friedman's fieldwork, with the support of the aforementioned fellowships, and by Brian D. Joseph's interviews in the early 1980s with some diaspora Judezmo speakers originally from Thessaloniki, there being very

2   Romance Languages in the Balkans

We first contextualize the mention here of a pair of Romance languages in the Balkans. There are actually a number of Romance languages in this part of the world, and five of them are of interest here, spanning, as indicated in Section 1, the two major branches of Romance: Eastern Romance and Western Romance. We exclude from consideration here both Dalmatian, once spoken along the coast of the Adriatic in what is now Croatia and on some of the Adriatic islands, as it is a poorly attested and now-extinct language (as of June 10, 1898 when the last user of the language, Tuone Udaina, died in an accident), and Italian, spoken still in those parts of Slovenia and Croatia that are adjacent to Italy.

Four eastern languages are relevant to our discussion, namely those that constitute the Balkan Romance branch within the east: the North Danubian Balkan Romance languages Romanian (see Maiden, 2021) and Istro-Romanian (see Loporcaro, Gardani and Giudici, 2021), and the South Danubian Balkan Romance languages Aromanian and Meglenoromanian.[3] These languages are the outgrowth of Latin, first brought to the Balkans in the third century AD, though it is a matter of some debate as to whether there is continuity in Romania with those Latin speakers. Of these, Aromanian is of particular interest here, for reasons that become clear in Section 3. One western language is at issue here, the Ibero-Romance language Judezmo, brought to the Balkans by Spanish-speaking Jews driven out of Iberia at the end of the 15th century.

Both Aromanian and Judezmo are spoken in various places in the Balkans, all of which were part of the Ottoman Empire. Aromanian is spoken in what is now Greece, North Macedonia, Albania, and southwestern Bulgaria – and, until recently, also what is now Kosovo – with diaspora communities in Serbia, Romania, and elsewhere. Judezmo, unlike Aromanian, is spoken both within and outside of the Balkans (especially North Africa, where it is also known as Haketia). In the Balkans, Judezmo was spoken in Ottoman towns, in what is now Greece, North Macedonia, Kosovo, Serbia, Montenegro, Bosnia-Hercegovina, the Dalmatian Coast, and Bulgaria; eventually, some Judezmo speakers migrated from Bosnia-Hercegovina and Bulgaria to Romania, settling mainly in Bucharest.

---

few speakers left in Thessaloniki now. Like all languages, Judezmo, where it is still spoken, continues to change, but the phenomena observed here are still part of the living language in at least some places, and represent the contact-induced changes that are the point of this paper.

3   Whether Moldovan constitutes a separate language within this branch distinct from Romanian is a question that is beyond the scope of the present contribution. See Dyer (1996; 1999), who notes that the Standard Moldovan has the same Wallachian dialect base as Standard Romanian. Since 2013, the official language of the Republic of Moldova has been Romanian.

Our attention is largely on the Judezmo and Aromanian of Greece, but we make reference to the other locales as appropriate for our argumentation.

3   Language Contact in the Balkans

The Balkans have long been a locus of language contact, from ancient times to the present. The point of departure for a comparison of Aromanian and Judezmo is the well-known outcome of intense contact among speakers of different languages in the Balkans alluded to in Section 1, the so-called Balkan sprachbund, an artifact of both language geography and language contact.[4] A sprachbund can be defined as a collection of geographically connected languages that through multilateral, multigenerational, mutual, multilingual contact over hundreds of years have come to share certain structural and lexical characteristics. For the Balkans, these characteristics are generally referred to as "Balkanisms". In the linguistic literature on the Balkans, Balkanisms are generally held to be structural features, elements of grammar, especially morpho-syntactic in nature, but the languages show convergence with regard to phonological features as well.

In addition to the structural side of Balkan sprachbund convergence, there is – as the definition given here indicates, and as noted by Trubetzkoy (1923; 1930) – a lexical side as well, as suggested already by Miklosich (1861), a key figure in 19th century Balkanistics. In Friedman and Joseph (2014; 2022), in a consideration of the lexical convergence of the sprachbund, we recognize a special set of conversationally based loanwords that arise under the sprachbund conditions adduced above, i.e., precisely the conditions that give rise to the structural convergence. We develop the notion of the "ERIC loan", an acronym standing for loans that are "*E*ssentially *R*ooted *I*n *C*onversation".[5] ERIC loans depend on – and thus demonstrate – intimate and sustained (both socially and temporally) everyday conversational interactions among speakers, and they include such items as the following:

(1)   Kinship terms
        – Numerals
        – Pronouns
        – Adpositions

---

4   Following Friedman and Joseph (2022), we treat this term as an assimilated loanword from German into English.
5   The acronym also pays homage to the late Eric Hamp, a mentor to both of us, the dean of Balkan linguistics and a Balkanist par excellence.

- Negatives
- Complementizers
- Discourse elements (connectives, attitudinal expressives, interjections, gestures)
- Vocatives
- Onomatopoeia
- Reduplication (especially of an expressive nature)
- Shared phraseology

These categories of ERIC loans in some instances involve closed class items and grammatical forms not usually thought of as (easily) borrowable, and yet they are borrowed throughout the Balkans, attesting to an intense and sustained kind of language contact in the region leading to the sprachbund.

As the brief characterization of Aromanian and Judezmo in Section 1 shows, not all of the languages that are located geographically in the Balkans show Balkanisms and ERIC loans to the same degree; rather, just a subset of these languages show a considerable number of these features. Accordingly, it is convenient to recognize a distinction between "languages of/in the Balkans" – a geographical notion – and "Balkan languages", those languages of/in the Balkans that show contact-induced convergent features, and thus participate in the contact that created the Balkan sprachbund. Of the languages of interest here, Aromanian, Meglenoromanian, and, Romanian are "Balkan languages" in this categorization, as well as of course, "languages of the Balkans", and Istro-Romanian is just a language of the Balkans; Judezmo is, of course, a language of the Balkans, but its status as a "Balkan language", in the sense adopted here, is less well defined, as the discussion below indicates. That is, returning to our basic claim, a comparison of Aromanian and Judezmo shows Aromanian to be very much a Balkan language, and Judezmo somewhat less so, and this difference is revealing regarding the Balkan sprachbund.

4   Comparing Aromanian and Judezmo

How does one substantiate a claim as to the degree of "Balkan-ness" of the languages in question, operating with a distinction between "language of the Balkans" and "Balkan language"? Since the latter is a language that shows "Balkanisms", a necessary first step is to assemble a set of Balkanisms and see where a given language falls with respect to that set. We caution, however, against any inference that there is a purely quantitative answer to the question of relative "Balkan-ness", since the judgment as to whether a particular

feature is found in a language is most often not a matter of simply seeing if some speaker or other accepts that feature; rather, a more nuanced concern for dialectology and history is needed. For instance, in Albanian, Aromanian, and Macedonian, there is a construction with an impersonal nonactive verb form in Albanian – in Aromanian and Macedonian with a functionally equivalent active verb form with a reflexive marker – and a pronominal dative of interest that gives the meaning 'I feel like VERB-ing', e.g., 'I feel like eating' in (2).[6]

(2)   a.   *më*   *hahet*   (Albanian)
           me.DAT   eat.3SG.NACT
      b.   *nji-si*   *mãcã*   (Aromanian)
           me.DAT-3.REFL   eat.3SG.PRS'
      c.   *mi*   *se*   *jade*   (Macedonian)
           me.DAT   REFL   eats.3SG.PRS

For most varieties of Greek, including the standard language, such a construction is not possible, but it does occur in the Greek of the area in the north of Greece around Kastoria, an area where Greek speakers have been in contact with Albanian, Aromanian, and Macedonian speakers, as in (3).

(3)   *mi*   *trojiti*   (Kastoria Greek)
      me.ACC   eat.3SG.NACT

So, does Greek have that construction? In a very real sense, it does since some dialects of Greek show it, but at the same time, one could say it does not. That is, in general in Greek, this particular construction does not occur, but for some speakers of Greek, in particular those who have been in contact with a language that has that construction, it does occur. Thus there is a qualitative dimension to any determination of whether a given feature is found in a language.

In the remainder of this section, we present a listing of some of the more prominently discussed Balkanisms and an indication of where Aromanian and Judezmo stand with respect to them. Of particular concern for the discussion here are those features for which one or both of the languages show "yes" in the listing below; if both languages show "no", then there is no contribution to a measuring of the degree of "Balkan-ness". But a "yes" indicates the possibility of a Balkan contact-based explanation for the convergent element in the language in question. However, for many of these features, the answer is not a simple yes or no, but rather is "yes, but" or "no, but", where the "but" reveals

---

6   See Papadamou and Papanastassiou (2013), Papadamou (2019) on this construction in Kastoria Greek and its counterparts in other Balkan languages; it is discussed also in Friedman and Joseph (2018; 2022: Ch. 7). These works are the sources for the data in (2) and (3).

the qualitative side to the evaluation of the occurrence of a particular feature. Moreover, as becomes clear below, in some instances, a "no" is actually a partial "yes" (as with the "feel-like" construction in (2) and (3)), so that qualitative comments are essential. Admittedly, this list lends itself to a purely quantitative, "scorecard"-like, interpretation of what it means to be more or less "Balkan", but the qualitative commentary that is provided is what we consider to be most important as it demonstrates that one cannot assess the presence or absence of a feature in a superficial manner.[7] With such caveats in mind, we turn now to the features and the relevant qualitative discussion, presenting them in the lettered items (A) through (Q) with numbered examples where appropriate, and relevant discussion.[8]

(A)   case reduction in the nominal system
        Aromanian   YES (but ...)
        Judezmo   YES (but ...)

For Aromanian, it must be noted that case distinctions are still present in the language, as it differentiates between a nominative-accusative form and a genitive-dative form for the definite article and certain pronouns. The specific genitive-dative merger is characteristic of many other Balkan languages and may in itself be a Balkanism. Nonetheless, Aromanian, like the rest of Balkan Romance, is anomalous among Romance languages in showing case distinctions at all, even if reduced from what is found in Latin. As for Judezmo, the lack of case distinctions is a total lack, but it is characteristic of Western Romance more generally, found in Portuguese, French, and Catalan, among other languages; thus this particular feature of Judezmo was brought to the Balkans from Iberia, so that overall, the apparent similarity between Aromanian and Judezmo as far as case reduction is concerned is of no consequence from a Balkanological and language-contact perspective.

(B)   enclitic definite article
        Aromanian   YES
        Judezmo   NO (but ...)

---

7   See Friedman and Joseph (2022: Sections 1.2.3, 3.3, 3.4.2.2) for a critique of the purely quantitative approach to the Balkan sprachbund. See also Friedman and Joseph (2017).
8   Given the space limitations, we cannot go into great detail here, both in the description of the feature in question and in the qualitative commentary; see Friedman and Joseph (2022) for a fuller discussion of all the Balkan features mentioned here.

In this case, the absence from Judezmo of the Balkan feature of a definite article that is enclitic within the noun phrase, as in Aromanian *om-lu bun* 'the good man' (literally 'man-the good'), must be viewed against the backdrop that the language came from Iberia to the Balkans with a well-developed article system.[9] Thus one might well wonder why it would show change in the direction of the Balkan system at all. This absence might also be connected with the chronology of the entry of Judezmo into the Balkans, a point taken up in Section 7 below.

(C)   analytic comparatives
          Aromanian   YES (but …)
          Judezmo   YES (but …)

Both Aromanian and Judezmo show analytic marking of adjectival comparison, a feature found in all of the Balkan languages. However, this feature is found as well all across the Romance languages, so that the Judezmo analytic comparative most likely was brought from Iberia. Thus as with (A), this similarity has no significance as far as the Balkan sprachbund is concerned.

(D)   possessive use of dative enclitic pronouns
          Aromanian   YES (but …)
          Judezmo   NO (but …)

The reason for the hedging here is that in addition to dative clitic pronouns used to mark possession, e.g., *ínima-ñ* 'heart-my', i.e., 'my heart' (Vrabie, 2000: 52), Aromanian also has possessive adjectives, e.g., *ínima a mea* 'heart my', i.e., 'my heart' (ibid.; see also Papahagi, 1974, s.v. *meu*), and furthermore, like Spanish, Judezmo allows dative clitic pronouns that can signal a possessive sense, as in (4).[10]

---

9   In this regard, Judezmo is somewhat like Greek, which does not participate in the enclitic definite article feature in the way that Albanian, Balkan Romance, and Balkan Slavic do, most likely because Greek had a well-developed article system dating from the Classical Greek period. Balkan Slavic and Latin, by contrast, both came to the Balkans without a definite article, to judge from the evidence of Classical Latin and Old Church Slavonic, so that the development of an article, enclitic or otherwise, can be viewed in that context, and the shared fact of enclisis in Balkan Romance and Balkan Slavic is, at the least, suggestive of contact-induced convergence.

10   We thank an anonymous reviewer for bringing such examples to our attention.

(4) *le*   *kemaron*   *la*   *kaza*
     him.DAT   burned.3PL   the   house
     'They burned his house'

While it may well be that the dative in (4) is a dative of interest or disadvantage, so that the possessive reading is more a matter of pragmatic inference than grammar per se, the more usual way of expressing possession in Judezmo is with possessive adjectives (e.g., *mi* 'my', *tu* 'your', *su* 'his/her'), like Spanish (as the texts in Crews (1935) clearly show). Moreover, the Aromanian datives attach to the possessed noun, as in other Balkan languages, whereas the Judezmo instances like (4) are verbal adjuncts, even if they can be construed as signaling nominal possession. Thus while there is some basis for saying that the languages agree on this feature, the situation is not clear-cut, and they do disagree on details regarding this means of expression for possession, with Aromanian siding with the Balkan pattern.

(E) the formation of a future tense based on a reduced, often invariant,
     form of the verb 'want'
     Aromanian   YES (but ...)
     Judezmo   NO (but ...)

This feature actually presents a number of interesting issues pertaining to how to determine the degree of agreement between the languages. The future tense in Judezmo, as in other Romance languages, is based on an inflected form of the verb 'have', a formation found in Vulgar Latin.[11] Interestingly, there are Balkan languages with a 'have'-based future, including Romanian, where it is an inheritance from Vulgar Latin, and Geg Albanian, where it possibly represents a calque from the Latin or from Slavic; moreover, the Balkan Slavic languages (Bulgarian and Macedonian) show a 'have'-based future when the verb is negated.[12] As for Aromanian, while its future is based on 'want', there are Aromanian dialects in close contact with Macedonian that use 'have' in negated futures, a formation calqued on what occurs in Macedonian. Thus while there is agreement between Aromanian and Judezmo regarding 'have' in future formations, that agreement appears to be greater only if one counts

11   As an anonymous reviewer has reminded us, contemporary Judezmo quite frequently uses a periphrastic future with 'go' as an auxiliary rather than a synthetic future of the sort described here.

12   When the verb *ima* 'have' is not negated, it can still be used to refer to the future in Balkan Slavic, but with a sense of obligation or threat.

the Aromanian dialects that have converged with Balkan Slavic regarding the negated future, and in general, Judezmo appears to show a structure parallel to that found in some other Balkan languages. In the final analysis, however, since the Judezmo future represents an inheritance from Vulgar Latin, it is harder to take any parallelism between it and other Balkan languages regarding the basis of the future as definitively a Balkan contact effect. Thus although there is disagreement with regard to the occurrence of a 'want'-based future, and there is a superficial convergence in the future tense regarding 'have'-based futures, that parallelism is only superficial, as the commentary shows, and is of no Balkanological significance, as it does not result from contact in the Balkans.

(F)   pluperfect with 'have' (sometimes in absence of a perfect)
        Aromanian   YES (but ...)
        Judezmo   YES (but ...)

Both languages here show a periphrastic pluperfect formed with the verb 'have' as an auxiliary, together with a past participle. However, such a construction reflects a late Latin formation and is found all across the Romance languages, even if there has been a semantic shift from perfect meaning to that of a simple past in some of them (e.g., French and Italian and also some dialects of Romanian, but not in Aromanian). Thus, as with (C) above, the parallelism in structure is a matter of shared Romance inheritance from Latin and the similarity between the languages on this feature is not Balkanologically significant.

(G)   reduction/replacement of infinitive[13]
        Aromanian   YES
        Judezmo   NO (but ...)

The facts here are that Aromanian, like Albanian, Bulgarian, Greek, Macedonian, and Romani, as well as other Balkan Romance languages, shows the absence of an infinitive, a grammatical category and form that once existed in the language, to judge from the evidence of Latin and the rest of the (non-Balkan) Romance languages. In place of an infinitive, Aromanian uses fully finite verbs (i.e., those marked for person and number of their subject) introduced by modal marker *s(i)* or an indicative complementizer *că*. Judezmo, by contrast, preserves the Latin, and the Ibero-Romance, infinitive (as shown by the many instances in the texts found in Crews (1935) and in the examples in Quintana Rodríguez (2006: 163–169)), so that the infinitive-less Aromanian

---

13   See also Section 6 for more on the infinitive in Judezmo.

differs significantly from the infinitive-rich Judezmo. Nonetheless, some uses of the subjunctive in Judezmo, e.g., in modal questions such as 'When might we come to get you?', mirror Balkan clauses with a bare subordinating marker (SM) unaccompanied by a controlling verb, as in (5a), a type which does not occur in either Modern Spanish or North African Judezmo, cf. (5b).

(5a) *kwando   ke   te   vengamoz   a tom-ar?* (Balkan Judezmo)
      when   that   you.ACC   come.1PL   to take-INF

versus:

(5b) *Cuando   quieres   que vengamos   a   recog-er-te?* (Modern Spanish)
      when   want.2SG   that come.1PL   to take-INF-you
      (literally: 'When do.you.want that we.come to take you?')

This distributional fact in itself suggests Balkan contact influence, and this suggestion becomes all the more compelling when parallel constructions in contact languages, such as Greek in (5c), are adduced.

(5c) *Póte   na   ʻrϑúme   na   se   párume?*   (Greek)
      when   SBJV   come.1PL   SM   you.ACC   take.1PL
      'When might we come to get you?'

In this respect, the finite formation that replaces infinitives in most of the Balkan languages takes on a usage in Judezmo parallel to that found throughout the Balkans, thus aligning Judezmo with the results of the loss of the infinitive in a language like Aromanian.

(H)   prepositional marking of personal
      direct objects
      Aromanian   YES (but...)
      Judezmo   YES (but ...)

There is a parallelism here, to be sure, but it turns out to be irrelevant in and of itself, for two reasons. First, the prepositions involved are different, with *a* occurring in Judezmo but *pi/pe* in Aromanian, and, second, this prepositional usage is found here and there across Romance. It is found, for instance, in some Italian dialects and Sardinian, and, importantly, in Iberian Spanish, where the preposition in question is also *a*. Thus the occurrence of this feature in Judezmo

and Aromanian would seem to continue a Romance construction, and in any case, for Judezmo, it is just continuing an Iberian Spanish construction.

(I)   doubling (cross-indexing) of an object by a weak ("clitic") pronoun
        Aromanian   YES
        Judezmo   YES (but …)

In this case, despite the structural parallelism, it is not at all clear that there is anything significant here as far as Balkan language contact is concerned, because this same construction occurs in Iberian Spanish; thus most likely, its occurrence in Judezmo reflects the language's Ibero-Romance origins and was brought to Balkans from Spain when Judezmo speakers took refuge in this region. However, as with the infinitive vs. subjunctive, there are Judezmo reduplications that are not typical of the Ibero-Romance type, given in (6a) and (7a) with Macedonian parallels in (6b) and (7b), all from Kolonomos (1995).

(6a) *Il   palu   tuertu   la   lumeri   lu   indireche*
      the   stick   crooked   the   fire   it.ACC   straightens

(6b) *kriv   stap   ogn-ot   go   ispravuva*
      crooked stick   fire-DEF   it.ACC   straightens
      'A crooked staff is straightened in the fire' (Kolonomos, 1995: 267)

(7a) *Al   hamor   kwandu   mas   l'   aroges   mas*
      to.DEF   donkey   how.much   more   it.ACC beg.2SG   more
      alvante   las   urezhe
      raise.3SG the   ears

(7b) *Magare-to   kolku   poveḱe   go   moliš*
      donkey-DEF   how.much   more   it.ACC   beg.2SG
      poveḱe   gi   diga   ushi-te
      more   them.ACC   raises.3SG   ears-DEF
      'The more you beg the donkey, the more it raises its ears.'

(J)   evidentiality
        Aromanian   YES (but …)
        Judezmo   YES (but …)

Evidentiality broadly speaking refers to the grammatical encoding of marking for source of information in a given statement or the speaker's attitude toward that source. Evidentiality as a broad category includes the admirative complex, where a form expressing attitude toward source can also express surprise or disbelief (see Friedman, 2012). It is found in several Balkan languages, in particular Bulgarian, Macedonian, Albanian, as well as Turkish; the standard view is that Turkish provided a model for the emergence of grammatical evidentiality in these other languages (see Friedman, 2018). Interestingly, this feature occurs in a limited way in Aromanian, in that it is found just in one dialect, that of Bela di Supră Frasheriote Aromanian in Macedonia (Friedman, 1994); it is however, a development resulting from contact with Albanian, so that it is a relatively recent, contact-induced development. As for Judezmo, some speakers of the language of Istanbul use the pluperfect as a calque on the Turkish (unwitnessed) past in -*miş*, as in this example from Varol (2001).

(8) *Kuando*   *estavan*   *en*   *l'*   *Amérika,*   *les*
     when   were.3PL   in   the   America   them.DAT
     aviya   *entra-do*   *ladrón*
     had.3SG   enter-PST.PTCP thief
     'When they were in America [i.e., absent], a thief (apparently) broke into [Turkish *girmiş*] their house.'

The distribution of evidentiality in both Aromanian and Judezmo therefore demonstrates that each language is responding to local conditions of contact, much as is the case with Greek and the impersonal construction discussed above and illustrated in (2) and (3). Moreover, the restricted nature of the occurrence of the feature in these languages raises the same questions as seen in the Greek case about how to judge a given language vis-à-vis a given feature. Still, the facts here show that each language in part, to a certain extent, has moved in the direction of a Balkan structural type, so that with regard to this feature, the agreement between the languages is of some Balkanological interest.

(K)   occurrence of a stressed mid-central vowel
       Aromanian   YES
       Judezmo   NO

Here the languages disagree, with no caveats, and Judezmo shows the absence of a feature that Aromanian exhibits. However, it is of some relevance with regard to the status of Judezmo vis-à-vis the sprachbund only insofar as

stressed schwa is considered a Balkan feature, which, as Hamp (1977) argues, it is not (cf. also Friedman and Joseph, 2022: Ch.5).

(L)   "clear" vowel system (i.e., with no "overlay" features such as nasalization or length)
        Aromanian   YES
        Judezmo   NO

Crews (1935) reports that for the Judezmo of Thessaloniki, sequences of *a/o* + *n* develop into "nasalized vowels in final position". Thus Judezmo deviates in this respect from what is found in other Balkan languages in having developed a nasal vowel, but then Geg Albanian has a rich system of nasals and long vowels, some Macedonian dialects have developed phonemic length, and length is also to be found in Lab and Çam, at the southern end of Albanian. As with (14), then, the presence of nasality in Judezmo and thus the absence of this "clear-vowel" feature from the language differentiates it from Aromanian, but not necessarily in a way that is Balkanologically significant (Hamp, 1977).

(M)   "Hissing" / "hushing" opposition (roughly: apico-dental / alveo-palatal) in fricatives and affricates, i.e., s/ʃ, ts/tʃ, and so also for voiced counterparts:
        Aromanian   YES (e.g., with [dz] vs. Romanian [dʒ])
        Judezmo   YES

With its [c], [dz], and [dʒ], Judezmo diverges from other Spanish dialects; [tʃ], by contrast, occurs in many Spanish dialects besides Judezmo. In this way, then, both Judezmo and Aromanian show movement in the direction of a Balkan phonological norm. The agreement seen in these two languages therefore gives each one a Balkan phonological system as far as affricates are concerned. This is thus a development of potential significance from a Balkanological standpoint.

(N)   Presence of at least two members of the set [ts tɕ tʃ]
        Aromanian   YES
        Judezmo   YES

As with the previous feature, the unqualified affirmative in each language can be taken to be Balkanologically important, showing the languages to both have a Balkan aspect to their respective phonological systems.

(O)   NT > ND (N = nasal; T = voiceless stop; D = voiced stop)
    Aromanian   YES
    Judezmo   NO

Here the languages disagree with no relevant qualification needed aside from the fact that the progressive voicing is localized within the Balkans; this feature is therefore a way in which Judezmo does not show a local Balkan phonology.

(P)   ø > a /#__C
    Aromanian   YES
    Judezmo   NO (but...)

The caveat for Judezmo here is due to the fact that even though forms like *amañana* 'tomorrow' (cf. Castillian Spanish *mañana*) do occur sporadically in the language, similar forms occur in Portuguese (cf. *amanha* 'tomorrow'); this suggests that the prothetic *a-* may be a Spanish or Portuguese dialectism brought to the Balkans and not a feature that arose on Balkan soil.

(Q)   ERIC loans
    Aromanian: YES
    Judezmo: YES

Both languages show a number of the conversationally based ERIC loans, though Aromanian includes a number of grammatical loans as well; a sampling of the relevant evidence is given here.

(8a)   widespread Balkan discourse items are found in both: *bre* 'hey you' (unceremonious term of address); *ayde* 'c'mon!', *na* 'here (it is); here ya go!', *aman* 'oh my; mercy!', but only one evaluative or connective type occurs in Judezmo (*zatén* 'indeed' (< Turkish *zaten*), whereas these are far more plentiful in Aromanian, e.g., *ağeaba* 'is it so?' (< Turkish *acaba*), *belchi* 'perhaps' (< Turkish *belki*), *ghio(i)a* 'as if' (< Turkish *gûya*), *sanchi* 'as if' (< Turkish *sanki*)

(8b)   a widespread Balkan taboo expression occurs in Judezmo: *asiktar* 'scram; go to hell' (from Turkish, actually stronger in force)

(8c)   bound morphology from Turkish occurs in Judezmo *-lik* e.g., *hanukalik* 'Chanukah present' (< Turkish qualitative or concrete *-lik*)

(8d)   in Aromanian, adpositions are borrowed, one even (*karşi* 'opposite' (< Turkish postposition *karşi*)) borrowed as a postposition; also, there is borrowing of pronouns (especially *-m* 'my' (< Greek *mu*)).

The agreement between Aromanian and Judezmo in this feature thus seems to be Balkanologically significant, and suggests a degree of integration into Balkan speech communities on the part of both languages. The grammatical loans in Aromanian, as in (8d), however, mark it as being more thoroughly Balkan in regard to this feature than is Judezmo.

5   Assessment

Taking stock of all the features surveyed in Section 4, we see that some features are inconclusive as to the degree of "Balkan-ness" of Judezmo, whereas others show what might be termed a trend in the direction of the language being fully "Balkan", especially in its phonology and lexicon, but also, as far as morphosyntax is concerned, in the way finite complementation is used, and, perhaps object reduplication. Taking them all together, the picture is pretty clearly one in which Aromanian is deeply embedded in the Balkan sprachbund, both structurally and lexically, whereas Judezmo is a peripheral member at best; Judezmo is lacking many features that have been identified as relevant for the Balkan sprachbund, whereas Aromanian has several Balkan features that Judezmo does not show and there are none that are found in Judezmo to the exclusion of Aromanian.

It must be noted, though, that the inconclusive features, in particular those that most likely were brought from Iberia to the Balkans by Judezmo speakers, are not completely irrelevant. That is, features such as analytic marking of adjectival comparative degree (cf. (C)), prepositional marking of direct objects (cf. (H)), and clitic object doubling (cf. (I)), even if not due to contact in the Balkans on the part of Judezmo speakers, nonetheless would make Judezmo appear to be structurally rather like its Balkan neighbors in typological terms, even if contact-induced processes of convergence were not at work. That is, whatever their origin, these features contribute to the overall "look" of the language as far as Balkan structural characteristics are concerned. Moreover, the even if these features, as part of the Romance inheritance of the language, predated the entry of Judezmo into the Balkans, their continued presence could have been enhanced by contact with Balkan languages possessing them; that is to say, contact effects are not simply a matter of gain or loss of a given feature.

6   On Causes of Differences – the Infinitive as a Case-Study

Shifting our focus somewhat, we offer a case-study into causation by way of examining why Judezmo shows only a subset of common Balkan characteristics.

To do so, we turn to developments with the infinitive, thus expanding the discussion of feature (G) above.

First, it is a fact that at least into the last decades of the 20th century, Judezmo of Thessaloniki had an infinitive with uses that parallel Castillian Spanish (Joseph, 1983: 252ff., and see the discussion and references in (G) above in Section 4). This is so even though many, if not most, speakers, in the 20th century at least, were bilingual in infinitive-less Standard Greek and in constant contact with monolingual speakers of Standard Greek, which was, after all, essentially the Greek of Orthodox Christian speakers. Moreover, the fact that the early Spanish starting point for Judezmo had an infinitive is no guarantee that the infinitive would persist, for there are Romance languages that have lost their infinitive through contact with a language with a restricted infinitive; in particular, Italian dialects in southern Italy show reduced infinitival usage as opposed to the rest of Italian (Rohlfs, 1958; see also Ledgeway, Schifano and Silvestri, 2021), possibly due to sustained contact with Southern Italy Greek (and/or Albanian (Arbëresh), see also Breu, 2021), which has an infinitive to a greater degree than the rest of Greek but much less so than a "standard" Romance language. It is thus a matter of some interest as to why Judezmo has retained its infinitive so robustly.

As a suggested answer, we note that Jewish languages in general are said to be conservative possibly due to the general segregation of Jewish communities. Such segregation would have created situations in which Jewish speakers would have less access to linguistic innovations found in the usage of coterritorial non-Jewish speakers or less willingness to adopt them. The Judeo-Greek of 16th century Constantinople offers a possible case in point, as it has archaic infinitival usage paralleling that of New Testament Greek (Joseph, 2000; 2019), and different from what occurs in the contemporary Greek of non-Jews. By contrast, Aromanian has been in the center of the Balkans, with at least the males in contact with Greek, Albanian, and/or Macedonian for centuries longer; Récatas (1934) describes gender-based village bilingualism in Aromanian communities in the Pindus region of Greece, for instance.

7   Conclusion: What we Learn about the Sprachbund

The foregoing has been a comparative exercise contrasting Aromanian and Judezmo, in which the fate of two different Romance languages was examined in their Balkan context. If we can generalize from this study, it seems that there are three lessons to draw.

First, chronology matters – having substantially more time for speakers to interact can make a difference to outcomes; the fact of less contact time for Judezmo than for Aromanian must surely have played a role in the extent to which possible structural changes in Judezmo under conditions of contact with various Balkan languages could have taken hold and been generalized throughout the language.

Second, structure matters – what a language starts with in terms of structural properties can make a difference for outcomes; the preservation of the Judezmo article as an element that could occur in first position in a noun phrase would seem to be a case in point.

Finally, and perhaps most importantly, social circumstances matter – being restricted in the nature of the interactions with speakers of other languages can make a difference to outcomes; the developments with the infinitive in Judezmo would be an example of this factor, as Judezmo speakers had more restricted access to other Balkan languages than Aromanian speakers did, and that seems to have made a difference with regard to the infinitive.

These three points converge in a way, for in their totality they offer the opportunity for interaction between speakers of different languages in a mutual, multi-lateral, multi-generational, multi-lingual mode what Friedman and Joseph (2022: Ch.8) refer to as the "four-M" model for language contact and sprachbund formation.

Abbreviations for categories absent from the Leipzig Glossing Rules

NACT　　non-active
SM　　　subordinating marker

References

Breu, Walter. 2021. Italo-Albanian: Balkan inheritance and Romance influence. *Journal of Language Contact* 14(1): 147–183.

Bunis, David. 2018. Judezmo (Ladino/Judeo-Spanish): A historical and sociolinguistic portrait. In Benjamin Hary and Sarah Bunin Benor (eds.), *Languages in Jewish Communities, Past and Present*, 185–238. Berlin: De Gruyter Mouton.

Crews, Cynthia M. 1935. *Recherches sur le judéo-espagnol dans les pays balkaniques.* Paris: E. Droz.

Dyer, Donald (ed.). 1996. *Studies in Moldovan: The History, Culture, Language and Contemporary Politics of the People of Moldova*. Boulder: East European Monographs / New York: Distributed by Columbia University Press.

Dyer, Donald. 1999. *The Romanian Dialect of Moldova: A Study in Language and Politics*. Lewiston: Edwin Mellen Press.

Friedman, Victor A. 1994. Surprise! Surprise! Arumanian has had an admirative! *Indiana Slavic Studies* 7: 79–89.

Friedman, Victor A. 2012. Perhaps mirativity is phlogiston, but admirativity is perfect: On Balkan evidential strategies. *Linguistic Typology* 16(3): 505–527.

Friedman, Victor A. 2018. Where do evidentials come from? In Alexandra Y. Aikhenvald (ed.), *The Oxford Handbook of Evidentiality*, 124–147. Oxford: Oxford University Press.

Friedman, Victor A. and Brian D. Joseph. 2014. Lessons from Judezmo about the Balkan Sprachbund and contact linguistics. *International Journal of the Sociology of Language* 226: 3–23.

Friedman, Victor A. and Brian D. Joseph. 2017. Reassessing Sprachbunds: A view from the Balkans. In Raymond Hickey (ed.), *Cambridge Handbook of Areal Linguistics*, 55–87. Cambridge: Cambridge University Press.

Friedman, Victor A. and Brian D. Joseph. 2018. Non-nominative and depersonalized subjects in the Balkans: Areality vs. genealogy. In Jóhanna Barðdal, Na'ama Pat-El and Stephen Mark Carey (eds.), *Non-Canonically Case-Marked Subjects. The Reykjavík-Eyjafjallajökull Papers*, 23–53. Amsterdam: John Benjamins.

Friedman, Victor A. and Brian D. Joseph. 2022. *The Balkan Languages*. Cambridge: Cambridge University Press.

Hamp, Eric P. 1977. On some questions of areal linguistics. In Ken Whistler et al. (eds.), *Proceedings of the 3rd Annual Meeting of the Berkeley Linguistics Society*, 279–282. Berkeley: Berkeley Linguistics Society.

Joseph, Brian D. 1983. *The Synchrony and Diachrony of the Balkan Infinitive: A Study in Areal, General, and Historical Linguistics*. Cambridge: Cambridge University Press.

Joseph, Brian D. 2000. Processes of Spread for Syntactic Constructions in the Balkans. In Christos Tzitzilis and Christos Symeonidis (eds.), *Balkanlinguistik: Synchronie und Diachronie*, 139–150. Thessaloniki: Aristotle University of Thessaloniki.

Joseph, Brian D. 2019. Can there be language continuity in language contact? In Edit Doron, Malka Rappaport Hovav, Yael Reshef, and Moshe Taube (eds.), *Linguistic Contact, Continuity and Change in the Genesis of Modern Hebrew*, 257–285. Amsterdam: John Benjamins.

Kolonomos, Žamila (ed.). 1995. *Sefardski odglasi*. Skopje: Ǵurǵa.

Ledgeway, Adam, Norma Schifano, and Giuseppina Silvestri. 2021. The negative imperative in Southern Calabria. *Spirito greco, materia romanza* again? *Journal of Language Contact* 14(1): 184–219.

Loporcaro, Michele, Francesco Gardani, and Alberto Giudici. 2021. Contact-induced complexification in the gender system of Istro-Romanian. *Journal of Language Contact* 14(1): 72–126.

Maiden, Martin. 2021. Establishing contact. Slavonic influence on Romanian morphology. *Journal of Language Contact* 14(1): 24–52.

Miklosich, Franz. 1861. *Die slavischen Elemente im Rumunischen*. Vienna: Kaiserliche-königliche Hof- und Staatsdruckerei.

Papadamou, Eleni. 2019. Northern Greek dialects of Kastoria as language varieties at the core of Balkan Sprachbund. In G. Papanastassiou et al. (eds.), *Studies in Greek Linguistics 39. Proceedings of the Annual Meeting of the Department of Linguistics, School of Philology, Faculty of Philosophy, Aristotle University of Thessaloniki April 19–21, 2018*, 793–801. Thessaloniki: Department of Linguistics, Aristotle University.

Papadamou, Eleni and George Papanastassiou. 2013. The position of the Northern Greek dialects of Kastoria among the Modern Greek dialects. In Mark Janse, Brian D. Joseph, Angela Ralli, and Metin Bagriacik (eds.), *MGDLT5. Proceedings of the 5th International Conference on Modern Greek Dialects and Linguistic Theory* (*Ghent, Belgium, September 20–22, 2012*), 389–400. Patras: University of Patras. https://pasithee.library.upatras.gr/mgdlt/article/view/2605/2842.

Papahagi, Tache. 1974. *Dicționarul dialectului aromân*. Bucharest: Editura Academiei Republicii Socialiste România.

Quintana Rodríguez, Aldina. 2006. *Geografía lingüística del Judeoespañol. Estudio sincrónico y diacrónico*. Bern: Peter Lang.

Récetas, B. 1934. *L'état actuel du bilinguisme chez les macédo-roumains du Pinde et le role de la femme dans le langage*. Paris: Librarie Droz.

Rohlfs, Gerhard. 1958. La perdita dell'infinitivo nelle lingue balcaniche. *Omagiu lui Iorgu Iordan cu prilejul împlinirii a 70 de ani*, 733–744. Bucharest: Academiei Republicii Populare Romîne.

Trubetzkoy, Nikolai S. 1923. Vavilonskaja bašnja i smešenie jazykov. *Evrazijskij vremennik* 3: 107–124.

Trubetzkoy, Nikolai S. 1930. Proposition 16. In Cornelis de Boer, Jacobus van Ginneken and Anton G. van Hamel (eds.), *Actes du Premier Congrès International des Linguistes à La Haye, du 10–15 Avril 1928*, 17–18. Leiden: A.W. Sijthoff.

Varol, Marie-Christine. 2001. Calques morphosyntaxiques du turc en judéo-espagnol: Mécanismes et limites. In Anaïd Donabédian (ed.), *Langues de diaspora, langues en contact*, special issue of *Faits de langues* 18: 85–99.

Vrabie, Emil. 2000. *An English-Aromanian* (*Macedo-Romanian*) *dictionary: With two introductory sketches on Aromanian*. University, Miss: Romance Monographs.

BRILL

Italo-Albanian: Balkan Inheritance and Romance Influence

Walter Breu
Professor of Slavic Studies, Department of Linguistics, University of
Konstanz, Konstanz, Germany
walter.breu@uni-konstanz.de

Abstract

This chapter deals with contact-induced change in Italo-Albanian and its effects on the Balkan inheritance of this minority language. The introduction is dedicated to the general characteristics of Albanian and its varieties from a historical, dialectological and geographic perspective, followed by a section on the historical and present situation of the Italo-Albanians. While Section 3 discusses the role of Balkanisms in Standard Albanian, Section 4 gives a general overview of the fate of these *Sprachbund* criteria in Italo-Albanian. In Section 5, contact-induced changes in the verb systems in single Italo-Albanian dialects are investigated, with special regard to the changes in the future tense, in the analytical perfect and verbal aspect, followed by a discussion of the innovative causative construction and other periphrases. Finally, changes in the domains of mood and voice in the Italo-Albanian dialects are described, most of them of recent date and, in part, not accepted by conservative speakers. All contact-induced developments in the minority language, as well as those parts of its grammar that have resisted foreign influence, are contrasted with their Standard-Albanian counterparts. As will be shown, many traditional Balkan features have been weakened or lost, whereas others have even expanded, but always in the direction of Romance models, to which Italo-Albanian functionally has adapted or which it has calqued.

Keywords

Italo-Albanian – Romance – Balkanism – contact-induced change – verbal aspect – causative – verbal periphrases – counterfactual mood – passive voice

1 Introduction

The history of the Albanians since the early middle ages has been a history of expansion and migration. From a very restricted territory in the Albanian mountains, with the Mati valley as its center, and possibly some other areas more to the east (Dardania) escaping complete Romanization, ancient Albanians expanded to a considerably vast territory in the south-west of the Balkan peninsula. Later on, further expansion and emigration included also territories outside the Balkans.

Nowadays, Albanian has an official status in Albania itself, in Kosovo, (North) Macedonia and Montenegro, but traditional Albanian-speaking minorities exist, for example, in Greece, Serbia, Croatia, Bulgaria and Ukraine, too, and they played an important role also outside Europe during the Ottoman-Turkish empire. Albanian was probably one of the main sources for the distribution of Balkanisms (Fiedler, 1992), which, as a matter of fact, developed on the basis of the mutual influence several languages exerted on each other, including Balkan Romance and Greek, less so Balkan Slavic.

In this paper I will be dealing with the Albanian-based minority language in Italy, which nowadays is threatened by a constant loss of native speakers, just like other alloglottic language islands of the Germanic (Walser, Cimbrian, Mocheno in northern Italy), Slavic (Resian in northern and Molise Slavic in southern Italy) and Greek (Griko/Grecanico in the extreme south of Apulia and Calabria) language families.

Italo-Albanian (ethnonym *Arbëresh*) enclaves are situated in several parts of southern Italy from Molise and Campania down to Apulia, Basilicata, Calabria and Sicily. They are the result of at least nine immigration waves from the 15th up to the 18th century, mainly from Greece as their starting point (less so from southern Albania), from where the Albanian emigrants brought their originally Greek-Orthodox faith and liturgy (Bartl, 1981). This corresponds to their mainly southern (Tosk) dialect base, while northern (Geg) elements are rather rare throughout Italo-Albanian (language name *Arbërisht*, abbreviated ARB), if they exist at all.[1] Continuous influence from Italian and southern Italo-Romance varieties has transformed the ARB varieties in many respects.

---

1 Possible candidates like the traces of a *me*-infinitive of the Geg type (see Section 3) could go back to a previously wider distribution of this construction in mainland Albania; see Altimari (2009). In the case of the Italo-Albanian *kam*-future, coinciding with the Geg future with respect to the auxiliary, I even propose a contact induced development on Italian soil (see Section 5.1).

Due to these developments, ARB could be claimed a third variety besides Tosk and Geg or even a micro-language in its own right.[2]

The two most-cited isoglosses separating Geg and Tosk (including Italo-Albanian) dialects come from historical phonology and go back to two innovations in Tosk, the so-called "rhotacism" of intervocalic *n > r* and the loss of nasal vowels, developing into schwa in the case of the accented nasal *â > ë* [ə]; see for example Geg *bâna* 'to make AOR.1SG': Tosk *bëra* or the Latin borrowing *arena* 'sand', giving *rânë* in Geg and *rërë* in Tosk.

Figure 1 is an overview of the main varieties of the Albanian language family, including *Arvanitika* in central and southern Greece, at present gradually dying out.

Common Albanian

Northern Albanian (Geg)   Southern Albanian (Tosk)

Standard Albanian   Greek Albanian (*Arvanitika*)   Italo-Albanian (*Arbërisht*)
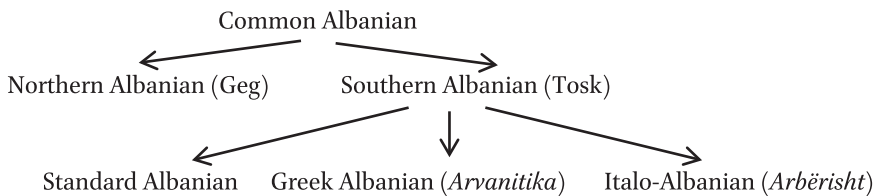
FIGURE 1   Main varieties of Albanian

This paper intends to show to what extent foreign Romance influence on Italo-Albanian has led to a loss of the Balkan traits Albanian traditionally has in common with other Balkan languages like Macedonian, Bulgarian, Romanian (and its south-Danubian varieties) and Greek, but also of specific Albanian inheritance, not found in the other Balkan languages. Cases of resistance against the foreign influence will also play a role.[3]

2   The Historical and Present Situation of the Italo-Albanians

As it seems, internal migration of Albanians in Italy occurred already immediately after their arrival in their new home country. People from different places mixed up, which means that there is no clear-cut correlation between specific places in their original home and the individual settlements in Italy. The same is, of course, true for their language/dialects. As of 2020, about fifty villages in southern Italy still have an Albanian-speaking population; others like Brindisi di Montagna in Basilicata and Mezzoiuso in Sicily as well as (only in the last

---

2   For a short overview of the situation and the structure of Italo-Albanian, see, for example, Savoia (2010).

3   See also Breu (2018c) for additional information on the topic of the present article.

decades) Villa Badessa in Abruzzo have lost it. The greater part of these municipalities concentrate in the Calabrian province of Cosenza. Italo-Albanian as a whole has transformed considerably with respect to the lexicon, but grammar has also been affected, mainly by means of the influences of Italian and its local varieties. This micro-language is, however, by no means homogeneous, mainly due to differences in the dialects of neighboring Romance-speaking villages, serving as models for contact-induced change. In the present article, I will in some cases refer to these differences, but I will also try to elaborate on common characteristics, in which Italo-Albanian as a whole differs from the Balkan-Albanian varieties. Nevertheless, many examples are based on single dialects, especially those of Frascineto in the Province of Cosenza (northern Calabria)[4] and of Portocannone in the Province of Campobasso (Molise).

The number of Italo-Albanian speakers in their traditional municipalities has continuously diminished in the last fifty years, due to their emigration to neighboring cities and to the North, in addition to language shift that has been especially strong in the last decades. Unofficial estimates run from twenty up to fifty thousand persons, using this micro-language with a certain degree of competence. The demographic figures of the official censuses normally do not consider linguistic data. Therefore, they do not reflect the real number of Italo-Albanian speakers, due to the influx of monolingual Italians, especially in regions near to the coast, in addition to language shift/loss. Nevertheless, they give some information as to growth and loss of the population in Italo-Albanian municipalities during the last 150 years. The official data of the Italian National Institute of Statistics ISTAT from Piana degli Albanesi (Sicily), Frascineto (Calabria) and Portocannone (Molise), summarized in Figs 2–4, may give some impression of the demographic development in these three municipalities. For the sake of comparison, the data for Montecilfone (Molise) are added in Fig. 5.

The difference between communities near to the coast like Portocannone, with their strong influx of monolingual Italians, and those in the hinterland like Montecilfone, likewise in Molise, becomes evident from the statistics for the latter (Fig. 5), with a continuous decline of the number of its inhabitants since 1951.[5]

---

4 For a first overview of the dialect of Frascineto, see Breu and Glaser (1979). For the Portocannone dialect, see the dictionary of Pignoli and Tartaglione (2007). An all-embracing comparison of the ARB-varieties is still missing. But see the work of L. Savoia for many details, for example Savoia (2008) and Manzini and Savoia (2015).

5 For a more detailed analysis of the Molise Albanian demographic situation (compared with neighboring Molise Slavic), including attempts at estimating the number of native speakers, see Breu (2018a).

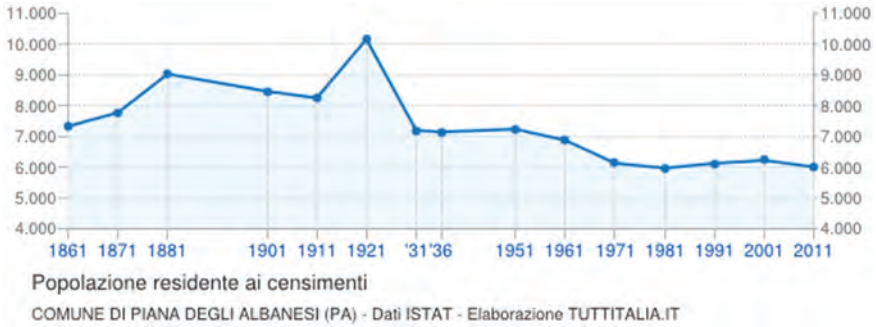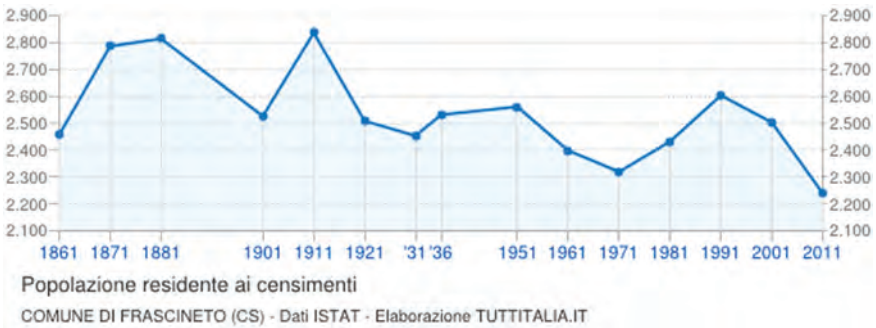FIGURE 2   Demographic data from *Piana degli Albanesi*, 1861–2011

FIGURE 3   Demographic data from *Frascineto*, 1861–2011

FIGURE 4   Demographic data from *Portocannone*, 1861–2011

ARB speakers are in a situation of "total language contact", which means that they are all bilingual, with Italian (and the local Italo-Romance varieties) dominating their vernacular as an umbrella language (*Dachsprache*), including

FIGURE 5   Demographic data from *Montecilfone*, 1861–2011

official use and writing.[6] To a growing extent Italian even enters their everyday communication. Actually, there is a great deal of variation with respect to language behavior and preservation in the different places, for example, with the ARB dialects in Molise, Campania, Apulia, and south-western Calabria being severely endangered, whereas the varieties in northern Calabria and the adjacent part of Basilicata (both of them conserving the Orthodox Greek liturgy) are relatively stable. See, for example, Breu (1991b) for an overview of the language behavior in most villages of the northern part of the Arbëria thirty years ago and Savoia (2010: Section 4), Breu (2018a) for a discussion of the actual linguistic and sociolinguistic situation in the Albanian villages of Molise.

3   Balkanisms in Albanian

In order to evaluate the role of linguistic change in ARB with respect to its Balkan features, the traditional position of Albanian in the Balkan linguistic area (*Sprachbund*) has to be considered. The best-known Balkan features (Balkanisms) found in most Balkan languages, though with variations, are:[7]

---

6   For the term "total" (or "absolute") language contact, see, for example, Breu (2011), especially Section 3, and Breu (2019: 385–386).

7   Many hypotheses exist concerning the development of these features, including a common Balkan substrate as well as Greek and Balkan-Romance influence. The concrete realization of the single Balkanisms in the different languages is by no means homogeneous, just like their presence or absence in a given language. The properties of a *Balkansprachbund*, if it exists at all, have been treated controversially in the literature; see, for example, Sandfeld (1930), Solta (1980), Fiedler (1989), Hinrichs (1999), Friedman (2006), Tomić (2006), Sims (2008), Kahl et al. (2012), and Trumper (2020). For an updated overview, see also Gardani, Loporcaro, and Giudici (2021).

1) The existence of a central vowel (schwa)
2) A common Balkan lexicon
3) A postponed definite article
4) Fossilization of modal verbs to particles based on the form of PRS.3SG
5) Formation of numbers 11–19 of the type "one upon ten"
6) Object doubling by means of clitic pronouns in the dative and accusative
7) A completely analytical system of comparison (adjectives, adverbs)
8) Very complex verb systems, including hypotactic particle constructions
9) A modal perfect
10) Lack of an infinitive
11) Future formation with the volitive modal particle based on 'will (to want)'
12) Lack of a nominal declension
    a) center: genitive-dative syncretism
    b) but: morphological expression of a vocative

Albanian shows almost all of these properties, though, of course, in their specific Albanian form. Therefore it has been claimed to be the most typical Balkan language; see Fiedler (1992) and Beci (2012). It has a stressed schwa (at least in Tosk and the standard), unlike, for example, Greek (feature 1), it has many Balkan words in common with Romanian, which, in part, are missing in other Balkan languages (feature 2), it has a postponed definite article unlike Greek (feature 3), it has more fossilized modal particles than, for example, Bulgarian and Macedonian (feature 4).

Its obligatory, probably Slavic-based, formation of the numerals 11–19 of the type *dy-mbë-dhjetë* 'twelve' (literally "two-upon-ten") is completely missing in Greek (feature 5), but it has not been extended to numerals higher than 20, unlike, for example, in Aromanian *două-spre-jingiţ* 'twenty-two' (literally "two-upon-twenty"). Albanian is the only Balkan language having traces of a vigesimal system for the tens: *njëzet* 'twenty (one score)', *dizet* 'forty (two score)'.

Object doubling by means of dative and accusative clitics is by far more grammaticalized than in Greek (feature 6). Likewise, in contrast to Greek, there is no alternative to the analytical system of comparison of adjectives and adverbs (feature 7) of the type *më shumë* 'more (much)'. The Albanian verb system is especially complex in the domain of mood, for example, in showing a morphological optative of the type *qoftë* 'may it be' and a jussive, expressed by means of the particle *lë* + subjunctive, e.g., *lë të jetë* 'let it be', unlike all the other Balkan languages (feature 8). This complexity also includes the morphological mediopassive of the type *blehet* 'is (being) bought', which Albanian has in common only with Greek and which normally does not figure among the Balkanisms. Unlike Greek and Romanian, it has a modal perfect, in its specific

realization as an admirative of the type *qenka* 'oh, s/he is' (feature 9), differing from the perfect, here *ka qenë* 'has been', mainly in the order of its components.

In contrast, the lack of an infinitive (feature 10) only concerns Tosk,[8] whereas Geg has an analytical infinitive of the type *me bâ* 'to make'. But unlike Romanian, Albanian as a whole indeed does not have any synthetic (morphological) infinitive.[9] A similar relativization holds true for the future, formed exclusively with the modal particle *do* 'will (to want)' + subjunctive (feature 11) in Tosk (and in the standard), whereas Geg traditionally prefers the inflected auxiliary *kam* 'to have' + its analytical infinitive, for example Tosk *do të punojë,* Geg *ka me punue* 's/he will work'.

As for the lack/loss of the nominal declension (feature 12), Albanian, unlike Balkan Slavic and Romanian, restricts the genitive-dative syncretism only to the morphological level, while syntactically these two cases are kept apart by means of the obligatory genitive connector,[10] missing in the dative (feature 12a). In general, Albanian has preserved case inflection, unlike Balkan Slavic and even more than Greek, as it has a distinct ablative, e.g., *grash* ABL.PL.INDF 'of/from women', and some dialects additionally show a locative, for example *në malt* LOC.SG.DEF 'in the mountain' (Gjinari, 2007: 263). But there is no vocative case, at least not in the standard and in the everyday vernacular (feature 12b), with the exception of rare agglutinative forms with postponed *-o* for persons, documented predominantly in folkloristic texts, e.g., *biro* VOC.SG.M for *bir* 'son' (Weigand, 1913: 28).

---

8   Even here, the final clause with *për* 'for' + subjunctive particle + participle, e.g., *për të punuar* '(in order' to work), is sometimes referred to as "infinitive".

9   The Romanian short infinitive, the only one used, for example, in modal constructions, is accompanied in many cases by the preposition *a*, for example, *a cânta* 'to sing', but it also appears without it, for example, in the analytical volitive future as in *voi cânta* 'I will sing' or when governed by *a putea* 'can' as in *poţi cânta* (in variation with the subjunctive *poţi să cânţi*) 'you can/may leave'. The so-called long infinitive, here *cântare* '(the) singing', is mostly used nominally.

10 Terminology varies with respect to this formally article-like linking element. Though it certainly is not an article, its form depends partially on definiteness, expressed by the definite article, in addition to other criteria like case and number agreement as well as word order, e.g., *burrat* (NOM.PL.DEF) *e fshatit* (GEN.SG.M.DEF) 'the men of the village': *disa burra* (NOM.PL.INDF) *të fshatit* (GEN.SG.M.DEF) '(some) men of the village' or *vajzës* (DAT.SG.F.DEF) *së mikut* (GEN.SG.M.DEF) 'to the daughter of the friend': *një vajze* (DAT.SG.F.INDF) *të mikut* (GEN.SG.M.DEF) 'to a daughter of the friend'. Terms like "genitive article" or "linking article" are obviously misleading.

4   Overview of the Fate of the Balkanisms in Italo-Albanian (ARB)

Italo-Albanian continues, in principle, to have a stressed schwa, often slightly nasalized, as in *zë* [zə̃] 'to take' (Balkan feature 1), whereas unstressed schwa has been given a new interpretation as a euphonic element in consonant clusters or was replaced by *e*. But at least in Calabria, many speakers, in some cases even whole dialects, replace stressed *ë* consequently with *o*, e.g., *ësht ~ osht* 'is', *ndë ~ ndo* 'if'. Some lexical Balkanisms of pre-Turkish times, especially substrate terms and Latin borrowings continue to exist (feature 2),[11] whereas others have been replaced by Italianisms. Phrase-like compounds like *kush isht isht* 'whoever it might be', literally "who (it) is, (it) is", have often been calqued from the Romance neighbors, which – just like in the case of lexical borrowings – in ARB attributes Italian and its varieties the same role Turkish had for the languages on the Balkans.

The postponed definite article has been kept in all its functions and continues to be inflected (feature 3). The Italian preponed article has not been borrowed, nor has its position been calqued.[12]

The traditional formation of the numerals 11–19 continues to exist in ARB (feature 5), e.g., *njëmbëdhjet* 'eleven', *trembëdhjet* 'thirteen', but it is severely threatened by Romance borrowings of the type *dhudhëç* 'twelve', *kuindhëç* 'fourteen', corresponding to Italian *dodici, quindici* in their local phonetic form (here Portocannone). As for the Albanian vigesimal numeral system, ARB has even expanded it to *trizet* 'sixty' and *katërzet* 'eighty', including odd decimals, missing in modern Balkan Albanian or, perhaps, it has preserved an older more complete system. Examples from Portocannone are: *trizet e di*

---

11   Pre-Latin substrate terms are normally restricted to Albanian and Romanian (Solta, 1980: 39–63), e.g., Alb./ARB *brez* (Rom. *brîu*) 'belt' and *katund* (Rom. *cătun*) 'village', conserved in Calabria, but not in Molise, where *katund* has been replaced by *hor*, borrowed from Greek. Turkish elements in the Balkan lexicon, like Alb. *bojë* 'color', do not play any role in ARB, apart from rare third-party borrowings (Mandalà, 2012), but Greek borrowings are even more frequent in Italo-Albanian than in Standard Albanian, due to the contact situation in Greece in pre-emigration times, e.g., *parathire* 'window', *dhjovasenj* 'to read'.

12   Italian articles appear, however, as parts of loanwords, especially in geographical denominations like *L-amerka* ← It. *l'America*, but also *l-universita* ← It. *l'università* 'the university'. Moreover, there are also rare cases of article-like functions of the Italian article attached to borrowed adjectives. Compare, for example, in the Portocannone dialect *lunku* ← It. *l'unico* 'the unique' in the definite phrase *lunku male çë kish kurra njohur* 'the only mountains which he had ever known' with the indefinite adjective *unku* ← It. *unico* 'unique' in *ti bëhe për mua **unku** te shekui* 'you become for me unique in the world'. For comparable cases in Sicily, see Mandalà (2005: 17). – The arrow "←" symbolizes borrowings, in contrast to phonetic or phonological developments, indicated by ">", "<".

'sixty-two = three score and two', *njëzet e dhjet* 'thirty = one score and ten', *dizet e katërmbëdhjet* 'fifty-four = two score and fourteen'. Romance borrowings are, however, excluded from this system, e.g., only *trendun* 'thirty-one', *sëtandaduj* 'seventy-two', etc. are possible and not **trizet e dhudhëç*.

Object doubling in the dative and accusative continues to exist in ARB (feature 6), but the specific rules seem to be slightly different, at least for the accusative. The preservation of object doubling by means of clitic pronouns and how it applies seems to be connected with the system of object doubling in southern Italian dialects. Further research has to be done in this field, in order to come to robust conclusions. In contrast, nothing has changed with regard to the analytical system of comparison (feature 7).

As for the verb system concerning features 8 through 11, ARB has remained rather complex, but with adaptations to the Romance tense, voice and aspect systems. Hypotactic particle constructions have even remarkably multiplied, perhaps as a reaction to the missing infinitive in contact with Italian varieties, having an infinitive. In many ARB dialects, the perfect has reduced its functions to a special type of presumptive and there is no volitive future, which in both cases seems to be due to contact-induced changes in the meanings of the verb *kam* 'to have'. These developments will be described in more detail in Sections 5.1 and 5.2.

With respect to the declension of nouns, it has, in principle, been preserved, in spite of the contact with caseless Romance varieties. Even the ablative has been kept, although with a reduction of its functions, and the locative, missing in standard Albanian also continues to be used. The morphological genitive-dative syncretism has not been broken up either (feature 12a), but the syntactic differentiation by means of the genitive connector has been weakened, in terms of a growing optionality of this element. On the other hand, there is a slight tendency to express the genitive analytically, i.e., based on the Italian prepositional model, but only in relatively fixed expressions with loanwords. Interestingly, in such cases the borrowed preposition *dhi* or *dhë* is used, corresponding to both Italian *di* 'of' and *da* 'from', which means that it also expresses ablative functions.[13]

---

13   See Mandalà (2005) for examples with *dhi* from Piana degli Albanesi (Sicily), where occasionally also the borrowed preposition *a* appears in dative-like constructions. An example with *dhë* from Portocannone is: *Ngë shkonjën kurra dhë modu*. 'They never go out of fashion.' In this dialect *dhë* also appears as part of the complex preposition (local adverb) *dhë-skuartu*, e.g., *dhë-skuartu kroghit* 'near the well'. Italian *da* 'from' has also been borrowed in Portocannone, as recorded for example in *jan fate da burrash* 'these are matters of men' (this is men's business), where *da* is pleonastic, as *burrash* is already the ABL.PL.INDF of *burr* 'man'.

Based on a southern Italian model, a neo-vocative developed, typically expressed by means of omitting all sounds following the stressed vowel of proper names and other denominations of persons, e.g., *Luixhi* NOM ≠ *Lui* VOC 'Luigi'. Thus, in a way, the Balkanism of a vocative (feature 12b) has been introduced by means of language contact, although in a very special form.

5 Contact-Induced Changes in the Italo-Albanian Verb Systems

In a situation of total language contact like that of ARB as a minority language, the concept of developing a grammatical diasystem becomes important, which claims that the grammars of the two languages in contact should become as uniform as possible. Two main procedures contribute to the development of such a diagrammar, the "adaptation of the semantic structure" (semantic calque), mainly by means of copying polysemies of the minority language to the majority language, and "loan translation" of periphrastic elements (syntactic calque). In the following, the effect of these procedures in the ARB verb system will be shown.[14]

5.1 *The Fate of the Future Tense*
The developments in the domain of the future tense are a typical case of the adaptation of the semantic structure of ARB as a recipient language, with regard to both lexicon and grammar. First of all, it should be noted that Italo-Albanian as a whole does not have a volitive Balkan future,[15] unlike the Tosk branch of Albanian. Given the predominantly southern provenience of the ancestors of modern Italo-Albanians, a connection of their *kam* future with the Geg one seems to be excluded.

If a volitional future existed in older times, the first step of the adaptation of the semantic structure to the southern-Italian future occurred in the lexicon, in so far as the verb *kam* 'to have' copied[16] the polysemy of its local Romance

---

14 For these concepts in the context of a comparison of Italian influence on Italo-Albanian and Molise Slavic, see Breu (2018b) and, with more details with respect to the concept of the "adaptation of the semantic structure", Breu (forthcoming).

15 Sometimes a volitive future has been claimed for Italo-Albanian, too, see for example Altimari (2005: 3–5), Savoia (2010). The results of my own field research show that at least in Frascineto and in Molise constructions of *dua* 'to want' + subjunctive are never interpreted or accepted as equivalents of the Italian future (not even with an additional connotation of volition).

16 The concept of "copying" instead of borrowing, calquing and other traditional terms of contact linguistics is widely used in Johanson's code-copying framework, for example

counterpart *avé* 'to have, must', thus acquiring an additional deontic meaning. The second step then regards grammar. More precisely, ARB calqued the deontic future of the surrounding Romance dialects with deontic *avé* as its auxiliary, thus getting a de-obligative *kam* future. As ARB did not have an infinitive, it could not follow the Romance model completely, but had to replace it in its usual way by combing *kam* hypotactically with the main verb in the subjunctive.[17]

ARB dialects differ with respect to the usage of the *kam* future, depending on the degree of deontic connotation. It seems weak in Calabria, but relatively strong in Molise. As a consequence the *kam* future is used freely, for example, in Frascineto (1a), whereas in Portocannone it is only possible if there is a high degree of necessity for the scheduled state of affairs or if the speaker intends to express future reference unambiguously without using a future adverbial as in (1b).[18]

Otherwise in both dialects the present tense is used to express the future, which is possible in local Romance dialects and in Standard Italian, too. For differentiating the future from the present, if necessary, time adverbials are added. At least in Frascineto, the *kam* future is also very common in epistemic statements like (1c).

(1a)   *Komungve,   menat   kem   vemi.*   (Frascineto)
       well   this.evening have.PRS.1PL come.PRS.SUBJ.1PL
       'Well, this evening we will/must come.'

(1b)   *Kat   kem   airin   t'   jem   keq,*
       have.PTL have.PRS.SUBJ.1SG   air.ACC.SG PTL be.PRS.SUBJ.1SG bad

       kat   kem   airin   të   vdes.   (Portocannone)
       have.PTL have.PRS.SUBJ.1SG   air.ACC.   PTL die.PRS.SUBJ.3SG
                                      SG
       'I will look like being sick... I will look as if I am dying.'

_____

        Johanson (2002). In the present paper it is restricted to the special case of transferring
        polysemic models from the dominant to the replica language as the reason for grammatical
        change, based exclusively on the semantic/functional level without any interference of
        concrete forms.

17   Note that the forms of the future auxiliary do not necessarily coincide completely with the
        paradigm of the full verb *kam*. In Frascineto, for example, the auxiliary loses its original final
        vowel in the 1st and 2nd person plural, with *kemi > kem* and *kini > kin*, and in the 2nd and 3rd
        person singular the particle *kat* is used, instead of *ke t, ka t*. In Portocannone, the particle *kat*
        is even used throughout the future.

18   For example, Lambertz (1914: 10–11) does not mention the de-obligative future at all. He
        simply states that future events in Molise Albanian are expressed by means of the present.

(1c) *jan   shum   gjind   jasht.   – **Kan**   **jen**
      be.3SG   many   people   outside   have.PRS.3PL   be.PRS.SUBJ.3PL
      ktje   per   t   bojen   preçsjunen.   (Frascineto)
      there   for   PTL   make.3PL   procession.ACC.SG.DEF
      'There are many people out there' – 'They will (probably) be there to
      make a procession.'

The de-obligative *kam*-future was documented in Italo-Albanian very early.
It shows up already in Matranga's Siculo-Albanian Catechism of 1592. As a
consequence, some authors claim this type of future to go back to a common
usage of a *kam*-future in Old Albanian, not restricted to the North (Altimari,
2005). Besides the problem that the ARB type of this future is characterized
by the subjunctive of the main verb and not its infinitive as in Old Albanian,
it seems important in this respect that the *kam*-future is completely absent
in Greek Albanian, which for historical reasons should be ARB's nearest rel-
ative. It only has a volitional future (Sasse, 1991: 227–228, 416–417). As older
texts in the Arvanitika varieties are missing, it is, however, unknown, whether
the de-obligative future existed there at the time of the emigration of the later
Italo-Albanians.

In any case, it may be concluded that even if an older *kam*-future survived
in Italy, it was not replaced by the volitional future, due to the existing Italian
model, whereas in the Balkan-Tosk varieties the Greek volitional future was
calqued, which eventually could be the overall source of this Balkanism. Be it
as it may, the result is the same: Italo-Albanian does not have this Balkanism,
either due to contact-induced innovation or contact-induced preservation.[19]

5.2   *The Fate of the Analytical Perfect*

The Albanian (active) perfect was already traditionally formed with the help of
the auxiliary HAVE, in combination with the participle. So there was no need of
a formal adaption to the structure of the Romance perfect, formed in the same
way. But the first, lexical step of the meaning extension of *kam* to its newly
acquired deontic meaning has to be considered in this case, too. Actually, a
type of linguistic drift occurred, in the sense that the polysemy of *kam* did not

---

19 Note that Molise Slavic, a Slavic micro-language in a similar contact situation as the
    neighboring ARB varieties in the Region of Molise, also has a de-obligative future going
    back to the Romance model, which, in principle, supports the claim of an innovation
    in Italo-Albanian. The position of this future type in the Molise Slavic system, however,
    differs from the ARB one, as here the volitional future, also found in related Croatian
    dialects, has been preserved, thus giving rise to an opposition of two modal futures (Breu,
    2018b: 220–221; Breu and Pila, 2018).

only affect the formation and deontic connotation of the future but also the semantics of the perfect. In other words, the meaning of the perfect auxiliary *kam* 'to have' extended to 'must'. As a consequence, an originally indicative perfect like *ka bën* 's/he has done' acquired a deontic reading, too. As it referred to the past and not to something still to happen, it developed a strong epistemic connotation, in the given case 's/he must (have) done', which eventually became its dominant meaning.[20] As a result, the perfect lost its indicative meaning altogether and turned into a new mood, a "presumptive" referring to the past. This development, not occurring in the local Romance dialects in question, was an internal change, clearly supported by the fact that the perfect in the contact varieties had only a very restricted function, excluding its usage as a (perfective) past tense, which was expressed only by the aorist. So, the ARB aorist could easily take over the remaining indicative functions of the perfect.

A simple example of a presumptive in the dialect of Frascineto is (2a), resulting from an assumption of the speaker, concluded from the given situation. If he had personally observed that the rain had stopped, he would have used the aorist *pundarti* instead of the perfect presumptive *ka pundartur*. The short dialogue between mother and son in (2b) presents both forms, the aorist, referring to the past without any connotation, and the presumptive perfect, expressing a (present) epistemic assumption regarding an event in the past: At first a mother claims to not have seen a certain photograph before. Her (passively bilingual) son contradicts her in Italian, by using the Italian (indicative) perfect. In her reply she gives in, by using the presumptive (with its epistemic perfect meaning).[21]

---

20   In addition to this basic motivation for the transition of the perfect from the indicative to the modal domain, a second one came from the epistemic future perfect, which originally had the structure type *have* (AUX inflected) + *have* (subjunctive) + participle, e.g., *ka t ket bën* 's/he will have done'. In this construction the subjunctive *t ket* was deleted, probably due to the model of the deletion of infinitives of auxiliaries governed by modal verbs like *volere* 'to want' in the Romance contact varieties (Rohlfs, 1969: 131). In our case this means: *deve aver fatto* 's/he must have done > *deve fatto* = ARB *ka bën*. Consequently, the future perfect became identical with the (present) perfect and transferred its epistemic function into this construction. Actually, the original future perfect has become obsolete, for example, in Frascineto. So, instead of **ka t ket fërnuar t bjer shi* 'it will **have** stopped raining / it must have stopped raining / it probably has stopped raining', only the presumptive perfect *ka fërnuar* (without the subjunctive *t ket*) is used, keeping the epistemic meaning of the future perfect (Breu, 1991a: 57–59).

21   This is a clear example of the presumptive not being restricted to the third person, a case which Altimari (1991: 54, 58) in this general restriction of the presumptive to the third person would judge as "amnesia".

(2a) *Ka*   *pundartur*   *të*   *bjer*   *shi.*

      have/must.PRS.3SG   stop.PTCP   PTL   fall.PRS.SUBJ.3SG   rain

      'It probably has stopped raining.' Literally: "It must (have) stopped raining."

(2b) *Nëng*   *e*   *pe.* –

      not   it.ACC   see.AOR.1SG

      Ma   *l'*   *hai*   *visto!* –

      but   it.ACC   have.PRS.2SG   see.PTCP

      E   *kam*   *parë.*

      it.ACC   have/must.PRS.1SG   see.PTCP

      'I haven't seen it (up to now). – But you have seen it. – I probably have seen it.'

      (~ 'I must have seen it.')

Only in part of the Italo-Albanian dialects is this change of the perfect into a presumptive observed. But as Map 1 shows, it is by far the greater part of the Arbëria from Molise down to northern Calabria.[22]

In the remaining parts of the Arbëria the functions of the perfect changed as well, but in a different way. More precisely, there are two different areas, in which either the temporal functions of the perfect have expanded in such a way that it has replaced the aorist (south-east), or in which the perfect has reduced its functions within the aspectual domain to an experiential perfect (extreme south).[23] Map 2 (based on Google Maps) shows the two isoglosses in Calabria, differentiating the modal (M) presumptive area in the northern part from the temporal (T) area in the south-east and the aspectual area (A) in the extreme south.

The different developments could be connected with differences in the structure of the ARB dialects, but also with linguistic differences between the respective contact varieties. As for internal ARB differences contributing to this

---

22   See Breu (2015: 207) for the map and Altimari (1991) for a survey on the distribution of the presumptive perfect.

23   In the south-eastern area (Province of Crotone in Calabria and Province of Taranto in Apulia), nowadays only the perfect is used, even for describing historical events or in stories and fairytales. In the southern area (southern Calabria and Sicily), on the other hand, even resultative situations require the aorist to the detriment of the perfect, e.g., in Piana degli Albanesi (Sicily): *Kapirta* (AOR.1SG) *atë që do thuash* 'I have understood (now) what you want to say'. See Breu (2015: 211–214) for more details.

MAP 1   The spread of the Italo-Albanian presumptive
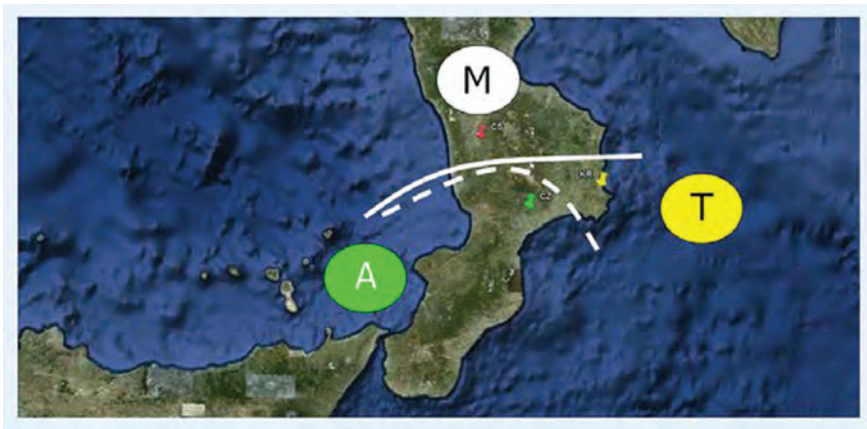
MAP 2   The two isoglosses separating modal, aspectual and temporal perfects

areal tripartition, the most important feature is certainly the formation of the mediopassive, which in the presumptive area is formed with *kam* (preceded by the reflexive particle *u*), just like the active voice, and where consequently

the perfect has developed its presumptive meaning, too, e.g., *u kan njohur* 'they probably have got acquainted' (M). However, where the mediopassive is formed with the auxiliary *jam* 'to be', e.g., *jam njohur* like in Standard Albanian (see below), the transformation of the perfect to a presumptive was completely blocked (T).

An additional factor to be considered is certainly the distribution of perfect and aorist in the Romance contact dialects. In most parts of the ARB presumptive area, complete events in the past can only be expressed by the aorist in Italian. In the extreme south (southern Calabria and Sicily), the aorist even comprises the resultative function, leaving for the perfect only its experiential function. This was the exact model for the reduction of the ARB perfect to an aspectual experiential. In contrast, the Italian dialects in the south-eastern coastal areas, for example in Crotone, seem to have a tendency towards replacing the aorist by the perfect.[24] This could be the reason why in the south-eastern ARB dialects, the perfect (whose development into a presumptive was blocked by the above-mentioned *jam*-perfect in the mediopassive) in the course of the last century has become a past tense grammeme, replacing the aorist.[25]

Interestingly enough, in Calabria the ARB isoglosses dividing the *kam*-perfect into three types seem to coincide also with other Romance isoglosses, especially with the isogloss separating the infinitive area from the southern area, where the infinitive (under Greek influence?) is avoided (Breu, 2015: 228). If this is just a coincidence has still to be investigated.

In any case, language change in the domain of the traditional ARB perfect is completely free of any Balkan heritage. If it is true that the development of a presumptive, i.e., a modal perfect, typologically unites Italo-Albanian with

---

24   For the historical distribution of the perfect (*passato prossimo*) and the aorist (*passato remoto*) in Italy, see Rohlfs (1969: 45–49); modern usage in local Italian is documented by the survey of Bertinetto and Squartini (1996) on the usage of these grammemes in eleven towns throughout Italy, respecting, among other things, various types of texts. Unfortunately, the network of informants in this survey is not fine-grained enough for the Arbëria, omitting for example the Crotone area. But it confirms at least the survival of the dominance of the aorist in the south, though Palermo and especially Cosenza show deviations. As for Crotone, informants confirm the preference of the perfect over the aorist in referring to past events. For the modern situation in Sicily, with influences from the (northern) standard but still a strong preservation of the old distribution of aorist and perfect in informal dialect-based usage, see Alfonzetti (2018).

25   It is worth noting that presumptive dialects are found also in places, where the Romance dialects have replaced the aorist by the perfect, as is the case in the coastal area of Molise. In such cases the most probable explanation is the assumption of a migration of the ARB population from more central territories to these areas in historical times, when they had already developed their presumptive, or, at least, prior to the changes in the corresponding Romance varieties.

those Balkan languages having modal perfects, too, it is also true that this is a bare coincidence and, what is more, valid only with respect to part of the ARB dialects. But just like in the Balkans, where the Bulgarian and the Macedonian renarrative (quotative) goes back to Turkish influence and should be clearly separated from the Albanian admirative, probably also due to Turkish influence in a different setting,[26] the Italo-Albanian presumptive has been induced by language contact based on a totally different model, in this case induced by a non-Balkan language.[27]

5.3   *Verbal Aspect*

In the domain of verbal aspect, beyond the genuine-Slavic derivational opposition of perfectivity expressed by prefixes and suffixes and preserved in Bulgarian and Macedonian, the Balkan verbal systems traditionally had much in common with the Romance opposition between imperfect and aorist in the past tense. On the other hand, analytical aspect forms are traditionally absent in the Balkan *Sprachbund*, which is, so to speak, a negative Balkanism, in which modern Albanian, however, does not participate. Actually, Albanian shows at least two analytical progressives, the gerund construction formed by means of the "auxiliary *jam* 'to be' + converb particle *duke* + participle", e.g., *është duke punuar* 's/he is working', and the construction with the particle *po*, added to the present and to the imperfect tense, e.g., *po punonte* 's/he was working' (Buchholz and Fiedler, 1987: 167–169).

Gerund constructions continue to express simultaneity in Italo-Albanian, but mostly reduced to adverbial phrases without an auxiliary, just like in Italian, often replacing temporal subordinate clauses. The converb particles used differ locally; see example (3) from Frascineto with the particle *ture*. The particle *po* is mainly restricted in ARB to periphrases with the function of an imminentive (see below).

---

26   See Friedman (2004) for both cases. As Jusufi (2016: 144–145) shows for the newly-developed renarrative in an Albanian dialect under Macedonian influence and Friedman (1994) for the rise of an admirative in an Aromanian dialect under Albanian influence, both of them in Macedonia, contact-induced developments in the domain of the traditional perfect are not rare in Balkan languages.

27   I do not agree with Altimari's (1991) hypothesis of the presumptive being a case of archaism, i.e., as the last residual of an originally general Albanian development of the perfect, connected with the admirative in modern Albanian and the renarrative in Balkan Slavic. Contrary to Altimari's opinion, its distribution in Italo-Albanian clearly forms an isogloss, in spite of the scattered settlements, and the principal areal influence of the Romance dialects is obvious. I admit that in some points newer developments inside Romance and the above-mentioned internal migrations of the Arbëreshë have blurred this picture to some extent; see Breu (2015) for more details.

(3) **Ture   mbuluar**,
     PTL   play.PTCP

     u   dojt   hipsha   mbi   tavolinin.
     I   want.PTL climb.IPRF.MPASS.1SG on   table.ACC.SG.DEF
     'When/while playing, I wanted to climb onto the table.'

In contrast, new progressive periphrases have been formed, based on south-ern-Italian models,[28] the most frequent being coordinated or subordinated constructions with the auxiliary *jam* 'to be', for example in northern Calabria and in Molise as in (4a) from San Martino di Finita (Province of Cosenza) and (4b) from Portocannone, but also in Sicily.[29]

(4a) *Mendre* **ësht**   *e*   *ja*   **rrfien**   *pra*
      while   be.PRS.3SG   COP   DAT.3SG:ACC.3SG tell.PRS.3SG   well

      l-urtmu   *ëndren*   *e*   *jëma, ...*
      DEF-last   dream.ACC.SG.DEF CON   mother.NOM.SG.DEF
      'Well, while mother is telling her/him her last dream, …'

(4b) *Ndjeja*   *mirë* *ke*   **ishi**   *e*   **suçëdiri**
      feel.IPRF.1SG well   that   be.IPRF.3SG COP   happen.IPRF.3SG
      ndogjagjë   *straurdhënarju.*
      something   extraordinary
      'I felt well that something extraordinary was happening.'

In Frascineto, a special paradigm of the auxiliary *jam* 'to be' is used in the coor-dinated progressive, conflating **je e* to *je* (PRS.2SG), **jemi e* to *jem e* (PRS.1PL) and **jini e* to *jin e* (PRS.2PL), resulting in the following present-tense paradigm: *jam e hin, je hin, ësht (~osht) e hin, jem e himi, jin e hini, jan e hinjen* 'I am enter-ing' etc. That this is not simply a case of phonetic simplification, but may be claimed a step towards (morphological) grammaticalization, is demonstrated by the fact that apart from the domain of the coordinative conjunction, an *e* does not conflate, e.g., *jam e e bie* 'I am beating him', where the second *e* is the

---

28 Cp. southern-Italian constructions of the type "*stare a* + infinitive" (Rohlfs, 1969: 133).

29 Less frequently than the coordinative conjunction *e*, the relative (and interrogative) pronoun *ç* 'which' has been recorded, for example, in Greci (Campania), e.g., *inja ç hanja* 'I was eating', and in Casalvecchio di Puglia (Apulia), Vena di Maida e S. Nicola dell'Alto (in different parts of Calabria). See Camaj (1971: 60–62) for Greci and the overview in Savoia (1991: 17), which also includes the rare usage of the subjunctive particle *të*. An interesting case is the asyndetic type *inja* (IPRF.1SG) *Ø disnja* (IPRF.1SG) 'I was dying', literally "I was I died", i.e., without any conjunction, recorded in San Martino besides the coordinated progressive in (4b).

clitic pronoun 3SG.ACC 'him/her'. But the most evident sign of grammaticalization is the reduction of the auxiliary to an uninflected particle *ish* in the imperfect, derived from the inflected form of the 3rd person singular, e.g., *ish e prisja* 'I was waiting', *ish e prisjem* 'we were waiting'.[30]

Another way of expressing processes is based on adverbs with the meaning 'now', e.g., *nanthi* in Portocannone, combined predominantly with the present of the verb, e.g., *nanthi zgjohem = jam e zgjohem* 'I am waking up', thus calquing a similar usage of southern-Italian *mo* 'now'.[31]

Besides the different kinds of progressive periphrases, there is an actional periphrasis expressing approaching events (preliminary phase). This so-called "imminentive" may under certain conditions also refer to already on-going processes, but only shortly after their beginning. It is formed in the dialect of Frascineto by means of the fully inflected auxiliary *jam* 'to be' + particle *po* or preposition *pë(r)* + subjunctive, initiated by the subjunctive particle *t(ë)*, e.g., *ësht* (PRS.3SG) *po ~ pë(r) t partirenj* (PRS.SUBJ.3SG) 's/he is about to leave / she is leaving'. At least the prepositional variant is a calque of the Italian model *sta per partire* 's/he is about to leave', etc.[32]

The ARB developments in the domain of aspectuality are clearly independent of any Balkan heritage. But the general tendency of Albanian towards the formation of analytical aspects, untypical for the Balkan *Sprachbund* as a whole, has possibly promoted the acceptance of Romance aspectual constructions by the Italo-Albanian varieties.[33]

---

30   The progressive is observed in other tenses, too, for example in the future, as in *Kat jet e bjer shi* 'It will be raining', mostly understood as an assumption. In this case the future particle *kat* (from *ka* = have.PRS.3SG + subjunctive particle *t*) requires the subjunctive in both verbal parts of the coordinated progressive, *jet* 'be.PRS.SUBJ.3SG' and *bjer* 'fall.PRS.SUBJ.3SG'.

31   A parallel usage is found in Molise Slavic, in this case by adding *sa* 'now', mainly to the present tense forms of the main verb, e.g., *sa parčivam* 'I am leaving' (Breu, 2011: 447), in contrast to the auxiliary-based progressive periphrasis, missing in this micro-language. Note that in neither of the two minority languages the Italian progressive of the type *sto partendo* 'I am leaving' exists, formed with the help of the auxiliary *stare* 'to be, to stay' + gerund, probably due to its absence in the local Romance varieties.

32   But note that in the imminentive, just like in the progressive periphrasis, the auxiliary *jam* 'to be' is used, essentially corresponding to Italian *essere* 'to be' and not *rri* 'to stay, to be somewhere', corresponding to *stare* in Italian. This is different from Molise Slavic, showing a parallel calque of the Italian imminentive, for example, *stoji za partit* 's/he is about to leave', but with the fully corresponding auxiliary *stat* 'to stay, to be somewhere' (and not *bit* 'to be') + *za* 'for' + infinitive. Actually, in place of *jam* 'to be', used in the progressive and the imminentive in Calabria and Molise, the auxiliary *rri* 'to stay' also appears in Italo-Albanian, more precisely in San Marzano di San Giuseppe (Apulia, in contact with Salentino varieties) and in Barile, Ginestra, Maschito (Basilicata); see Savoia (1991: 395–396) and Altimari (2020) for more details.

At least the progressive cannot be separated from the particle constructions to be discussed in the following paragraphs. This follows, for example, from its tendency to form more or less particle-like forms of the auxiliary *jam* 'to be'. The same is true for the auxiliary *kam* in the formation of the future.

5.4 *The Development of a Causative*

The causative as an analytic grammeme is a new development of Italo-Albanian, too. With regard to its formation, the causative construction presents a wide spectre of possibilities, which all have in common the usage of *bënj* 'to make', mostly fossilized as a particle (causative marker), in combination with a linking element and the inflected form of the main verb.[34] In Standard Albanian, causation is expressed by means of the fully inflected verb *bëj* 'to make', governing the main verb in the subjunctive.

Simple examples from Frascineto with intransitive (or intransitively used) main verbs are (5a) and (5b). Here the causative particle *bën*, originally the form of the 3rd singular present, combines with the coordinative conjunction *e* 'and' and the main verb. Variants of *bën* in Frascineto are *bon ~ bin*.[35] The inflected main verb bears all grammatical information, i.e., besides the reference to person, number and mood also tense/aspect characteristics like "present tense" in (5a) und "aorist" in (5b). The agent of the causation is not indicated in the construction proper. It may be expressed optionally, like *u* 'I' in (5a-b), but normally it is concluded from the wider context. In the imperfect of these two sentences the corresponding form would be *bën e pinej* (*~pij;* IPRF.3SG) and *bën e partirnej* (IPRF.3SG), with the same fossilized particle.

---

33   Interestingly, southern Italian models of the type *stare a* 'to stay at', in this case from Salentino, are the bases for Italo-Greek progressive periphrases, too, including the coordinative type (Ledgeway, Schifano and Silvestri, 2018).

34   Linking elements can be the conjunction *e* 'and' (coordination), the subjunctive particle *të* (subordination) or an empty element (asyndetic construction). The main verb may be in the subjunctive or in the indicative, in part depending on the linking element in question. For a comparative description of the causative in several ARB dialects, see Savoia (1989) and Manzini and Savoia (2007).

35   Other forms of the causative particle in different dialects are, for example, *men(ë), bit(ë), mit(ë), pit(ë),* in part containing the particle *t(ë),* requiring the subjunctive of the main verb (Savoia, 1989).

(5a)   (*U*)   **bën**   *e*   **pi**   *Maria.*
       (I.NOM)   CAUSE   COP   drink.IND.PRS.3SG   Maria.NOM.SG.DEF
       'I make (~am making) Maria drink.'

(5b)   (*U*)   **bon**   *e*   **partirti**   *mëma.*
       (I.NOM)   CAUSE   COP   leave:AOR.3SG   mother.NOM.SG.DEF
       'I made mother leave.'

With transitive verbs the subject of the full verb may likewise be put after the
conjunction as in (6a) or at the end of the construction: ~ *U bin e pi një botiljë
ver Maria* 'I made Maria drink a bottle of wine'. A similar example is sentence
(6b) from San Martino, with the nominative *qeni* 'the dog' in three functions:
object of the causation, subject of the main verb in the causative construction
and subject of the coordinated verb *ha* 'to eat'.

(6a)   *U*   **bin**   *e*   *Maria*   **pi**
       I.NOM   CAUSE   COP   Maria.NOM.SG.DEF   drink.IND.PRS.3SG

       një   *botiljë*   *ver.*
       a   bottle   wine
       'I made Maria drink a bottle of wine.'

(6b)   *Ndëse*   *e*   *raxhojim*   **bin**   *e*
       if   her.ACC   enrage.IPRF.1PL   CAUSE   COP

       kalarej   *qeni*   *e*   *na*   *haj.*
       come.down.IPRF.3SG   dog.NOM.SG.DEF.M   COP   us.ACC   eat.IPRF.3SG
       'If we enraged her (=mother), she would make the dog come down
       and he would devour us.'

Even in a dialect, in which the causative particle incorporates the subjunctive
particle *t(ë)*, the main verb in the causative construction may be coordinated
with the causative particle, as for example *bit* (< *bën* + *të*) in Portocannone. But,
interestingly enough, the main verb in this dialect remains in the subjunctive
in spite of being preceded by the coordinative conjunction *e*; see *jet* (be.SBJV.
PRS.3SG) in (7a).

(7a)   *Isht*   *atë*   *çë*   **bit**   *e*   *jet*
       be.IND.PRS.3SG   that   REL   CAUSE   COP   be.SBJV.PRS.3SG

një	*dit*	*dëversu*	*te*	*tjert*	*dita.*
a	day	different	from	other.PL	day.NOM.PL

'It is that what makes a day be different from other days.'

This strange combination of coordination and subordination may also be the reason for the possibility of a variation between nominative and accusative case in the object of the causation (= subject of the main verb), at least for some speakers. Thus in (7b), instead of the definite nominative *prinxhëpi*, the corresponding definite accusative *prinxhëpin* was also claimed to be acceptable.[36]

(7b)	*Idea*	*e*	*murrës*	*elefande*	***bit***	*e*
	idea.NOM.SG.DEF	CON	herd.GEN.SG	elephants	CAUSE	COP

	qeshi	*prinxhëpi*	*i*	*vogël.*
	laugh.AOR.3SG	prince.NOM.SG.DEF	CON	little.

'The idea of a herd of elephants made the little prince laugh.'

Further research on the causative construction of Portocannone seems an important task. At present, it could be classified as an intermediate stage between the original Albanian construction with the fully inflected causative verb *bëj* 'to make', governing the subjunctive and the causative in the Frascineto dialect, in which the causative marker (particle) *bën ~ bin ~ bon* has lost all additional grammatical properties.

There are Italo-Albanian dialects still nearer to the original Albanian construction[37] than that of Portocannone, for example, the ARB dialect of San Nicola dell'Alto in the north-eastern Calabrian Province of Crotone. Turano (1989) gives a detailed description of the causative of this dialect. For the argumentation and the examples presented in the following, see Turano (2005: 24–26).

In San Nicola dell'Alto the causative verb has completely kept its inflection and still governs the full subjunctive, subjunctive particle included, and without any coordinative conjunction. But regarding the syntactic characteristics of the causative constructions it has nevertheless adapted to Italian. While in a Standard Albanian example like (8a) the object of causation *djalin* 'the boy' is in the accusative, doubled by the clitic object marker *e*, and immediately follows the causative verb, the only possibility in the San Nicola dialect

---

36   This example is from the translation of Antoine de Saint-Exupéry's *Le petit prince* into the
      dialect of Portocannone, discussed with informants.

37   See Savoia (1989: 324–329) and Manzini and Savoia (2007: 349–352) for a general comparison
      of the Italo-Albanian and Balkan-Albanian ways of expressing causation.

is (8b). Here the object of causation *jali* is in the nominative (not doubled by the object clitic) and located at the end of the construction, thus functioning as the subject of the subordinate subjunctive clause.

(8a) *Mësuesi*   *e*   *bën*   *djalin*
      teacher.NOM.SG.DEF 3SG.ACC   make.PRS.3SG boy.ACC.SG.DEF
      të   *lexojë*   *librin.*   (Standard Albanian)
      PTL   read.PRS.SUBJ.3SG book.ACC.SG.DEF
      'The teacher makes the boy read the book.'

(8b) *Mënjeshtri*   *bon*   *të*   *ghojirnj*
      teacher.NOM.SG.DEF make.PRS.3SG   PTL   read.PRS.SUBJ.3SG
      ghibrin   *jali.*   (San Nicola dell'Alto)
      book.ACC.SG.DEF   boy.NOM.SG.DEF
      'The teacher makes the boy read the book.'

Actually, a construction similar to (8b) is possible in Standard Albanian, too, but in San Nicola it is the only one. A third type of causative construction found in Standard Albanian, i.e., by means of a complement clause like *Mësuesi bën që djali të lexojë librin* 'The teacher makes the boy read the book'(literally: The teachers makes that the boy reads the book)', introduced by the complementizer *që* 'that' (which would be *çë* in ARB) + subject + subjunctive, is also excluded in San Nicola. On the other hand, just like in the dialect of San Nicola, in Italian there is also only one fully grammatical way of expressing a causation, namely a construction with the infinitive and the object of causation as indirect object (dative): *Il maestro fa leggere il libro al ragazzo.*[38] In view of the typological difference between the two varieties in contact, with the subordination via subjunctive (ARB) vs. infinitive (Italian) and their respective syntactic characteristics, it may be claimed that the causative constructions in question in Italian and ARB exactly match, whereas the additional possibilities of Standard Albanian are excluded in both of them.[39]

Coming back to the causative particle in Frascineto with complete coordination, it is obvious that in this case full grammaticalization took place with

---

38   Here I exclude the only marginally acceptable construction *il maestro fa che legga il libro il ragazzo*, corresponding literally to (8b); see Turano (2005: 25).

39   Actually, the parallelisms do not stop here. So, the Italian construction with the object of causation as an indirect object has been calqued in San Nicola by a corresponding dative and it may even be expressed in both varieties in contact by an agentive prepositional phrase of the type ARB *ka djali,* It. *dal ragazzo* 'by the boy'. Both possibilities are excluded in Standard Albanian. See Turano (2005: 26–29) for other examples and more details.

only one inflected verb (the main verb) and a purely analytical marker of the function CAUSE, leading so to speak to an analytical verb form. In contrast, it could also be argued that the ARB causative shows an exactly inverse (mirrored) image of the Italian causative construction, in both cases with only one inflected verb: in Italian the causative verb, in ARB the main verb. In both languages there is a corresponding uninflected element, the infinitive in Italian and the causative particle in ARB. In other words, ARB shows both: a preservation of the Balkanism of the lack of an infinitive and a calque of the foreign model of inflecting only one of the verbal units in the causative construction. From a point of view of language economy, both languages are on a par, and it seems obvious that this inverse image was the maximum way by which ARB could adapt to Italian, given its lack of an infinitive, on the one hand and the typological characteristics it had as a Balkan language, on the other (Breu, 1994: 379–381).

It is worth noting that the Frascineto particle construction may be coordinated as a whole with inflected *lë* 'to let' as in (9). In this case the causative marker is used more or less pleonastically.

(9) *Na   e   lam   e   **bin**   e   **partirti.***
    we   3SG.ACC let.AOR.1SG COP   CAUSE   COP   leave.AOR.3SG
    'We let her/him leave.' Literally: "We let and made and s/he left."

5.5 *Other Periphrases*

In Italo-Albanian, particle constructions are more widespread than in Standard Albanian. Most modal particles existing in Balkan Albanian have been preserved, for example, *mund* 'can', but not *do* 'will' (for forming the volitive future). On the other hand, new particles have developed, derived in part from the 3rd person singular of the present, besides the causative marker *bën,* used in all tenses, for example also *zë* 'to start', still limited to the present. Another source is the 3rd person singular of the aorist, e.g., *zu* 'to start', as in *u zu e dola* 'I started to go out', corresponding to present-tense *zë*. Others are derived from the imperfect, in addition to the already mentioned particle *ish*, forming the imperfect of the progressive, for example *disht* and (possibly mixed with the present) *do(j)t* 'to want', e.g., *u dojt/disht shkruaja* 'I wanted to write'.

All these particles have been, at least partially, grammaticalized in the sense of forming periphrases in the domains of actionality, irreality or causativity. Like in the case of the progressive, several possibilities for linking them with the main verb exist, coordination with the conjunction *e* 'and', subordination by means of the subjunctive and, less frequently, asyndetic juxtaposition.[40]

A further construction in this domain, worth being mentioned explicitly, is the deontic unreal construction, formed with the particle *kisht*, derived from the 3rd person singular of the imperfect of *kam* 'to have, must'. A typical example is (10), in which the particle combines with the pluperfect, formed with the fully inflected imperfect of the auxiliary *kam* and the participle of the main verb.

(10) *U*   ***kisht***   *u*   *kisha*   *zgjuar*   *mëpar.*
      I.NOM   PTL   REFL   have.IPRF.1SG get.up.PTCP   earlier
      'I should have got up earlier.'

Particle constructions appear also in combination with the coordinated progressive, like *jem e* in (11) with the above-mentioned ingressive particle *zë* 'to start'.

(11) *Na*   *jem*   *e*   ***zë***   *e*   ***dalmi.***
      we   be.PRS.1PL   COP   PTL   COP   exit.PRS.1PL
      'We are starting to leave.'

Just like in the case of the coordinative progressive, not all auxiliaries in periphrases are reduced to particles, with differences from dialect to dialect. For example, in the dialect of San Martino the (abbreviated) example (12) was recorded, showing the inflected auxiliary *zura* (AOR.1SG), instead of the ingressive aorist-based particle *zu*, here in combination with the gerundial progressives *tuke sërritur* and *tuke qar*.

(12) ***Zura***   *tuke*   *sërritur*   *e*   *tuke*   *qar,*
      start.AOR.1SG PTL   scream.PTCP COP   PTL   cry.PTCP

      ika   *atej*   *ku*   *ish*   *mëma.*
      run.AOR.1SG   there   where   be.IPRF.3SG mother.NOM.SG.DET
      'I started crying and shouting, I ran there where mother was.'

Moreover, speakers sometimes insist in a periphrasis instead of an expected simple verb. An example from Portocannone is (13) with the suppletive verb *bie* 'to fall' (aorist *ra* 's/he fell'), where according to the informants adding the aorist of the actional auxiliary *vete* 'to go' is obligatory, in order to give this

---

40   As only the 2nd and 3rd person singular present have a subjunctive form different from the indicative (with the exception of *jam* 'to be' and *kam* 'to have'), it is in many cases undecidable if the main verb is subordinated or linked asyndetically. For a general classification of verbal periphrases in the dialect of Frascineto, see Breu (1994), for their position in a Balkan-orientated typology, see Fiedler (1989). See also Breu (2008) for similar constructions in San Costantino Albanese (Basilicata).

sentence the intended durative meaning, expressed here also by the adverbial *dal e dal* 'slowly'.

(13)   **Vajti**   *e*   **ra**   *dal e dal*
           go.AOR.3SG   COP   fall.AOR.3SG   slow and slow

           si   **vete**   *e*   **bije**   *një*   *lis.*
           how   go.PRS.3SG   COP   fall.PRS.3SG   INDF   tree.NOM.SG
           'He fell slowly like a tree falls.'

Particle constructions are, in principle, typical for the Balkan linguistic area. This is especially true for subordination. In this regard, the new subordinating particle constructions presented in Italo-Albanian could be claimed to be just an extension of an existing type. But as claimed above, this extension was certainly promoted by the model of the infinitive constructions in the Romance contact varieties, in which also only one element is inflected (e.g., the modal verb but not the main verb in the infinitive). It is worth noting that in this domain Italo-Albanian developed its Balkan characteristics still further under the influence of non-Balkan contact varieties, which again relativizes the role of Balkanisms (see also Gardani, Loporcaro, and Giudici, 2021: Section 3).

The causative of the Frascineto type seems to have developed without any direct model. But it obviously is the combination of the two tendencies towards coordination and the spin-off of grammatical information in particles, together with the contact-supported principle of inflecting only one component in modal constructions. As for coordination, it is present in traditional Albanian, too, for example in relatively rare actional constructions (manners of action) like *mori e tha* 's/he suddenly said' (literally: "s/he took and said") and *zu e pi* 's/he started drinking' (literally: "s/he started and drank").[41] So, also in this point a historically given starting point led to further expansion in the contact situation.

5.6   *Recent Developments in Italo-Albanian Grammars*

While the Italo-Albanian developments discussed so far seem to be rather old, as they are found in the oldest documentations of the respective dialects and are normally widely used by the speakers of the older generation, there are also more recent ones, taken from my current fieldwork. Their acceptance varies from speaker to speaker and from dialect to dialect. Unlike the well-integrated

---

41   According to Gjinari et al. (2007: 430), the second example is locally restricted. On the role of coordinated (paratactic) verb forms as a common feature of the Balkan languages, see Sandfeld (1930: 196–199).

older changes, going mainly back to models from local Romance dialects, the younger ones have been calqued from Standard Italian in its colloquial form.

5.6.1   The Imperfect as a Counterfactual Mood

The first development to be mentioned here is the expression of counterfactuality by means of the imperfect, as a result of its extension from a tense/aspect (past imperfective) to the category of mood.

The traditional way of expressing counterfactuality in Standard Albanian is by means of the indicative pluperfect in the protasis and the conditional pluperfect in the apodosis of hypothetical sentences. This corresponds to the Standard-Italian combination of the subjunctive pluperfect (*congiuntivo trapassato*) in the protasis and the conditional pluperfect (*condizionale passato*) in the apodosis, which in colloquial speech may be replaced by the imperfect in both clauses of the counterfactual construction.

An example of the Italo-Albanian adaptation to the counterfactual imperfect of colloquial Italian is (14a).[42] The complete parallelism in the two languages in contact is demonstrated by the same glosses in the second line in (14a) for both the first line in Italo-Albanian and the third line in Italian.

(14a) *Ndë* **vije**   *puru ti,* **shurbejem**   *bashk.*   (ARB)
          if   come.IPRF.2SG also   you work.IPRF.1PL together
          Se   **venivi**   *anche tu,* **lavoravamo**   *insieme.* (Italian)
          'If you had come, too, we would have worked together.'

Example (14b) presents a mixed type of the counterfactual construction, with the colloquial imperfect in the protasis and the traditional pluperfect in the apodosis.

(14b) *Ndë*   *nëng*   **birja**   *qiçin,*
          if   not   lose.IPRF.1SG   key.ACC.SG.DEF

          kishem   **hijtur**   *mbjatu.*
          have.IPRF.1PL   enter.PTCP   at.once

---

42   The conditional pluperfect of Standard Albanian (*do të kisha ardhur* 'I would have come') corresponds to the indicative pluperfect in Italo-Albanian, due to the lack of the particle *do* for the formation of the volitional future, including the future in the past, which serves also as a conditional in Albanian. Therefore the ARB counterfactual hypothetic period traditionally has the pluperfect in both clauses, like *Ndë* **kishe ardhur** *puru ti,* **kishem** **punuar** *bashk*, which is the conservative equivalent to (14a). The corresponding traditional Italian construction is: *Se* **fossi venuto** *anche tu,* **avremmo lavorato** *insieme.*

'If I had not lost the key, we would have entered at once.'

Just like in colloquial Italian, the counterfactual imperfect is used also in single clauses governed by modal verbs like *mund* 'can' in (14c), corresponding to It. *potere*. In this case, the counterfactual imperfect finds in the main verb in ARB, in contrast to Italian, where it appears in the modal verb, as in the Italian translation of (14c): ***Potevano*** (can.IPRF.3PL) *mostrarci* (INF+us) *come si arriva in cima*. The infinitive in Italian and the modal particle in ARB are their uninflected counterparts. This constellation reminds of the inverse (mirrored) correspondences in the causative expressions.

(14c) *Mund*   *na*   ***buthtojen***
       can.PTL   us   show.IPRF.3PL

       si   *arrivohet*   *nd*   *kriet.*
       how   arrive.PRS.MPASS.3SG   in   head.ACC.SG.DET
       'They could have shown us, how to reach the summit.'

The counterfactual imperfect is still avoided by conservative Italo-Albanians, and its frequency varies also in the single dialects. For example, even younger speakers in Portocannone use it only rarely, while it is quite common in Frascineto.

5.6.2   Contact-Induced Change in the Domain of Voice

In the domain of voice, Italo-Albanian has fully adapted to the traditional system of two passives in Italian, the participial passive with the auxiliary *essere* 'to be' and the reflexive passive. Though traditional Albanian also has several possibilities for the expression of states of affairs in the passive voice, there are typological differences with respect to their Italian counterparts. First of all, the two Italian passives are free variants, competing with each other, whereas Albanian has a complementary distribution, as, for example, the passive present and the imperfect are expressed by means of a synthetic medio-passive, while the perfect of the passive is expressed by means of the participial construction.

As for the changes that have occurred in the Italo-Albanian voice system since the immigration, I cannot go into details here, particularly as the passive formation does not belong to the traditional balkanisms. But two important developments, for example in the dialect of Frascineto, should be noted. The first one is the adaptation of the Italo-Albanian medio-passive to the Italian reflexive (dynamic) event passive in rejecting an explicit agent. In the same

way as it is ungrammatical in Italian to say *la casa si (REFL) compra da me 'the house is (being) bought by me', literally: "the house buys itself by me", it is impossible in Italo-Albanian to say *shpia bjehet (PRS.MPASS.3SG) ka u,[43] though in Standard Albanian the corresponding medio-passive expression shtëpia blehet (PRS.MPASS.3SG) nga unë is perfectly acceptable (Buchholz and Fiedler, 1987: 186–187). In contrast, the Italian participial passive of the type la casa è comprata (PTCP.SG.F) da me 'the house is (being) bought by me' is compatible with an explicit agent and may be used both as an event passive and as a stative passive, expressing the result of an action in the past.

Traditionally, the Albanian passive formed by means of "present or imperfect of jam 'to be' + participle" in Albanian only expresses states (Buchholz and Fiedler, 1987: 192–193). This is also the preferred function of this passive in Italo-Albanian in examples like (15) from Frascineto.

(15) *Shpia   je   e   bjejtur   ka* (variant *te*) *u.*
     house.NOM.SG.DET is.PRS.3SG CON.NOM.SG.F buy.PTCP by   me
     'The house is (=has been) bought by me.'

Just like Albanian, ARB uses in its stative passive instead of the invariant participle its adjectival counterpart, by adding an adjective connector, allowing for number and gender agreement, like the feminine-singular connector *e* in (15). This corresponds to the agreement characteristics of the Italian past participle (in both its event and stative function), here *comprata* (PTCP.SG.F). Starting out from this situation, the Italo-Albanian stative passive has extended its function to an event passive in the sense of 'is being bought by me', at least for some speakers. This is a clear adaptation of the semantic structure of this construction to the polysemy of its Italian counterpart. In both passive functions an agent is allowed, again like in Italian.

While in the present the interpretation of the adjectival passive construction as an event passive, based on the Italian model, is still rather restricted in ARB, it has become the only possibility, when referring to past events. In Albanian, traditionally participial (not adjectival) forms with the present and the imperfect of *jam* 'to be' like *është blerë* 'it was bought', *ishte blerë* 'it had

---

43 Most Italo-Albanian dialects have the same restriction with respect to the impossibility of an explicit agent for the medio-passive as Frascineto, with the exception of the central Basilicata dialects like Barile, showing a transition between the traditional usage of the passive constructions and the contact-influenced changes (Turano, 2011: 36). In both contact languages the medio-passive refers to both single and habitual events. But, for example, in the dialect of Frascineto it is possible to disambiguate these two functions by using the copulative periphrasis for current processes like *Kjo shpi **ësht e bjehet*** 'This house is being bought'.

been bought' are forms of the perfect and the pluperfect, alternating with the synthetic mediopassive in the present (*blehet*) or the imperfect (*blehej*) and the reflexive mediopassive, for example, in the aorist (*u ble*). In the Italo-Albanian dialects that have preserved the formation of the passive perfect with *jam* these temporal relations are still valid, but, interestingly, the participle obligatorily transforms into an adjective, preceded by the adjective connector agreeing in gender and number (Turano, 2011: 32–35), thus extending the traditional parallelism between ARB and Italian in the stative present to this environment. In a similar way, in the presumptive area, where the indicative event passive referring to the past is expressed exclusively by the Aorist of *jam*, only the adjectival participle seems to be possible, at least in Frascineto, e.g., *shpia qe* (be.AOR.3SG) *e* (CON.SG.F) *bjejtur* (PTCP) *ka Maria* 'the house was / had been bought by Maria' (Frascineto), with an agreement structure parallel to Italian and contrasting with Standard Albanian.

While these developments in the domain of traditional passives already belong to the very core of Italo-Albanian grammar, there is also an innovation in the passive types themselves, more precisely, the development of a venitive passive as a calque of the Italian participial passive with the auxiliary *venire* 'come'. In this case, a construction like Italian *viene comprata* 'she is (being) bought' (ongoing or habitual event), literally "she comes bought" has an exact counterpart in the ARB venitive passive *vjen e bjejtur*. Like its Romance model – and unlike the medio-passive as the ARB counterpart of the Italian reflexive passive – it allows for an agent; see example (16) from Frascineto. Like in the participial *jam*-passive, the participle is used in its adjectival form, agreeing with the subject in gender and number, just like the passive participle in Italian.

(16) *Ajo*   *bika*   *vinej*   *e*   *mbuluar*
     this.SG.F pile.NOM.SG.DEF come.IPRF.3SG CON.NOM.SG.F cover.PTCP

     ka   *Vinxhendzi.*
     by   Vincenzo.NOM.DEF
     'This pile was (being) covered by Vincenzo.'

As yet, the venitive passive is still not very frequent, and some speakers reject it altogether.[44] But just like in the case of the counterfactual imperfect, they are accepted by less conservative persons, who use them in colloquial speech, as it seems, more freely in Frascineto than in Portocannone. In the descriptions of other dialects it is simply not mentioned.

The introduction of agreement into the event passive is unexpected from a common Albanian point of view. It is a clearly contact-induced Romance feature, distancing Italo-Albanian from its genetic cognates, which is also true for the emergence of a venitive passive.

6   Conclusions

Contact-induced change in Italo-Albanian has in part reduced its Balkan characteristics, as in the cases of the volitional future and the development of aspectual periphrases. But the opposite is also true, as in some cases in which Albanian traditionally did not participate in a specific Balkanism Italo-Albanian now has the feature in question. Examples presented here were, for example, the development of a vocative or a tendency towards the loss of the morphosyntactic differentiation between genitive and dative.

But all such differentiations and adaptations have to be taken as casual, as they simply depend on the structure of Italian as the dominant model for contact-induced change in Italo-Albanian, contrasting by chance with the peculiarities of "Balkan grammar" or matching them. This also generally speaks against simply counting common features of Balkan languages as such to characterize the Balkan linguistic area, but in favor of explaining every single Balkanism in terms of the mutual influences of the languages on the Balkan peninsula. The same is true for grammatical developments in Italo-Albanian that have to be investigated individually, in order to determine the possible role of foreign models in each single case. If such developments lead to a "de-Balkanization" of Italo-Albanian is a rather secondary question.

Two other criteria are of central importance for contact linguistics, namely, the preservation of a traditional Albanian structure (be it a Balkanism or not) against the model of the dominant varieties and contact-induced change by adapting to a foreign model, in spite of the resistance of the inherited Albanian structure. Among its resistant characteristics Italo-Albanian shows such typical Albanian features like the postponed definite article or the declension of nouns, but also the lack of a (morphological) infinitive. In contrast, for instance, the morphosyntax of the passive has adapted to a large degree to

---

44   In Molise Slavic, also banning the addition of an agent in the reflexive passive, the situation is in many respects parallel to ARB, though in the case of its perfective imperfect even older speakers use the newly developed venitive passive. For a detailed description of the Molise-Slavic passives, compared with their Italian counterparts, see Breu and Makarova (2019).

the Romance type, and the imperfect has developed a counterfactual function, based on a model of colloquial Italian.

However, in many cases there are grammatical developments in Italo-Albanian not contradicting neither the Albanian language type nor the specific properties of its contact varieties, but simply uniting them to a new compatible structure, different from both. An excellent example of this type is the development of a causative, based on the Albanian heritage of forming grammatical particles and on the Romance property of having only one inflected verb in modal constructions. As Italo-Albanian has been influenced by different Italian varieties, including different local dialects, the results of such common developments may form isoglosses within the Italo-Albanian linguistic area, as in the case of the perfect, having developed into three different types, depending on both inherited differences and differences in the Romance models.

Abbreviations

Glosses follow the Leipzig Glossing Rules when applicable.
ABL ablative; ACC accusative; AOR aorist; AUX auxiliary; CON connector; COP copula; DAT dative; DEF definite; F feminine; GEN genitive; IND indicative; INDF indefinite; IPRF imperfect; LOC locative; M masculine; MPASS mediopassive; NOM nominative; REFL reflexive; PFV perfective; PL plural; PRS present; PTCP participle; PTL particle; SG singular; SUBJ subjunctive; VOC vocative.

References

Alfonzetti, Giovanna. 2018. Usi e funzioni del passato remoto nell'italiano di Sicilia. In Annamaria Chilà and Alessandro De Angelis (eds.), *Capitoli di morfosintassi delle varietà romanze d'Italia: teoria e dati empirici*, 17–30. Palermo: Centro di studi filologici e linguistici siciliani.

Altimari, Francesco. 1991. Alcune osservazioni sul passato «presuntivo» dell'arbëresh. In Addolorata Landi and Patrizia Del Puente (eds.), *La lingua albanese nell'Italia meridionale. Studi e prospettive*, 53–65. Salerno: Università degli Studi di Salerno.

Altimari, Francesco. 2005. Il 'futuro necessitativo' dell'albanese d'Italia: Influenza italo-romanza o arcaismo balcanico? In Walter Breu (ed.), *L'influsso dell'italiano sul sistema del verbo delle lingue minoritarie*, 1–12. Rende: Università della Calabria.

Altimari, Francesco. 2009. Tracce di antichi infiniti dell'albanese nell'arbëresh d'Italia. In Lenka Scholze and Björn Wiemer (eds.), *Von Zuständen, Dynamik und*

Veränderung bei Pygmäen und Giganten. Festschrift für Walter Breu zu seinem 60. Geburtstag, 157–175. Bochum: Brockmeyer.

Altimari, Francesco. 2020. Ballkanizma dhe italo-romanizma në disa ndërtime foljore aspektore të arbërishtes së Italisë [Balkanisms and italo-romanisms in some verb constructions of the Arbërisht of Italy]. In Rexhep Ismajli (ed.), *Studimet albanistikë në Itali. Albanistic studies in Italy*, 69–82. Prishtina: ASHAK.

Bartl, Peter. 1981. Fasi e modi dell'immigrazione albanese in Italia. In Pier Fausto Palumbo (ed.), *Congressi sulle relazioni tra le due Sponde adriatiche 2, I rapporti demografici e popolativi*, 199–212. Rome: Centro di studi sulla storia e la civiltà adriatica.

Beci, Bahri. 2012. Konsiderata të përgjithshme për rolin e shqipes në formimin e bashkësisë gjuhësore ballkanike [General considerations on the role of Albanian in the formation of the Balkan linguistic area]. In Rexhep Ismajli (ed.), *Shqipja dhe gjuhët e Ballkanit* [Albanian and the Balkan languages], 145–161. Prishtina: Akademia e Shkencavet.

Bertinetto, Pier Marco and Mario Squartini. 1996. La distribuzione del Perfetto Semplice e del Perfetto Composto nelle diverse varietà di italiano. *Romance Philology* 49(4): 383–419.

Breu, Walter. 1991a. Das italoalbanische 'Perfekt' in sprachvergleichender Sicht. In Francesco Altimari, Gabriele Birken-Silverman, Martin Camaj, and Rupprecht Rohr (eds.), *Atti del Congresso Internazionale di Studi sulla Lingua, la Storia e la Cultura degli Albanesi d'Italia* (*Mannheim, 25–26 giugno 1987*), 51–66. Rende: Università della Calabria.

Breu, Walter. 1991b. Zur aktuellen Situation in den nördlicheren italoalbanischen Kolonien. In Walter Breu, Hans-Jürgen Sasse and Rolf Ködderitzsch (eds.), *Aspekte der Albanologie*, 1–16. Berlin: Harrassowitz.

Breu, Walter. 1994. Forme verbali perifrastiche arbërisht. In Francesco Altimari and Leonardo M. Savoia (eds.), *I dialetti italo-albanesi. Studi linguistici e storico-culturali sulle comunità arbëreshe*, 365–385. Rome: Bulzoni. [Palermo 1982[1]].

Breu, Walter. 2008. *Aspetto verbale ed aspettualità nel dialetto italoalbanese di San Costantino Albanese*. In Francesco Altimari and Emilia Conforti (eds.), *Omaggio a Girolamo de Rada: Atti del V Seminario Internazionale di Studi Italo-Albanesi*, 93–110. Rende: Università della Calabria.

Breu, Walter. 2011. Language contact of minority languages in Central and Southern Europe: a comparative approach. In Bernd Kortmann and Johan van der Auwera (eds.), *The Languages and Linguistics of Europe. A Comprehensive Guide*, 429–452. Berlin: De Gruyter Mouton.

Breu, Walter. 2015. Der Präsumptiv im Arbrisht: Formen und Funktionen, Entwicklung und grammatische Klassifikation. In Bardhyll Demiraj (ed.), *Sprache und Kultur der Albaner. Zeitliche und räumliche Dimensionen*, 205–231. Wiesbaden: Harrassowitz.

Breu, Walter. 2018a. La situazione linguistica nei paesi *arbëreshë* del Molise. In Lucija Šimičić , Ivana Škevin and Nikola Vuletić (eds.), *Le isole linguistiche dell'Adriatico*, 169–197. Canterano: Aracne Editrice.

Breu, Walter. 2018b. On the influence of Italian on the Grammar of Molise Slavic and Italo-Albanian. In Thede Kahl , Iliana Krapova and Giuseppina Turano (eds.), *Balkan and South Slavic Enclaves in Italy: Languages, Dialects and Identities*, 215–236. Newcastle upon Tyne: Cambridge Scholars Publishing.

Breu, Walter. 2018c. Das Italoalbanische zwischen balkanischem Erbe und kontaktbedingter Entwicklung. In Francesco Altimari , Gëzim Gurga and Shaban Sinani (eds.), *Matteo Mandalà dhe albanologjia sot. Në 60-vjetorin e lindjes*, 77–96. Tirana: Fast Print.

Breu, Walter. 2019. Morphosyntactic changes in Slavic micro-languages: The case of Molise Slavic in total language contact. In Andrii Danylenko and Motoki Nomachi (eds.), *Slavic on the Language Map of Europe*, 385–432. Berlin: De Gruyter Mouton.

Breu, Walter. forthcoming. Semantische Strukturanpassung im romanisch-slavischen Sprachkontakt. In Malinka Pila and Pino Marco Pizzo (eds.), *Transferkonzepte 2017* (= *Arbeiten und Texte zur Slavistik*). Mainz: Franke und Timme.

Breu, Walter and Elvira Glaser. 1979. Zur sprachlichen Situation in einer italo-albanischen Gemeinde. *Münchner Zeitschrift für Balkankunde* 2: 19–50.

Breu, Walter and Anastasia Makarova. 2019. Typologie des Passivs im Moliseslavischen: Bewahrung, Umbau und Innovation im totalen slavisch-romanischen Sprachkontakt. In Bernhard Brehmer and Holger Kuße (eds.), *Slavistische Linguistik 2016/2017*. *Wiener Slawistischer Almanach* 83: 7–60. Frankfurt am Main: Lang.

Breu, Walter and Malinka Pila. 2018. Vzaimodejstvie buduščego vremeni s glagol'nym vidom v molizsko-slavjanskom i rez'janskom mikrojazykax. In Tatiana Milliaressi (ed.), *La relation temps/aspect: approches typologique et contrastive*, 129–138. Lille: Éditions du Conseil Scientifique de l'Université de Lille.

Buchholz, Oda and Wilfried Fiedler. 1987. *Albanische Grammatik*. Leipzig: VEB Verlag Enzyklopädie.

Camaj, Martin. 1971. *La parlata albanese di Greci in provincia di Avellino*. Florence: Olschi.

Fiedler, Wilfried. 1989. Zum Verhältnis von arealer Linguistik und Arealtypologie (am Beispiel der Balkansprachen). *Zeitschrift für Phonetik und sprachliche Kommunikations forschung* (*ZPSK*) 42: 304–321.

Fiedler, Wilfried. 1992. Ist das Albanische, „dem Typus nach", die entscheidende Sprache des ganzen balkanischen Gebietes? *Albanica: A Journal of Albanological Research and Criticism* 3–4: 78–95.

Friedman, Victor A. 1994. Surprise! Surprise! Arumanian has had an admirative. *Indiana Slavic Studies* 7: 79–89.

Friedman, Victor. 2004. The typology of Balkan evidentiality and areal linguistics. In Olga M. Tomić (ed.), *Balkan Syntax and Semantics*, 101–134. Amsterdam: Benjamins.

Friedman, Victor A. 2006. Balkans as a linguistic area. In Keith Brown (ed.), *Encyclopedia of Language and Linguistics*, volume I, 657–672. Oxford: Elsevier.

Gardani, Francesco, Michele Loporcaro and Alberto Giudici. 2021. In and around the Balkans: Romance languages and the making of layered languages. *Journal of Language Contact* 14(1): 1–23.

Gjinari, Jorgji, Bahri Beci, Gjovalin Shkurtaj, and Xheladin Gosturani. 2007. *Atlasi dialektologjik i gjuhës shqipe. Vëllimi I* [Dialectological atlas of the Albanian language. Volume I]. Naples: Università degli Studi di Napoli L'Orientale.

Hinrichs, Uwe. 1999. Die sogenannten Balkanismen als Problem der Südosteuropa-Linguistik und der allgemeinen Sprachwissenschaft. In Uwe Hinrichs (ed.), *Handbuch der Südosteuropa-Linguistik*, 429–462. Wiesbaden: Harrassowitz.

Johanson, Lars. 2002. Contact-induced change in a code-copying framework. In Mari C. Jones and Edith Esch (eds.), *Language Change: The Interplay of Internal, External, and Extra-linguistic Factors*, 285–314. Berlin: Mouton De Gruyter.

Jusufi, Ljumnije. 2016. *Zajazi dhe e folmja e tij* [Zajazi and its Dialect]. Prishtina: Om.

Kahl, Thede, Michael Metzeltin and Helmut Schaller (eds.). 2012. *Balkanismen heute – Balkanisms Today – Balkanizmy segodnja*. Vienna: LIT Verlag.

Lambertz, Maximilian. 1914. Albanische Mundarten in Italien (Mit einer Kartenskizze). *Indogermanisches Jahrbuch* 1913(1): 1–30.

Ledgeway, Adam, Norma Schifano and Giuseppina Silvestri. 2018. The expression of progressive aspect in Grico: Mapping morphosyntactic isoglosses in an endangered Italo-Greek variety. *Transactions of the Philological Society* 116(2): 179–217.

Mandalà, Matteo. 2005. Siculo-arbëresh e siciliano a contatto: alcune verifiche. In Walter Breu (ed.), *L'influsso dell'italiano sulla grammatica delle lingue minoritarie. Problemi morfologici e sintattici*, 13–21. Rende: Università della Calabria.

Mandalà, Matteo. 2012. Mbi turqismat në të folmet arbëreshe [On turkisms in Italo-Albanian dialects]. In Rexhep Ismajli (ed.), *Shqipja dhe gjuhët e Ballkanit* [Albanian and the Balkan Languages], 277–287. Prishtina: Akademia e Shkencavet.

Manzini, M. Rita and Leonardo M. Savoia. 2007. Causatives in Albanian. In M. Rita Manzini and Leonardo M. Savoia (eds.), *A Unification of Morphology and Syntax. Investigations into Romance and Albanian dialects*, 335–385. London: Routledge.

Manzini, M. Rita and Leonardo M. Savoia. 2015. Notes on the contact between Italo-Albanian and Romance (Calabrian, Lucanian) varieties: Borrowings, code-mixing and convergence. *Hylli i Dritës* 35(1–2): 92–116.

Pignoli, Maria Luisa and Guido Tartaglione. 2007. *Dizionario albanese molisano* (*Parlate di Portocannone e Ururi*). Rende: Università della Calabria.

Rohlfs, Gerhard. 1969. *Grammatica storica della lingua italiana e dei suoi dialetti. Sintassi e formazione delle parole*. Turin: Einaudi.

Sandfeld, Kristian. 1930. *Linguistique balkanique. Problèmes et résultats*. Paris: Champion.

Sasse, Hans-Jürgen. 1991. *Arvanitika: Die albanischen Sprachreste in Griechenland. Teil 1*. Wiesbaden: Harrassowitz.

Savoia, Leonardo M. 1989. Alcune caratteristiche del causativo arbëresh. In Pietro Manali (ed.), *Le minoranze etniche e linguistiche. Atti del 2° Congresso Internazionale. I*, 363–420. Piana degli Albanesi: Comune.

Savoia, Leonardo M. 1991. Alcuni elementi per una classificazione dei dialetti arbëreshë. In Addolorata Landi and Patrizia Del Puente (eds.), *La lingua albanese nell'Italia meridionale*, 13–52. Salerno: Università degli Studi di Salerno.

Savoia, Leonardo M. 2008. *Studi sulle varietà arbëreshe, con la prefazione di Francesco Altimari e un contributo di M. Rita Manzini*. Rende: Università della Calabria.

Savoia, Leonardo M. 2010. Albanese, comunità. In Raffaele Simone (ed.), *Enciclopedia dell'italiano*, 33–37. Rome: Istituto dell'Enciclopedia Italiana.

Sims, Andra D. 2008. Review of: Tomić (2006). *Journal of Slavic Linguistics* 16(2): 331–347.

Solta, Georg R. 1980. *Einführung in die Balkanlinguistik mit besonderer Berücksichtigung des Substrats und des Balkanlateinischen*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Tomić, Olga M. 2006. *Balkan Sprachbund. Morphosyntactic features*. Dordrecht: Springer.

Trumper, John. 2020. The role of albanoid groups (Albanian and Italoalbanian) in language transmission (Italy and Calabria as exemplum). In Rexhep Ismajli (ed.), *Studimet albanistikë në Itali. Albanistic studies in Italy*, 101–108. Prishtina: ASHAK.

Turano, Giuseppina. 1989. *La costruzione causativa nell'albanese di S. Nicola dell'Alto. Una applicazione della teoria 'Government and Binding'*. Tesi di laurea, Università di Firenze.

Turano, Giuseppina. 2005. L'influsso dell'italiano sulla sintassi dell'arbëresh. Il caso di S. Nicola dell'Alto in provincia di Crotone. In Walter Breu (ed.), *L'influsso dell'italiano sulla grammatica delle lingue minoritarie. Problemi di morfologia e sintassi*, 23–41. Rende: Università della Calabria.

Turano, Giuseppina. 2011. Continuità e innovazione nella morfosintassi dell'arbëresh. La realizzazione della morfologia verbale Non-attiva. In Walter Breu (ed.), *L'influsso dell'italiano sul sistema del verbo delle lingue minoritarie. Resistenza e mutamento nella morfologia e nella sintassi*, 17–39. Bochum: Brockmeyer.

Weigand, Gustav. 1913. *Albanesische Grammatik im südgegischen Dialekt* (*Durazzo, Elbassan, Tirana*). Leipzig: Johann Abrosius Barth.

The Negative Imperative in Southern Calabria.
Spirito Greco, Materia Romanza Again?

Adam Ledgeway
Professor of Italian and Romance Linguistics, Faculty of Modern
and Medieval Languages and Linguistics, University of Cambridge,
Cambridge, UK
anl21@cam.ac.uk

Norma Schifano
Lecturer in Modern Languages, Department of Modern Languages,
University of Birmingham, Birmingham, UK
n.schifano@bham.ac.uk

Giuseppina Silvestri
Lecturer in Linguistics, Department of Linguistics, University of California,
Los Angeles, USA
g.silvestri@ucla.edu

Abstract

The aim of this article is to investigate a special case of suppletion in the paradigm of the negative imperative in some dialects of southern Calabria. First, we show how these paradigms involve the extension of an original infinitival desinence to a present indicative verb, giving rise to a hybrid imperatival form (Section 2). Second, we claim that this pattern of suppletion does not represent a Romance-internal development but, rather, the outcome of contact-induced change and, in particular, the influence of the local Greek sub-/adstrate (Section 3). Furthermore, we show that these hybrid patterns also provide significant evidence for the formal morphosyntactic equivalence between competing Greek finite and Romance non-finite forms of subordination, a typical Balkanism (Section 4). Finally, we demonstrate that the extension of the Romance infinitival desinence according to an underlying Greek model yields in synchrony an alternation between a suppletive positive imperative and a true negative imperative, a typologically very rare formal opposition (Section 5).

Keywords

negative imperative – suppletion – infinitive – Italo-Greek – Calabrian – language contact

1 Introduction

Much attention has been devoted in the literature to the study of imperatival morphology in Romance, and in particular the distribution and nature of true vs. suppletive forms exhibited by positive and negative paradigms,[1] as exemplified in Table 1.

TABLE 1   Some canonical Romance imperative paradigms[2]

		Positive	Negative
Piedmontese	2SG	true	true[3]
	2PL	[no data]	[no data]
modern central Occitan	2SG	true	suppletive [=SBJV]
	2PL	[no data]	[no data]
Spanish, Sardinian	2SG	true	suppletive [=SBJV]
	2PL	true	suppletive [=SBJV]
French	2SG	(true)	(true)
	2PL	suppletive [=IND]	suppletive [=IND]
Catalan	2SG	true	suppletive [=SBJV]
	2PL	suppletive [=IND]	suppletive [=IND]
Italian	2SG	true	suppletive [=INF]
	2PL	suppletive [=IND]	suppletive [=IND]
Romanian	2SG	true	suppletive [=INF][4]
	2PL	suppletive [=IND]	suppletive [=IND]

1   See Zanuttini (1994; 1997: Section 4.3), Rivero (1994a; 1994b), Rivero and Terzi (1995), Silva-Villar (1998), Poletto and Zanuttini (2003), Portner and Zanuttini (2003), Manzini and Savoia (2005: 389–487), Maiden (2006), Ionescu (2019), a.o. For further differences exhibited by positive vs. negative imperatives cross-linguistically, see Han (1999; 2001), Tomić (1999), Aikhenvald (2008: 18–20).

2   See Rivero (1994b: 91–92), Zanuttini (1997: 109–113), Manzini and Savoia (2005: 461–467), Maiden (2006).

3   The same true > true pattern can be found in Latin for both persons (e.g., (nē) credite (NEG) imagine.IMP.2PL '(don't) imagine!', later followed by several other strategies, including true > suppletive (e.g., *lauda* praise.IMP.2SG 'praise!', *noli laudare* NEG praise.INF 'don't praise!'). See Ionescu (2019: Section 4) for an overview.

4   The infinitive is also attested in old French and in Romansch (Rohlfs, 1968: 356; Tekavčić, 1972: 417), and across numerous Italo-Romance varieties (cf. discussion in Section 4). For

Following Zanuttini (1997: 105), true imperatives are "verbal forms that are unique to the paradigm of the imperative, in the sense that they are different from any other verbal form used for the same person in any other verbal paradigm", whereas this is not the case with suppletive (or surrogate; cf. Rivero, 1994a) imperatives which involve "verbal forms that are used in the imperative but are morphologically identical to a form used for the same person in another paradigm" such as subjunctives or infinitives (cf. also Rivero, 1994a: 103; Zanuttini, 1994: 119; Isac, 2015: chs 2, 9–10). As Table 1 illustrates, suppletive forms prove particularly common in the negative paradigm, a distribution frequently interpreted as a consequence of the so-called "negative imperative puzzle" (Alcázar and Saltarelli, 2014: Section 2.6.1),[5] which highlights the fact that true imperatives typically prove incompatible with (preverbal) negators.[6] Formally, this empirical generalization has been explained in terms of an intervention effect of the preverbal negator: assuming true imperatives are licensed by raising to C° (Rivero, 1994b; Graffi, 1996), this movement is blocked by the intervening preverbal negator or by the negator itself lexicalizing the C position (see further the discussion in Section 4).

Focusing specifically on Italo-Romance, we can identify following Manzini and Savoia (2005: 389ff.) five distinct suppletive types for the 2SG negative imperative, as summarized in Table 2.[7] Type A (cf. 1b) involves the use of the 3sg indicative, which also functions as the 2sg positive imperative (cf. 1a), whereas in Types B-E positive and negative imperatives display distinct forms. In Type B the negative imperative is expressed by the morphological infinitive (2b), in Type C by the imperfect (3b) or present (4b) subjunctive, in Type D by the auxiliary STAND followed by the lexical infinitive (5b), and in Type E by the gerund optionally introduced by an infinitival auxiliary BE or GO (6b).

---

    further discussion of the suppletive and novel true uses of (the long form of) the infinitive in Romanian, see the discussion around examples (28)-(29) in Section 5.

5 Cf. also Han (1998; 2001), Zeijlstra (2006), Manzini and Savoia (2005: Section 7.2), Cavalcante (2011).

6 As correctly pointed out by an anonymous reviewer, a related point is the fact that imperatives often feature a different negator than the one used with indicative verbs, not only in cases of subjunctive suppletion for the negative paradigm, where the same irrealis negator is carried over to the imperative (e.g., standard modern Greek), but also in languages featuring a true negative imperative (e.g., the second person singular in Ancient Greek).

7 We gloss over the further internal classification they make in terms of clitic placement and type/position of negation, as well as differences between indicative and the positive imperative forms. We take the second person singular as the model to classify true vs. suppletive negative forms, since cross-linguistically suppletive forms seem to enter paradigms through this particular person first (cf. WALS, map 70A, where only 2 languages out of 547 have true imperative forms for the second person plural but not for the second person singular; http://wals.info/feature/70A#2/19.3/148.4).

TABLE 2   Italo-Romance synthetic and periphrastic suppletive forms for 2SG negative
imperative

Synthetic			Periphrastic	
Type A NEG+IND	Type B NEG+INF	Type C NEG+PRS/ IPFV.SBJV	Type D NEG+*stare* (*a*) 'stand (to)'+INF	Type E NEG+(*essere/ andare* 'be/go') +GER

(1) a. *'klɔma*   *el.*   (Donat; Manzini and
         call.PRS.IND.3SG [also IMP.2SG]   him   Savoia, 2005: 446)
         'He calls him. / Call him!'

     b. *'bɪtʃa*   *'klɔma*   *el.*
         NEG   call.PRS.IND.3SG [also IMP.2SG]   him
         'Don't call him!'

(2) a. *'kɛrda*   *to*   *'fre.*   (La Pli de Mareo; Manzini and Savoia,
         call.IMP.2SG your   brother 2005: 390)
         'Call your brother!'

     b. *no*   (*pa*)   *le=*   *kɛr'de.*
         NEG   PRT   him=   call.INF
         'Don't call him!'

(3) a. *ʻcama*   *tu*   *'paðre.* (Avigliano Umbro; Manzini and Savoia,
         call.IMP.2sg   your   father   2005: 462)
         'Call your father!'

     b. *nu*   *llo=*   *ca'massi.*
         NEG   him=   call.IPFV.SBJV.2SG
         'Don't call him!'

(4) a. *'mutti=li.*   (Dorgali; Manzini and Savoia, 2005: 463)
         call.IMP.2SG=him
         'Call him!'

b.   *nɔn=lu*   *'muttaca.*
     NEG=him   call.PRS.SBJV.2SG
     'Don't call him!'

(5)   a.   *'tʃama*   *to*   *'frelu.*   (Calizzano; Manzini and Savoia,
           call.IMP.2SG   your   brother   2005: 451)
           'Call your brother!'

      b.   *nɛ*   *s'ta=lu*   *a*   *tʃa'mɔ.*
           NEG   stand.PRS.IND.2SG=him   to   call.INF
           'Don't call him!'

(6)   a.   *cama='jullə.*   (Miglionico; Manzini and Savoia, 2005: 456)
           call.IMP.2SG=him/her/them
           'Call him/her/them!'

      b.   *na*   *lə=*   *ca'mannə.*
           NEG   him/her/them=   call.GER
           'Don't call him/her/them!'

Within this scenario we consider below a special case of suppletion in the paradigm of the negative imperative in some dialects of southern Calabria (Section 2). The relevant paradigms are special in several key respects. First, they involve the extension of an original infinitival desinence to a present indicative verb, giving rise to a hybrid imperatival form which exceptionally marries together finite and non-finite inflexional markings. Second, the synchronic comparison of these southern Calabrian varieties allows us a rare opportunity to reconstruct in diachrony the emergence and extension of the relevant suppletive pattern across different persons. Third, the patterns of suppletion in southern Calabria just outlined do not represent a Romance-internal development but, rather, the outcome of contact-induced change and, in particular, the influence of the local Greek sub-/adstrate on the surrounding Romance varieties (Section 3). The resultant system of formal paradigmatic oppositions thus reproduces an underlying Greek model, not a Romance one, giving rise to a case of what Rohlfs famously termed *spirito greco, materia romanza* ('Greek spirit, Romance material'). At the same time, these Greek-Romance hybrid patterns also provide significant evidence for the formal morphosyntactic equivalence between competing Greek finite and Romance non-finite forms of subordination, a typical Balkan feature, inasmuch as extension of the infinitival desinence never penetrates those imperatival forms introduced by a Greek-style modal subordinator (Section 4). Finally, the extension of the Romance

infinitival desinence according to an underlying Greek model, although initially a manifestation of a suppletive pattern, will be shown to yield in synchrony a novel true imperatival pattern and, in turn, an alternation between a suppletive positive imperative and a true negative imperative, a typologically very rare formal opposition (Section 5).

2   Negative Imperative in Southern Calabria

2.1   *Mosorrofa*

Loporcaro (1995) identifies a unique pattern in the negative paradigm of the extreme southern Italian dialect of Mosorrofa (province of Reggio Calabria). Drawing on data from Crucitti (1988), he reports a peculiar inflexional morphology for the first and second person plural, both featuring an (optional) *-ri* ending. The relevant data set is exemplified in Table 3 with the first-conjugation verb *parrari* 'speak' (adapted from Loporcaro, 1995: 349).

While the p(ositive) paradigm displays two patterns, namely one which is syncretic with the indicative (pattern p(a): 2SG,[8] 1/2PL) and one which exhibits the irrealis modal particle *mi* plus the present indicative (pattern p(b): 3SG/PL), the n(egative) paradigm operates three patterns, namely one which is syncretic with the indicative (partial pattern n(a_1): 1/2PL without *-ri*), the *mi*-form (pattern n(b): 3SG/PL), and one with the infinitive (pattern n(c): 2SG, and optionally pattern n(c_2): 1/2PL).[9] To explain the optionality of the *-ri* suffix with the first and second persons plural, Loporcaro (1995) correctly argues that *-ri* should be analysed as the erstwhile infinitival ending (cf. *parr-a-**ri*** 'stem-thematic.vowel-INF') which has spread to the first and second persons plural via analogical extension from the second person singular because of paradigmatic pressure to restore the structural symmetry between the positive and negative paradigms. In particular, by extending *-ri* to the first and second persons plural, a paradigm with two patterns (cf. (n(c_2): 2SG, 1/2PL vs. n(b): 3SG/PL) is restored

---

8   Note that the second singular of first-conjugation verbs is syncretic with the third singular of the present indicative.

9   Rohlfs (1968: 355) reports negation + irrealis modal particle also for the second person singular and plural in southern Calabria (cf. also discussion in Section 4):

(i)   *Nommu*   *cadi/caditi!*
       NEG.that$_{\text{IRREALIS}}$ fall.PRS.IND.2SG/PL
       'Don't fall!'

TABLE 3   Imperative in Mosorrofa

	Positive	Pattern	Negative	Pattern
2SG	*parra*	a	*non parra**ri***	c
3SG	*mi parra*	b	*non mi parra*	b
1PL	*parramu*	a	*non parramu(**ri**)*	a_1 (c_2)
2PL	*parrati*	a	*non parrati(**ri**)*	a_1 (c_2)
3PL	*mi parrinu*	b	*non mi parrinu*	b

for the negative imperative exactly mirroring that of the positive imperative (cf. two p(a) patterns: 2SG, 1/2PL vs. p(b): 3SG/PL). According to Loporcaro, this extension takes place in the negative imperative because this is the only finite paradigm to include an infinitive. Moreover, it is attested exclusively in these dialects (but not in other Romance varieties which also exhibit the infinitive in the second person singular) because of their well-known alternation (and functional equivalence; see Ledgeway, 1998, cf. also De Angelis, 2013: 2 n.2) between finite *mi*-clauses and infinitival (viz. *-ri*) clauses in subordinate contexts.[10]

In short, Loporcaro's analysis presupposes a Romance-internal development, according to which extension of erstwhile infinitival *-ri* is driven by a system-internal symmetry to establish and transfer the distributional p(a) pattern of the positive imperative to the corresponding cells of the negative imperative paradigm, viz. the n(c_2) pattern. The role of contact is indirectly acknowledged in accounting for the lack of analogous extensions in other Romance varieties which exhibit the infinitive in the second person singular, insofar as these lack the *mi*-clauses/infinitive alternation which is only attested in these areas of historical Greek substrate. In the next section, however, we shall see on the basis of new data from a wider selection of southern Calabrian dialects that the role played by contact in this syntactic domain is more direct and is in fact the main trigger for the observed remodelling of the relevant paradigms.

2.2   *Cardeto and Gallicianò*

Our recent fieldwork investigations in southern Calabria confirm the extension of *-ri* in Mosorrofa, but also reveal that the same pattern is attested in Cardeto and Gallicianò (province of Reggio Calabria).[11] Significantly, however, our

---

10   For an overview and relevant bibliography, see Ledgeway (2016: 1018–1019, 1023–1027).

11   All examples from our own fieldwork appear unmarked, whereas examples from other sources are marked as such. In citing Romance data from our fieldwork we use a very

TABLE 4   Imperative in Mosorrofa, Cardeto and Gallicianò Calabrian

Mosorrofa, Cardeto, Gallicianò Calabrian		Gallicianò Calabrian		Mosorrofa, Cardeto	
Positive	**Pat.**	**Negative**	**Pat.**	**Negative**	**Pat.**
2SG	= IND.3SG (1 conj.) = IND.2SG (2/3 conj.)	a	*non*+INF-*ri*	$c_{(1)}$	*non*+INF-*ri*	$c_{(2)}$
3SG	*mi*+IND.3SG	b	*non*+*mi*+IND.3SG	b	*non*+*mi*+IND.3SG	b
1PL	= IND.1PL	a	*non*+IND.1PL	a_1 (d)	*non*+IND.1PL (+-*ri*)	a_1 (c_2)
2PL	= IND.2PL	a	*non*+IND.2PL (+-*ri*)	a_1 (c_1)	*non*+IND.2PL (+-*ri*)	a_1 (c_2)
3PL	*mi*+IND.3PL	b	*non*+*mi*+IND.3PL	b	*non*+*mi*+IND.3PL	b

investigations bring to light for the Calabrian dialect of Gallicianò a paradigm for the negative imperative which, to date, has gone unnoticed. Specifically, in this dialect *-ri* extends optionally to the second person plural but never to the first person plural, witness the selection of examples in (7)-(9) (see also Ledgeway Schifano and Silvestri, in prep.: ch. 3). Table 4 offers an overview of the attested patterns.

(7)   Positive (Mosorrofa, Cardeto, Gallicianò Calabrian)[12]
    a. *Danci*   *u*   *libbru*   *a*   *Mmaria!* (Cardeto)
       give.PRS.IND.3SG=DAT   the   book   to   Maria
       'Give Maria the book!' (2SG)

---

broad orthographic representation largely based on Italo-Romance practices. For the transliteration and transcription of Italo-Greek forms we adopt here, with some modifications, the relatively simple system used in Papageorgiadis (n.d.) in his adaptation of Karanastasis (1997). All examples taken from published sources are reproduced in their original orthographic form.

12   For the third persons, not attested in our corpus, we follow the literature, which reports the *mi* + indicative pattern for southern Calabrese, e.g., *Mi scrivi!* 'let him write!' (Rohlfs, 1968: 355; cf. also Loporcaro, 1995).

(7)   b.   *Cucinamu!*   (Gallicianò Calabrian)
            cook.PRS.IND.1PL
            'Let's cook!'

       c.   *Pighiattivillu*   *u   vinu   russu!* (Cardeto)
            take.PRS.IND.2PL=you.2PL=it.3SG the wine   red
            'Take the red wine!' (2PL)

(8)   Negative (Gallicianò Calabrian)[13]
       a.   *Non   jiri   a   la   casa!*
            NEG   go.INF   to   the   house
            'Don't go home!' (2SG)

       b.   *Non   nci=   gridamu(*ri)!*
            NEG   them.DAT=   shout.PRS.IND.1PL(.ri)
            'Let's not shout at them!'

       c.   *Non   nci=   gridati(ri)!*
            NEG   them.DAT=   shout.PRS.IND.2PL(.ri)
            'Don't shout at them!' (2PL)

       d.   *Li   figghioli   non   mi   toccanu   nenti!*
            the   kids   NEG   that_IRREALIS touch.PRS.IND.3PL nothing
            'Don't let the kids touch anything of mine!'

(9)   Negative (Mosorrofa, Cardeto)[14]
       a.   *Non   nci=   gridari   u   figghiolu!* (Mosorrofa)
            NEG   him.DAT=   shout.INF   the kid
            'Don't shout at the kid!' (2SG)

       b.   *Non   jjimu   me   cattamu   u   pane!* /
            NEG   go.PRS.IND.1PL that_IRREALIS buy.PRS.IND.1PL the bread
            non   gridamuri!   (Cardeto)
            NEG   shout.PRS.IND.1PL.ri
            'Let's not go and buy the bread / let's not shout!' (2PL)

---

13   For the *non mi* + indicative pattern of the third person singular in southern Calabrese, see
      Rohlfs (1968: 355), e.g., *Num mi curri nuddu* 'Nobody run!' (cf. also Loporcaro, 1995).

14   On the *mi* + indicative pattern for the third person singular, our corpus includes examples
      from nearby localities, e.g., *Non mi tocca nente!* 'Don't let him/her touch anything of mine!'
      (Chorìo di Roghudi).

c. *Non   nci=   gridatiri   u   figghiolu!*[15] (Mosorrofa)
   NEG   him.DAT=   shout.PRS.IND.2PL.ri   the kid
   'Don't shout at the kid!' (2PL)

d. *E   figghioli   non   me   gridanu!*   (Cardeto)
   the kids   NEG   that$_{\text{IRREALIS}}$ shout.PRS.IND.3PL
   'Don't let the kids shout (at me)!'

The facts for Gallicianò therefore call into question the validity of Loporcaro's original account based solely on the dialect of Mosorrofa. More specifically, if *-ri* extends from the second person singular to restore a two-pattern paradigm driven by the distributional symmetry of the *mi*/infinitive alternation exhibited in embedded contexts (viz. the n(c_2) and n(b) patterns: 2SG, 1/2PL vs. 3SG/PL), its failure to extend to the first person plural, as witnessed in Gallicianò, produces a system which still retains three patterns and where the *mi*/infinitive alternation is not wholly restored (viz. partial n(a_1) pattern, (partial) n($c_{(1)}$) pattern and n(b) pattern: 1(/2)PL vs. 2SG(/PL) vs. 3SG/PL).

3   Language Contact: Greko Sub-/Adstrate

In light of our observations regarding the more nuanced distribution of the *-ri* morpheme alongside alternative patterns with *mi*, the diachronic processes by which this erstwhile infinitival ending was extended must be reconsidered. More specifically, we claim that the driving factor in the extension of *-ri* is the role played by the underlying local Greek sub-/adstrate, viz. Greko, which is shared by all three of the villages exhibiting this phenomenon. Indeed, our parallel investigations in the surviving Greko-speaking villages in the province of Reggio Calabria, namely Bova, Chorìo di Roghudi, Roghudi as well as Gallicianò itself, have brought to light a number of patterns in the imperative paradigm which, we argue, have played a role in shaping the observed Romance forms and their distributions.

We start by considering the paradigm for the positive imperative in Greko as reported in traditional sources (Falcone, 1973: 288; Rohlfs, 1977: 107–109;

---

15   Data reported in Loporcaro (1995: 348, fn.36) show that *-ri* is always optional in Mosorrofa too.

TABLE 5   Positive imperative in Greko

	Positive imperative	Pattern
2SG	IMP (-e/-a, 1st/2nd conj.)	γ
3SG	*na* 'THAT$_{\text{IRREALIS}}$' + SBJV	β
1PL	= 1PL IND	δ
2PL	IMP (-ete/-ate, 1st/2nd conj.)	γ
3PL	*na* 'THAT$_{\text{IRREALIS}}$' + SBJV	β

Katsoyannou, 1995: 324; Karanastasis, 1997: 82–83; Violi, 2004: 72), which is given in Table 5.[16]

We observe 3 distinct patterns. Pattern (γ) characterizes the second person singular and plural where we find dedicated imperatival forms built on the erstwhile perfective stem followed by distinctive singular/plural person formatives, as illustrated in (10a-b). These, in turn, formally contrast with the corresponding (exhortative/jussive) subjunctive forms (11a-b), also built on the erstwhile perfective stem and introduced by the irrealis modal particle *na* (cf. southern Calabrian *mi* above), and with the corresponding indicative forms built on the original imperfective stem (12a-b).[17]

(10)   a.   *Grázz-e!*   (Bova; Rohlfs, 1977)
            write.IMP.2SG
            'Write!'

       b.   *Grázz-ete!*
            write.IMP.2PL
            'Write!'

(11)   a.   *Na*   *grazz-i!*   (Bova; Rohlfs, 1977)
            that$_{\text{IRREALIS}}$ write.SBJV.2SG
            '(That) you should write!'

       b.   *Na*   *grázz-ite!*
            that$_{\text{IRREALIS}}$ write.SBJV.2PL
            '(That) you should write!'

_____

16   For the first person plural, Karanastasis (1997: 83) reports the use of the subjunctive.
17   See Katsoyannou (1995: 286, 292) on an *-ete* variant for the second person plural form of the subjunctive (11b) and present indicative (12b).

(12)   a.   *Gráf-i(s)(e).*   (Bova; Rohlfs, 1977)
            write.PRS.IND.2SG
            'You write.'

       b.   *Gráf-ite.*
            write.PRS.IND.2PL
            'You write.'

The second pattern (β) is suppletive and is found with the third persons which employ the subjunctive (13a-b), a verb form based on the earstwhile perfective stem introduced by the modal irrealis particle *na* largely comparable to the third-person Romance forms introduced by *mi* reviewed above (cf. the p(b) pattern in Table 4).

(13)   a.   *Na*   *cámi!*   (Bova; Rohlfs, 1977: 109)
            that_IRREALIS   do.SBJV.3SG
            'Let him do it!'

       b.   *Na*   *cámusi!*
            that_IRREALIS   do.SBJV.3PL
            'Let them do it!'

The third and final pattern (δ) is unique to the first person plural which suppletively employs the corresponding present indicative form (14), albeit marked by distinct prosodic (viz. intonational) properties.

(14)   *Gráfome!*   (Greko; Rohlfs, 1977: 109)
       write.PRS.IND.1PL
       'Let's write!'

As for the negative imperative, Rohlfs (1977: 193) reports for the second persons singular and plural the suppletive use of the subjunctive paradigm optionally introduced by the irrealis modal particle *na* but obligatorily marked by the distinctive irrealis negator *mi* (cf. realis negator (*d*)*en*).

(15)   a.   *Mi*   *písi!*   (Greko; Rohlfs, 1977: 193)
            NEG   drink.SBJV.2SG
            'Don't drink!'

       b.   *Mi*   *klázzite!*
            NEG   cry.SBJV.2PL
            'Don't cry!'

TABLE 6   Greko imperative (our corpus)

Positive	Pat.	Negative	Pat.
2SG IMP (-e/-a)	γ	(na 'THAT$_{\text{IRREALIS}}$') mi 'NEG'+SBJV	β
3SG na 'THAT$_{\text{IRREALIS}}$'+SBJV	β	(na 'THAT$_{\text{IRREALIS}}$') mi 'NEG'+SBJV	β
1PL IND.1PL	δ	den 'NEG'+IND.1PL	δ
		((ia) na '(for) THAT$_{\text{IRREALIS}}$') mi 'NEG'+SBJV	β
		den 'NEG' èkhome na 'have.IND.1PL THAT$_{\text{IRREALIS}}$'+SBJV	ε
2PL IMP (-ete/-ate)	γ	(na 'THAT$_{\text{IRREALIS}}$') mi 'NEG'+SBJV	β
3PL na 'THAT$_{\text{IRREALIS}}$'+SBJV	β	(na 'THAT$_{\text{IRREALIS}}$') mi 'NEG'+SBJV	β

While our own recent investigations confirm the traditional picture reported above for the Greko positive imperative, the situation for the negative imperative proves more variable, as exemplified in (16)-(17) and summarised in Table 6.

(16)   Positive (Greko)[18]

    a. *Kame*   *sirma!*   (Gallicianò)
       do.IMP.2SG   quickly
       'Hurry up!'

    b. *Kherònnome*   *na*   *kàmome*
       start.PRS.IND.1PL   that$_{\text{IRREALIS}}$ make.SBJV.1PL
       to   *fajì!*   (Chorìo di Roghudi)
       the   food
       'Let's start to prepare the food!'

    c. *Piateto*   *tundo*   *krasì!*   (Bova)
       take.IMP.2PL=it   this   wine
       'Take this wine!'

(17)   Negative (Greko)

    a. *Na*   *mi*   *kuddise!*   / *Mi*   *platèssise*
       that$_{\text{IRREALIS}}$ NEG   shout.SBJV.2SG   NEG   speak.SBJV.2SG
       fitta!   (Gallicianò)
       loudly
       'Don't shout!' / 'Don't speak loudly!'

---

18   For the third persons, not attested in our corpus, we follow the literature mentioned above, which reports the *na* + subjunctive pattern, e.g., Bovese *Na kami!* 'Let him/her do it!', *Na grázzusi!* 'Let them write!' (Rohlfs, 1977: 107, 109).

b. *San   arrivespi   i   Maria, mi*
    when   arrive.PRS.IND.3SG the Maria NEG
    anjì   tìpote!   (Chorìo di Roghudi)
    touch.SBJV.3SG nothing
    'When Maria arrives, don't let her touch anything!'

c. *Den   kuddìzome!*   (Gallicianò)
    NEG   shout.PRS.IND.1PL
    'Let's not shout!'

d. *Ia   na   mi   pame   grìgora!* (Bova)
    for   that_IRREALIS NEG   go.SBJV.1PL   quickly
    'Let's go not quickly!'

e. *Na   mi   kuddime!*   (Gallicianò)
    that_IRREALIS NEG   shout.SBJV.1PL
    'Let's not shout!'

f. *Mi   tu=   kuddìome   panta   sta*
    NEG   them.DAT= shout.SBJV.1PL always   to.the
    pedìa   =ma!   (Bova)
    kids   =our.GEN
    'Let's not always shout at our kids!'

g. *Den   èkhome   na   tavrime*
    NEG   have.PRS.IND.1PL that_IRREALIS beat.SBJV.1PL
    tu   pedìu!   (Gallicianò)
    the.DAT kid.DAT
    'Let's not hit the kid!' (lit. we do not have to hit the kid)

h. *Na   mi   kuddite!*   (Gallicianò) /
    that_IRREALIS NEG   shout.SBJV.2PL
    mi   kuddite!   (Chorìo di Roghudi)
    NEG   shout.SBJV.2PL
    'Don't shout!'

i. *An   ertun   ta   pedìa, mi*
    if   come.PRS.IND.3PL the kids   NEG
    njìun   tìpote!   (Gallicianò)
    touch.SBJV.3PL nothing
    'If the kids come, don't let them touch anything!'

We thus see a general extension in the negative paradigm of suppletive pattern p(β) consisting of the subjunctive (built on the erstwhile perfective stem) to all persons, indeed the only option outside of the first person plural. In the first person plural, however, we find as many as three different strategies: the (β) pattern, including a variant in which the irrealis modal particle is reinforced by the preposition *ia* 'for' (cf. southern Calabrian *pe mmi/pemmi* 'for(.)SBJV.PRT', standard modern Greek για να 'for SBJV.PRT'), the simple (δ) pattern in conjunction with the realis negator *den*, and finally pattern (ε) involving the deontic modal *ekho* 'have' followed by a subjunctive clause (cf. the Italo-Romance Types D and E in Table 2).[19]

If we now compare the imperatival paradigms in Greko with the forms of the negative imperative in the Romance dialect of Gallicianò (viz. Gallicianò Calabrian) in Table 7, we witness an interesting parallel. In particular, we note that in Gallicianò Calabrian the extension of *-ri* from the second person singular to the second person plural, but crucially not to the first person plural, viz. the n(c_1) pattern, restores the underlying p(γ) pattern of the Greko positive imperative where we see that just the second persons singular and plural share dedicated imperatival forms. Significantly, this distribution of dedicated imperatival forms limited to just the second persons singular and plural is not otherwise found in the local Romance varieties (cf. Table 4), but is robustly attested in Greko, the contact model language which we believe to have shaped this innovative Romance replica pattern n(c_1).[20] Given these facts, it is our claim that it can hardly be coincidental that Greko was lost considerably earlier in Mosorrofa and Cardeto than in Gallicianò, where the language survives to the present day. In particular, we argue that the differential extension of *-ri* within the negative imperative paradigm of the Calabrian dialects of Gallicianò (the n(c_1) pattern in Tables 4 and 7) on the one hand, and Mosorrofa and Cardeto (the n(c_2) pattern in Table 4) on the other, reflects the fact that in conservative Gallicianò, where

---

19   According to Karanastasis (1997: 85), the first person plural can be conveyed via different strategies also in the positive paradigm of medio-passive verbs in Greko.

20   One might still object that the Gallicianò Calabrian pattern can be more naturally interpreted as an endogenous, rather than an exogenous, change in which the extension of *-ri* to the second person plural simply represents a natural, language-internal grouping of the second persons. However, if this were the case, then we must ask why this same natural endogenous extension is never attested in the thousands of other Italo-Romance dialects (nor Romansch, Daco-Romance or early langue d'oïl varieties) which also suppletively employ the infinitive in the second person singular negative imperative, but, rather, only in those dialects which are or have recently been in contact with Greko.

TABLE 7   Imperative in Greko (all villages) and Gallicianò Calabrian (our corpus)

		Greko			Gallicianò Calabrian	
	Positive	Pattern	Negative	Pattern	Negative	
2SG	IMP (-e/-a)	γ	(*na*) *mi*+SBJV	β	*non*+INF-***ri***	c_1
3SG	*na*+SBJV	β	(*na*) *mi*+SBJV	β	*non*+*mi*+IND.3SG	b
1PL	IND.1PL	δ	*den*+IND.1PL	δ	*non*+IND.1PL	d
			((*ia*) *na*) *mi*+SBJV	β		
			den èkhome na	ε		
2PL	IMP (-e/-a)	γ	(*na*) *mi*+SBJV	β	*non*+IND.2PL+-***ri***	c_1
3PL	*na*+SBJV	β	(*na*) *mi*+SBJV	β	*non*+*mi*+IND.3PL	b

the underlying Greko pattern p(γ) (cf. Tables 5–7) is still present, *-ri* is only extended to the second person plural, whereas in innovative Mosorrofa and Cardeto, where the Greko distributional model p(γ) is no longer observable,[21] *-ri* has been analogically extended to the first person plural. In turn, this produced, not by chance, a symmetrical distribution with the Romance (viz. Calabrian) positive imperative which also groups together second person singular, second person plural and first person plural within a single pattern (cf. the p(a) pattern in Tables 3 and 4).[22] If this contact-induced analysis is correct, witness further the fact that there are crucially no varieties in which *-ri* extends to the first person plural but not to the second person plural, we are witnessing yet another case of what Rohlfs aptly called *spirito greco, materia romanza* whereby an original Greek PAT(tern) has been recreated through Romance MAT(erial) (cf. Matras and Sakel, 2007).

4   Competition between *-ri* and *mi*

Alongside the negative forms of the Romance imperative incorporating (the extension of) infinitival *-ri* in the second persons singular and plural, as well

---

21   In Cardeto Greko was lost in the latter half of the nineteenth century (cf. Morosi, 1878: 1; Rohlfs, 1977: XIX n.9a; Martino, 1980: 7–8; Squillaci, 2017: 170 n.30). Similarly, in the census of 1861 Greko was reported still to be spoken among members of the community in Mosorrofa, though not as robustly as in Cardeto, before dying out before the end of the century (Martino, 1980: 6–7).

22   A cursory examination of the patterns described in Manzini and Savoia (2005: 388ff.) reveals the (synchronic) spreading of a suppletive form according to the hierarchy 2SG>2PL>1PL to be very common across Italo-Romance.

as the first person plural in Mosorrofa and Cardeto, we also find competing formations with these same three persons involving the irrealis modal particle *mi* plus the present indicative (18a-c; cf. Rohlfs, 1968: 355; Ledgeway, 1998: 47), otherwise typical of the third persons (cf. the p/n(b) patterns in Tables 3 and 4).[23] The result is a Calabrian replication of the Greko pattern n(β) in Table 7 illustrated in Table 8.

(18) a. *Non*   *mi*   *ti=*   *permetti!*   (Mosorrofa)
        NEG   that_IRREALIS   you=   permit.PRS.IND.2SG
        'Don't you dare!'

     b. *Non* *mi*   *nci=*   *lu*   *dicimu!* (Gallicianò Calabrian)
        NEG that_IRREALIS him.DAT= it.ACC= say.PRS.IND.1PL
        'Let's not tell him!'

     c. *Non*   *mi*   *nci=*   *minati!*   (Gallicianò Calabrian)
        NEG   that_IRREALIS him.DAT= beat.PRS.IND.2PL
        'Don't beat him!'

Notably, however, whenever these suppletive forms introduced by the irrealis modal particle are employed in the second persons singular and plural, as well as in the first person plural in Mosorrofa and Cardeto, they invariably prove incompatible with the extension of *-ri*, witness the representative second person plural examples in (19a-b).

(19) a. *Non* *mi*   *nci=*   *minati(*ri)!*   (Gallicianò Calabrian)
        NEG that_IRREALIS him.DAT= beat.PRS.IND.2PL(.ri)
        'Don't hit him!'

     b. *Non*   *nci=*   *minati(ri)!*
        NEG   him.DAT=   beat.PRS.IND.2PL(.ri)
        'Don't hit him!'

TABLE 8   Greko and southern Calabrian negative imperative

	Greko		Southern Calabrian	
	Negative	**Pattern**	**Negative**	**Pattern**
2SG	(*na* 'THAT$_{\text{IRREALIS}}$') *mi* 'NEG'+SBJV	β	*non*+*mi*+IND	b
			non+INF-*ri*	c$_{(1(/2))}$
3SG	(*na* 'THAT$_{\text{IRREALIS}}$') *mi* 'NEG'+SBJV	β	*non*+*mi*+IND	b
1PL	((*ia*) *na* '(for) THAT$_{\text{IRREALIS}}$') *mi* 'NEG'+SBJV	β	*non*+*mi*+IND	b
	den 'NEG' + IND.1PL	δ	*non*+IND.1PL	a$_1$ (d)
	den 'NEG' *èkhome na* 'have.IND.1PL'+SBJV	ε	*non*+IND.1PL-*ri*	c$_2$
2PL	(*na* 'THAT$_{\text{IRREALIS}}$') *mi* 'NEG'+SBJV	β	*non*+*mi*+IND	b
			non+IND.2PL	a$_1$
			non+IND.2PL+-*ri*	c$_1$
3PL	(*na* 'THAT$_{\text{IRREALIS}}$') *mi* 'NEG'+SBJV	β	*non*+*mi*+IND	b

It is natural then to ask why *-ri* has not been extended to examples such as (19a) if, as we have argued in Section 3, *-ri* is extended in conjunction with *non* + 2PL indicative to restore an underlying Greko formal distribution (viz. the p(γ) pattern in Table 7), and then subsequently extended by analogy to the first person plural in Mosorrofa and Cardeto following the loss of Greko in these two localities in line with the p(a) pattern in Table 4. The answer, we argue, lies in the functional structure and licensing of imperatival clauses. Specifically, we adopt here the idea widespread in the literature (Rivero, 1994a; 1994b; Graffi, 1996; Zanuttini, 1997; Manzini and Savoia, 2005: 388) that imperatival clauses display a reduced functional structure. In particular, while declaratives are standardly argued to project a full array of functional projections associated with the T-domain (20a), imperatival clauses (20b) are assumed to lack this same series of functional projections (cf. also Tortora, 2014: ch.3, Section 6). Not by chance, the absence of T-related functional projections in (second-person singular) imperatival clauses is correlated with the frequent traditional observation that one of the most notable characteristics of the imperative is its absence of any inflexional marking or, at the very least, very minimal inflexional marking in accordance with a widespread cross-linguistic tendency (Bybee, 1985: 173; Floricic, 2008: 10; Ledgeway, 2014). Theoretically, we can interpret the observed inflexional impoverishment of the imperative in terms of the mechanisms of feature transmission and inheritance (Chomsky, 2007; 2008): whereas phi-features that originate on the phase head, viz. C°, are usually 'transferred' down to T° in root declaratives, in the absence of T°

and related functional structure in imperatives these same features fail to be passed down – or, to borrow Ouali's (2008) terminology are 'kept' – such that the imperatival verb is forced to raise to C° to license its inflexional features (Rivero, 1994a; 1994b; Rivero and Terzi, 1995; Manzini and Savoia, 2005: 388). As a consequence, any clitics are stranded *in situ* within the *v*-VP complex from where they subsequently encliticize, not syntactically, but phonologically at PF to the imperatival verb now raised to C°.

(20) a. [$_{TP}$ *Gli=*   *avete*   [$_{v\text{-}VP}$ *servito*   *la   cena.*]] (Italian)
              him.DAT= have.PRS.IND.2PL   serve.PTCP the dinner
          'You have served him dinner.'

     b. [$_{CP}$ *Servite*   [$_{v\text{-}VP}$ *=gli*   ~~*servite*~~ *la   cena!*]]
             serve.IMP.2PL   =him.DAT   the   dinner
          'Serve him dinner!'

In negative imperatives, by contrast, the presence of the sentential negator instantiates a functional head whose presence in the clause necessarily forces the projection of the T-domain, otherwise absent in positive imperatives. As a consequence, negative imperatival clauses are therefore predicted to be inflexionally richer than affirmative imperatival clauses since they automatically come with T-related functional positions to host the inflected verb and any accompanying clitics, as exemplified by the Italian example in (21).

(21)   [$_{CP}$... [$_{TP}$ *Non   la=*   *servite*   [$_{v\text{-}VP}$ *la=* ~~*servite!*~~]]] (Italian)
                 NEG   it.ACC=   serve.IMP.2PL
       'Don't serve it!'

Further direct proof of this analysis can be seen in numerous Italian dialects (cf. Type D in Table 2) where, in contrast to the positive imperative, the T° head is exceptionally lexicalized in the negative imperative through an overt auxiliary, a reflex of STARE 'stand' (22a), selecting an infinitival complement (Zanuttini, 1994; 1997: 150–54; Manzini and Savoia, 2005: Section 7.2; Ledgeway, 2019). It is logical therefore to assume that so-called suppletive cases of the simple infinitive such as Italian (22b) employed in the second person singular negative imperative (cf. the n(c) pattern in Tables 3 and 4) simply involve a null auxiliary (Kayne, 1992; Portner and Zanuttini, 2003; Zanuttini, 1994; 1997: 118ff.).

(22)   a.   [_TP *non   stá*   [_*v*-VP *parlare!*]] (Padua; Zanuttini, 1997: 119)
                 NEG stand.PRS.IND.2SG   speak.INF
           'Don't speak!'

       b.   [_TP   *Non   ø*   [_*v*-VP *parlare!*]]   (Italian)
                   NEG   AUX.PRS.IND.2SG   speak.INF
           'Don't speak!'

In the light of this evidence, we can conclude that extension of the infinitival marker *-ri* in the relevant dialects of southern Calabria to the second person plural (23c) and, in Mosorrofa and Cardeto, to the first person plural (23b), implies its selection by a null auxiliary in TP along the lines of the second person singular (23a). Consequently, forms such as *parrati-ri* and *parramu-ri* which are selected by a null auxiliary must be reanalysed as inflected infinitives, in contrast to the original competing forms without *-ri* (viz. *parrati, parramu*) which, in the absence of a null auxiliary, involve V-raising to the T-domain (23d-e).

(23)   a.   [_TP   *Non   ø*   [_*v*-VP *parrari!*]]   (sth. Calabrian)
                   NEG   AUX.PRS.IND.2SG   speak.INF

       b.   [_TP   *Non ø*   [_*v*-VP *parramuri!*]]   (sth. Calabrian)
                   NEG AUX.PRS.IND.1PL   speak.INF.1PL

       c.   [_TP   *Non   ø*   [_*v*-VP *parratiri!*]]   (sth. Calabrian)
                   NEG   AUX.PRS.IND.2PL   speak.INF.2PL

       d.   [_TP   *Non   parramu* [_*v*-VP   ~~*parramu!*~~]]   (sth. Calabrian)
                   NEG   speak.PRS.IND.1PL

       e.   [_TP   *Non   parrati* [_*v*-VP   ~~*parrati!*~~]]   (sth. Calabrian)
                   NEG   speak.PRS.IND.2PL

Note that there is an important structural difference between this southern Calabrian example of an inflected infinitive and the well-known standard cases of the inflected infinitive reported for other Romance varieties such as Portuguese, Galician, old Leonese, central-northern Sardinian, and old Neapolitan (for an overview and references, see Ledgeway, 2012: 293–294). In the former case, the infinitival marker is the outermost morpheme following the person/number marker (cf. *parra-mu-ri* 'speak-AGR.1PL-INF'), whereas

in the latter the person/number marker is attached to the infinitival marker which is, in turn, suffixed to the verb stem (cf. Portuguese *fala-r-mos* 'speak-INF-AGR.1PL'). These internal structural differences clearly relate to the very distinct paths by which these forms arose, namely adjunction of infinitival *-ri* to a present indicative verb form (viz. *parramu* > *parramu+-ri* > *parramuri*) vs. reanalysis of the Latin imperfective subjunctive (Ibero-Romance, Sardinian: *FABULAREMUS* > Portuguese *falarmos*; cf. Maurer, 1968; Jones, 1992; 1993: 78–82) or pluperfect indicative (*PARABOLA(UE)RAMUS* > Old Neapolitan *parlar(e)m(m)o*; cf. Loporcaro, 1986).

Returning to the incompatibility of *-ri* with the competing negative imperatival forms introduced by the irrealis modal particle *mi* (cf. 19a and the n(b) pattern in Table 8), we now have a principled explanation for this observation. Assuming with Ledgeway (1998; 2007; 2013: 4 n.9) and Damonte (2010) that Calabrian *mi* is a T-element, its incompatibility with infinitival *-ri* follows without further stipulation since *mi* would be competing for the same T-position as the null auxiliary required to license infinitival *-ri*. The complementary distribution of *mi* and the extension of infinitival *-ri* illustrated in (24a-b) therefore falls out naturally.[24]

(24) a. [~TP~ *Non*   **ø**   [~v-VP~ *parratiri!*]]
                NEG   AUX.PRS.IND.2PL   speak.PRS.IND.2PL.INF
      b. [~TP~ *Non*   **mi**   [~v-VP~ *parrati!*]]
                NEG   that~IRREALIS~   speak.PRS.IND.2PL
          'Don't speak!'

5   Extension of *-ri*: a Hybrid Pattern?

In view of the contact-induced developments and associated distributional patterns considered above, it is now time to return to the distinction between true and suppletive imperatives introduced at the outset of this article. Above we noted following Zanuttini (1997: 105) that true imperatives involve "verbal forms that are unique to the paradigm of the imperative, in the sense that they are different from any other verbal form used for the same person in any other verbal paradigm." Given this definition, we are led to conclude that extension of *-ri* from the second person singular to the second (and first) person(s) plural

---

24   There are at least two other possible competing explanations for the observed incompatibility of *-ri* with *mi*-forms of the negative imperative. One, a functional-based approach, would

TABLE 9　　*Mi-* and *-ri* patterns in Mosorrofa, Cardeto, Gallicianò Calabrian

	Grammar A	Pattern	Grammar B	Pattern
2SG	*non* + INF-*ri*	$g_{(1(/2))}$	*non **mi*** + IND	b
3SG	*non **mi*** + IND	b	*non **mi*** + IND	b
1PL	*non* + IND (+ -*ri*)	a (g_2)	*non **mi*** + IND	b
2PL	*non* + IND (+ -*ri*)	a (g_1)	*non **mi*** + IND	b
3PL	*non **mi*** + IND	b	*non **mi*** + IND	b

as part of the emergence of an $n(c_{(1(/2))})$ pattern gives rise to the concomitant genesis of novel true negative imperatives, as schematized in Table 10.

　　Whereas the use of the Romance infinitive in the second person singular negative imperative is standardly interpreted as suppletive (cf. Type B in Table 2),[25] it is highly questionable whether this analysis is still applicable for

---

be to argue, following Ledgeway (1998: Section 7), that synchronically southern Calabrian *mi*-clauses behave and should be analysed as inflected infinitival clauses in which proclitic *mi*, just like the infinitival suffix -RE (> southern Calabrian -*ri*) also licensed under T° (i.a; cf. Kayne, 1991), functions as an infinitival marker generated under T° in conjunction with a verb inflected for person and number (i.b).

(i)　a.　[$_{TP}$ *parla-re*　[$_{AgrSP}$　-*mo*　　　　[$_{v\text{-}VP}$　~~*parla-*~~]]]　(old Neapolitan)
　　　　　speak-INF　　　　-AGR.1PL

　　b.　[$_{TP}$ ***mi***　[$_{AgrSP}$　*parra-mu*　　　[$_{v\text{-}VP}$　~~*parra-*~~]]]　(sth. Calabrian)
　　　　　that$_{IRREALIS}$　　speak- AGR.1PL

Given the functional and structural equivalence between proclitic *mi* and suffixal -RE (> -*ri*), both analysed as infinitival T-markers, *mi* is predicted to be incompatible with the extension of an additional infinitival marker. Another approach would be to view the *mi*-pattern and *ri*-pattern as the outputs of two originally distinct grammars, as in Table 9. We have seen above that in Grammar A extension of -*ri* to just the second person plural (and subsequently in Mosorrofa and Cardeto also to the first person plural) is driven by contact-induced internal pressure to restore the distributional pattern of the underlying Greko positive imperative (cf. the p(γ) pattern in Tables 5 and 6), hence its exclusion from the (first person plural and) third persons displaying *mi*-forms. In Grammar B, by contrast, the negative imperative is built on *mi*-forms in all persons (cf. 18a-c) in replication of the Greko negative imperative n(β) pattern in Tables 6 and 7. The infinitive then has no foothold in the negative imperative paradigm of Grammar B, not even in the second person singular, from which infinitival -*ri* could extend its distribution to come into contact with *mi*-forms.

25　Recall that the use of the infinitive in the second person singular negative imperative represents a suppletive use, since the infinitive may equally mark the second person singular (as well as all other persons) in embedded contexts as the complement to raising and control predicates.

TABLE 10   Summary of Greko and southern Calabrian imperative paradigms

	Greko		**Gallicianò Calabrian**		**Mosorrofa/Cardeto**	
	Positive	**Negative**	**Positive**	**Negative**	**Positive**	**Negative**
2SG	**true**	suppletive	**true**	**suppletive** > **true**	**true**	**suppletive** > **true**
3SG	suppletive	suppletive	suppletive	suppletive	suppletive	suppletive
1PL	suppletive	suppletive	suppletive	suppletive	suppletive	**true**
2PL	**true**	suppletive	suppletive	**true**	suppletive	**true**
3PL	suppletive	suppletive	suppletive	suppletive	suppletive	suppletive

many, if not all, southern Calabrian dialects. In the dialects of this area there is an ongoing and now highly advanced retreat of the infinitive (cf. Ledgeway, 2013: 4, 17–18; 2016: 1024–1025; Squillaci, 2017: chs 4,5; Ledgeway, Schifano and Silvestri, in prep.) such that its use as an embedded verb form is now, at best, a residual syntactic feature limited to the complement of a dwindling handful of functional predicates (essentially CAN; cf. Romanian),[26] all of which now overwhelmingly select a so-called finite *mi*-clause. The result is that synchronically this verb form is now predominantly, if not wholly in the speech of many speakers, no longer to be considered functionally an infinitive since its distribution is now almost exclusively restricted to the second person singular negative imperative. Given this loss of its subordination uses and its increasing restriction to the second person singular negative imperative, the output of an ongoing and well-advanced process of refunctionalization (cf. Smith, 2005; 2011), it is natural to reinterpret the erstwhile infinitive as having shifted from suppletive (25a) to true (25b) imperative in the Romance negative paradigms of Table 10.

(25)   *Non*   *parra-ri!*   (Mosorrofa)
       a.   NEG   speak-INF   (suppletive)
       b.   NEG   speak-IMP(.2SG)   (true)
                      'Don't speak!'

---

26   Indeed, this presumably explains the exceptional persistence of the infinitive in the second person singular negative imperative where, as argued in Section 4 (cf. 23a), it is selected by a null auxiliary.

By the same token, the extension of *-ri* to the second person plural (26a) and, in turn, to the first person plural (27a) through the emergence of the $n(c_{1(/2)})$ pattern further reinforces and extends this novel 'true' negative imperative pattern (26b, 27b), inasmuch as the distribution of all the relevant *-ri* forms is unique to the paradigm of the negative imperative.

(26)   *Non*   *parramu-ri!*   (Mosorrofa)
   a.   NEG   speak.PRS.IND.1PL-ri
   b.   NEG   speak.IMP.1PL
       'Let's not speak!'

(27)   *Non*   *parrati-ri!*   (Mosorrofa)
   a.   NEG   speak.PRS.IND.2PL-ri
   b.   NEG   speak.IMP.2PL
       'Don't speak!'

Also relevant here is the striking parallel found in old Romanian. Following Mării (1969) and Zamfir (2005; 2007), Maiden et al. (forthcoming: Section 6.3.4) observe how in sixteenth- and seventeenth-century Romanian the 'long' form of the infinitive in *-re* used to mark the second-person singular imperative (28a),[27] alongside the short form of the infinitive still used today (28b), is extended to mark the second-person plural with concomitant adjunction of the second-person plural marker *-ţi* (29b), alongside the older suppletive second-person present indicative form (29a) still in use today. Differently from the southern Calabrian case where the erstwhile infinitival suffix *-ri* is extended to the second-person plural present indicative form (viz. *nu parrati > nu parratiri*), in old Romanian we see an extension of the entire infinitival form, not just the infinitival suffix, to the second person plural to which the distinct second-person plural marker is then added. In this respect, the internal structure of the old Romanian relevant form resembles the canonical forms of the inflected infinitive considered above (cf. Pt. *fala-r-des* 'speak-INF-2PL'). However, the southern Calabrian and old Romanian cases come together in several key respects: i) they both involve the genesis of a novel second-person plural 'true' imperative through extension of the infinitival suffix or the infinitive; ii) extension (initially) excludes the first-person plural (which is distinctly marked in (old) Romanian suppletively by a preverbal irrealis modal marker along the lines of the n/p(b) pattern), bringing together the second persons singular and plural to restore an already existing distributional pattern (cf. *nu*

---

27   Although not used in official orthography, in the examples below we use a grave accent to indicate stress placement for expository purposes.

cântà!: *nu cântàṭi!* vs. *nu cântàre!*: *nu cântàreṭi!*); and iii) on a par with south-ern Calabrian, (old) Romanian is characterized by a heavily reduced use of the subordinating uses of the infinitive which is increasingly restricted to just the second-person singular negative imperative, a necessary step for the reanalysis of the suppletive use of the infinitive as an innovative true imperative.

(28) a. *Nu*   *cânt-àre!*   (Old Romanian)
          NEG   sing-INF
      b. *Nu*   *cântà!*
          NEG   sing.INF
          'Don't sing!'

(29) a. *Nu*   *cântàṭi!*   ((Old) Romanian)
          NEG   sing.PRS.IND.2PL
      b. *Nu*   *cântàre-ṭi!*
          NEG   sing.INF-2PL
          'Don't sing!'

Significantly, then, we see that, although the infinitive suppletively marks the second-person singular negative imperative in thousands of (especially Italo-) Romance varieties, it is only extended in those varieties where the subordinat-ing uses of the infinitive are radically attrited – under language contact with Greko in the case of southern Calabrian, and under language contact with other varieties of the Balkan Sprachbund (cf. Joseph, 1983; 2010; Friedman, 2006; 2011; Tomić, 2006; Friedman and Joseph, 2017; 2021; Krapova and Joseph, 2019, a.o.; see also Gardani et al., 2021) and, in particular, Greek in the case of Romanian –, such that infinitival morphology is free to be exaptively reinter-preted as a 'true', dedicated imperatival marker and from there extended to the second-person plural according to an already salient (language-internal or -external) paradigmatic distribution. We thus see that a so-called typi-cal Balkanism has independently undergone a very similar development in two areas – southern Calabria and Romania – which have not otherwise been in contact with each other. This, in turn, underlines how so-called original Balkanisms can give rise, through time, to new "second-generation" Balkanisms, both within and outside the Balkan Sprachbund proper, thereby further reinforcing the linguistic cohesion of such varieties.

Within a wider Romance typology, the resulting opposition between true and suppletive imperatives in southern Calabrian (and now also in old Romanian) presents us with a unique distribution not previously recorded for Romance (cf. Table 1), as illustrated in Table 11. Table 11 presents all the logi-cal combinations of true and suppletive forms for the Romance positive and

negative imperative. Type (i) is found in varieties such as Piedmontese (30) which presents, at least for the 2SG, a true imperatival form – indeed the same distinctive form – in both the positive and negative paradigms, an apparent counterexample to the so-called negative imperative puzzle. Type (ii), by contrast, shows a formal opposition between the positive and negative paradigms through the alternation between true and suppletive forms, respectively, where the latter can be an infinitive as in Italian (31b) or a subjunctive as in Gascon (32b).

(30) a. *Parla!*   (positive, true) (Piedmontese; Zanuttini, 1997: 111)
           speak.IMP.2SG
           'Speak!'

      b. *Parla*   *nen!*   (negative, true)
           speak.IMP.2SG   NEG
           'Don't speak!'

(31) a. *Parla!*   (positive, true)   (Italian)
           speak.IMP.2SG
           'Speak!'

      b. *Non*   *parlare!*   (negative, suppletive)
           NEG   speak.INF
           'Don't speak!'

(32) a. *Cante!* (positive, true) (Béarnais Gascon; Puyau, 2013: 91–92)
           speak.IMP.2SG
           'Sing!'

      b. *Ne*   *càntes*   *pas!*   (negative, suppletive)
           NEG   sing.PRS.SBJV.2SG NEG
           'Don't sing!'

TABLE 11   Combinations of true/suppletive forms in Romance positive/negative imperative

	Positive	Negative	Example
(i)	true	true	Piedmontese (30)
(ii)	true	suppletive	Italian (31), Gascon (32)
(iii)	suppletive	suppletive	Catalan (33), Venetan (34)
(iv)	suppletive	true	?

The remaining two combinations in Table 11 both involve a suppletive paradigm in the positive imperative. In type (iii), the suppletive positive imperative combines with a suppletive paradigm in the negative which can either be the same as that of the positive imperative as in Catalan (33a-b) or distinct as in Venetan (34a-b).

(33)　　a.　*Parleu!*　(positive, suppletive)　(Catalan; Zanuttini, 1997: 109)
　　　　　　　speak.PRS.IND.2PL
　　　　　　　'Speak!'

　　　　b.　*No*　　　　*parleu!*　　　　　　　(negative, suppletive)
　　　　　　　NEG　　　　speak.PRS.IND.2PL
　　　　　　　'Don't speak!'

(34)　a.　*ʧa'mɛi=lu*　(positive, suppletive)　(Calizzano; Manzini and Savoia, 2005: 451)
　　　　　　speak.PRS.IND.2PL=him
　　　　　　'Call him!'

　　　b.　*nɛ*　　　*s'tɛi=lu*　　　　　　　　*ʧa'mɔ*　　　(negative, suppletive)
　　　　　　NEG　　　stand.PRS.IND.2PL=him　　call.INF
　　　　　　'Don't call him!'

However, the fourth and final logical combination (iv), which marries together a suppletive positive paradigm with a true negative paradigm, has until now remained merely a theoretical possibility. Given, however, our hypothesis about the creation of novel true negative imperatives in southern Calabrian (Mosorrofa, Cardeto, Gallicianò Calabrian) as a concomitant of the contact-induced (almost complete) loss of subordinating uses of the infinitive and the extension of *-ri*, the empirical gap predicted by the typology presented in Table 11 can now be filled with the southern Calabrian second (and first) person(s) plural (35), as well as the old Romanian case noted by Maiden et al. (forthcoming) reviewed above (cf. 28–29).

(35)　a.　*Parrati!* /　　　　*Parramu!*　(positive, suppletive)　(sth. Calabrian)
　　　　　speak.PRS.IND.2PL　speak.PRS.IND.1PL
　　　　　'Speak! / Let's speak!'

　　　b.　*Non*　　*parratiri* /　　　*parramuri!*　(negative, true)
　　　　　NEG　　　speak.IMP.2PL /　　speak.IMP.2PL
　　　　　'Don't speak / let's not speak!'

The significance of the southern Calabrian (and old Romanian) data in confirming the predicted combination (iv) of Table 11 fills not only a Romance gap, but is also of wider typological significance. In particular, a combined examination of WALS maps 70A (The Morphological Imperative; cf. Auwera, Lejeune, Pappuswamy, and Goussev, 2013) and 71A (The Prohibitive; cf. Auwera, Lejeune and Goussev, 2005; Auwera, 2010) – see http://wals.info/combinations/70A_71A#/-18.771/304.080 – shows that the combination (iv) found in southern Calabrian is indeed typologically extremely rare. For the value 'no second person imperative', map 70A reveals that 122 out of 547 languages do not display morphologically dedicated second-person positive imperatives at all (⇒ suppletive positive imperative), whereas the value 'special imperative + normal negative' for map 71A reveals that out of 495 languages there are just 55 languages where the negative imperative uses a verbal construction different than the positive imperative whilst maintaining the same negative marker as in declaratives (⇒ true or suppletive negative imperative). If we then compare and combine the results for both values from both maps, we can isolate through map 71A languages where the negative imperative is distinct from the positive imperative and then cross-reference this group of languages against just those languages in map 70A where the positive paradigm is suppletive, such that the negative imperative in the same varieties must either be true or suppletive, but in any case distinct from the positive imperative. This yields just three possible languages out of a total of 474, namely Gooniyandi (Australia), Svan (Georgia) and Trumai (Brazil),[28] which, on a par with southern Calabrian, potentially combine suppletive and true forms in the positive and negative paradigms of the imperative, respectively.

However, a closer look at the relevant descriptions shows that in both Gooniyandi (McGregor, 1990: 542–543) and Svan (Tuite, 1997: 42; 2018: 61) the negative imperative is suppletive, and not true: in the former the verb assumes the present definite form and in the latter either the present or future indicative (when introduced by the negator *nom*) or the optative or conjunctive (when introduced by the negator *nosa*). The picture in Trumai proves more complex where the verb shows very little morphology, with relevant grammatical categories mainly expressed by accompanying particles or auxiliaries (Monod-Becquelin, 1975: 98). Indeed, the imperative is variously marked by the particles *wana, waki, wa* and *wanach* (Monod-Becquelin, 1975: Section 3.4), the distribution of which is determined by polarity, transitivity and the animacy of the Patient argument (O). In the positive imperative the uninflected verb

---

28   If we include the combination 'special negative + special imperative' in map 71A, we obtain a further 15 possible languages, which is still a considerably reduced sample.

is preceded by *wana* with intransitives, and by *wa* and *waki* with transitives in accordance with the animacy or otherwise of the O, respectively (Monod-Becquelin, 1975: 111, 250). In the negative imperative, by contrast, the relevant particles must now all follow the verb, namely *wanach* with intransitives and *waki* with transitives (Monod-Becquelin, 1975: 113-112, 251–252). We see therefore that a formal distinction between true and suppletive imperatives is harder to recognise in this case, since there is some overlap in the distribution of the positive and negative particles, e.g., transitive *waki* and potentially also positive intransitive *wana* if related to negative intransitive *wanach*. Syntactically, however, there is a sharp distinction between the preverbal and postverbal positions of the positive and negative imperatival particles. The overall picture is therefore inconclusive in several respects: (i) the positive imperative can hardly be described as 'suppletive', in that the uninflected verb is marked off by distinct preverbal particles; (ii) the negative imperative can hardly be described as 'true', in that there is some formal overlap in the distribution of imperatival particles in positive and negative imperatives; and (iii) the distinction between positive and negative imperatives is robustly marked by the respective preverbal vs. postverbal position of the particle.

Consequently, we tentatively conclude that the dialects of Gallicianò, Mosorrofa and Cardeto, together with old Romanian, constitute to date the only secure examples of combination (iv) in Table 11, an otherwise unattested option, the exceptional presence of which in southern Calabria can plausibly be explained as the hybrid outcome of contact between indigenous Greek and Romance grammars.

6   Conclusions and Summary

The data discussed in this article have shown how language contact between indigenous Greek (viz. Greko) and Romance (viz. southern Calabrian) grammars has led to the creation of a hybrid Romance negative imperative paradigm which marries together traditional finite verb forms (marked for person and number) with an erstwhile infinitival ending. At the basis of this paradigm is a Greek ad-/substrate model where the distinctive marking of the second persons singular and plural in the positive imperative is transferred to the southern Calabrian negative imperative through the extension of the infinitival ending *-ri* from the second person singular to the second person plural and, in turn, by analogy to the first person plural in Mosorrofa and Cardeto.

The synchronic contrast in the extension of *-ri* in Gallicianò Calabrian (2SG > 2PL) on the one hand and in Mosorrofa and Cardeto (2SG > 2PL > 1PL) on the other effectively rules out a Romance-internal motivation for the original 2SG > 2PL stage but, rather, points to a contact-induced change where Romance MAT(erial) has reproduced a Greek PAT(tern) in line with Rohlfs' common mantra *spirito greco, materia romanza* (for discussion, see Ledgeway, 2006; 2013; and Ledgeway, Schifano and Silvestri, in prep. on Greek-Romance contact, see also Ralli, 2021).

As is often the case in situations of language contact, the particular changes affecting the southern Calabrian paradigm of the negative imperative give rise to typologically non-linear and exceptional developments which ultimately distort the expected and regular Romance 'type'. Here we have seen two key examples. Firstly, we have observed how the extension of infinitival *-ri*, first to the second person plural under the influence of a Greek distributional model and then, in Mosorrofa and Cardeto, under the influence of a Romance model to the first person plural, leads formally to the creation of apparently new inflected infinitival forms. Secondly, we are forced to recognise in these same innovative inflected infinitival forms, including the bare infinitive in the second person singular following the almost complete loss of its subordinating functions, the emergence of a novel 'true' paradigm for the negative imperative since all the relevant forms are unique to the negative imperative. At the same time, the emergence of this novel paradigm also brings about an otherwise extremely rare typological opposition between suppletive and true forms in the positive and negative paradigms of the imperative, respectively, which is otherwise only found in a specific period of old Romanian. Once again, we see that this rare typological pattern is not the output of a linear Romance-internal development, but, rather, the consequence of language contact between Greek and Romance grammars, the resolution of which leads to this otherwise typologically extremely rare, but logically predicted, distribution of true and suppletive forms.

Finally, in the present article we have shown that, although the infinitive is suppletively employed to mark the second-person singular negative imperative across several (Italo-)Romance varieties, its extension to the plural is only attested in those varieties where subordinating uses of the infinitive are radically attrited, such as in the southern Calabrian dialects under investigation here and in old Romanian. This case study therefore also offers a significant contribution to the study of the so-called Balkan Sprachbund by bringing to light a new feature associated with a typical Balkanism (viz. the reduced use of

the infinitive), namely the possibility of reinterpreting the infinitival marker as a true, dedicated imperatival marker.[29]

Acknowledgements

References

Aikhenvald, Alexandra. 2008. Imperative: The grammar of a clause type. Omotic utterance type, mood and attitude markers and linguistic typology. Unpublished ms., University of Leiden.

Alcázar, Asier and Mario Saltarelli. 2014. *The Syntax of Imperatives*. Cambridge: Cambridge University Press.

Auwera, Johan van der. 2010. Prohibition: Constructions and markers. In Dingfang Shu and Ken Turner (eds.), *Contrasting Meaning in Languages of the East and West*, 443–475. Tübingen: Narr.

Auwera, Johan van der and Ludo Lejeune, with Valentin Goussev. 2005. The prohibitive. In Matthew Dryer, Martin Haspelmath, David Gil, and Bernard Comrie (eds.), *The World Atlas of Language Structures*, 290–293. Oxford: Oxford University Press.

Auwera, Johan van der and Ludo Lejeune, with Umarani Pappuswamy and Valentin Goussev. 2013. The morphological imperative. In Matthew Dryer and Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/70, accessed on 2019-03-08).

Bybee, Joan. 1985. *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam: Benjamins.

Cavalcante, Rerisson. 2011. Negative imperatives in Portuguese and other Romance languages. In Julia Herschensohn (ed.), *Romance Linguistics 2010. Selected Papers*

---

29   In addition to the limited use of the infinitive, other 'Balkanisms' observed in the Calabrian varieties under investigation here are the genitive/dative syncretism (cf. the so-called *dativo greco* 'Greek-style dative') and the syntax of possessives in indefinite nominal phrases (cf. Ledgeway, 2013; Ledgeway, Schifano and Silvestri, 2018; 2020a, 2020b, in prep.).

from the 40th Linguistics Symposium on Romance Languages, 205–219. Amsterdam: Benjamins.

Chomsky, Noam. 2007. Approaching UG from below. In Uli Sauerland and Hans Martin Gärtner (eds.), *Interfaces + Recursion = Language?: Chomsky's Minimalism and the View from Syntax- Semantics*, 1–29. Berlin: Mouton.

Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos Otero and Maria Luisa Zubizarreta (eds.), *Foundational Issues in Linguistic Theory. Essays in Honor of Jean-Roger Vergnaud*, 133–166. Cambridge, Mass.: MIT Press.

Crucitti, Pasquale. 1988. *Vocabolario e frasario Mosorrofano-Santagatino*. Soveria Mannelli: Calabria Letteraria Editrice.

Damonte, Frederico. 2010. Matching moods: Mood concord between CP and IP in Salentino and Southern Calabrian subjunctive complements. In Paola Benincà and Nicola Munaro (eds.), *Mapping the Left Periphery*, 228–256. Oxford: Oxford University Press.

De Angelis, Alessandro. 2013. *Strategie di complementazione frasale nell'estremo Meridione italiano*. Messina: SGB Edizioni.

Falcone, Giuseppe. 1973. *Il dialetto romaico della Bovesia*. Milan: Istituto lombardo di scienze e lettere.

Floricic, Franck. 2008. The Italian verb-noun anthroponymic compounds at the syntax / morphology interface. *Morphology* 18: 167–193.

Friedman, Victor A. 2006. Balkans as a linguistic area. In Keith Brown (ed.), *Encyclopedia of Language and Linguistics*, 2nd edn., 657–672. Oxford: Elsevier.

Friedman, Victor A. 2011. The Balkan languages and Balkan linguistics. *Annual Review of Anthropology* 40(1): 275–291.

Friedman, Victor A. and Brian D. Joseph. 2017. Reassessing Sprachbunds: A view from the Balkans. In Raymond Hickey (ed.), *The Cambridge Handbook of Areal Linguistics*, 55–87. Cambridge: Cambridge University Press.

Friedman, Victor A. and Brian D. Joseph. 2021. *The Balkan Languages*. Cambridge: Cambridge University Press.

Gardani, Francesco, Michele Loporcaro and Alberto Giudici. 2021. In and around the Balkans: Romance languages and the making of multi-layered languages. *Journal of Language Contact* 14(1): 1–23.

Graffi, Giorgio. 1996. Alcune riflessioni sugli imperativi italiani. In Paola Benincà, Guglielmo Cinque, Tullio De Mauro, and Nigel Vincent (eds.), *Italiano e dialetto nel tempo. Saggi di grammatica per Giulio C. Lepschy*, 133–148. Rome: Bulzoni.

Han, Chung-Hye. 1998. *The Structure and Interpretation of Imperatives: Mood and Force in Universal Grammar*. Pennsylvania: University of Pennsylvania Institute for Research in Cognitive Science Technical Report Series (number 74).

Han, Chung-Hye. 1999. Cross-linguistic variation in the compatibility of negation and imperatives. *West Coast Conference on Formal Linguistics* 17: 265–279.

Han, Chung-Hye. 2001. Force, negation and imperatives. *The Linguistic Review* 18: 289–325.

Ionescu, Emil. 2019. Negative imperatives in Eastern Romance languages: Latin heritage and Romance innovation. *Zeitschrift für romanische Philologie* 135: 845–865.

Isac, Daniela. 2015. *The Morphosyntax of Imperatives*. Oxford: Oxford University Press.

Jones, Michael. 1992. Infinitives with specified subjects in Sardinian. In Christiane Laeufer and Terrell Morgan (eds.), *Theoretical Analyses in Romance Linguistics*, 295–309. Amsterdam: Benjamins.

Jones, Michael. 1993. *Sardinian Syntax*. London: Routledge.

Joseph, Brian. 1983. *The Synchrony and Diachrony of the Balkan infinitive*. Cambridge: Cambridge University Press.

Joseph, Brian D. 2010. Language contact in the Balkans. In Raymond Hickey (ed.), *Handbook of Language Contact*, 618–633. Oxford: Blackwell.

Kayne, Richard. 1991. Romance clitics, verb movement and PRO. *Linguistic Inquiry* 22: 647–686.

Kayne, Richard. 1992. Italian negative infinitival imperatives and clitic climbing. In Liliane Tasmowsky and Anne Zribi-Hertz (eds.), *Hommages à Nicolas Ruwet*, 300–312. Ghent: Communication and Cognition.

Karanastasis, Anastasios. 1997. *Grammatikí ton Ellinikon Idiomaton tis Kato Italias*. Athens: Academy of Athens.

Katsoyannou, Marianne. 1995. *Le parler gréco de Gallicianò* (*Italie*). PhD dissertation, Université Paris VII.

Krapova, Iliyana and Brian Joseph (eds). 2019. *Balkan Syntax and* (*Universal*) *Principles of Grammar*. Berlin and Boston: De Gruyter Mouton.

Ledgeway, Adam. 1998. Variation in the Romance Infinitive. The Case of the Southern Calabrian Inflected Infinitive. *Transactions of the Philological Society* 96: 1–61.

Ledgeway, Adam. 2006. The dual complementiser system in Southern Italy: *Spirito greco, materia romanza?*. In Anna Laura Lepschy and Arturo Tosi (eds.), *Rethinking Languages in Contact: The Case of Italian*, 113–126. Oxford: Legenda.

Ledgeway, Adam. 2007. Diachrony and finiteness: Subordination in the dialects of Southern Italy. In Irina Nikolaeva (ed), *Finiteness: Theoretical and Empirical Foundations*, 335–365. Oxford: Oxford University Press.

Ledgeway, Adam. 2012. *From Latin to Romance. Morphosyntactic Typology and Change*. Oxford: Oxford University Press.

Ledgeway, Adam. 2013. Greek disguised as Romance? The case of Southern Italy. In Mark Janse, Brian Joseph, Angela Ralli, and Metin Bagriacik (eds.), *Proceedings of the 5th International Conference on Greek Dialects*, 184–228. Patras: University of Patras.

Ledgeway, Adam. 2014. La morphologie flexionnelle de l'impératif des dialectes de l'Italie méridionale: La distribution de la métaphonie. In Jean Léo Léonard (ed),

Morphologie flexionnelle et dialectologie romane: Typologie(s) et modélisation(s). Mémoires de la Société linguistique de Paris XXII, 13–33. Louvain: Peeters.

Ledgeway, Adam. 2016. Clausal complementation. In Adam Ledgeway and Martin Maiden (eds.), *The Oxford Guide to the Romance Languages*, 1013–1028. Oxford: Oxford University Press.

Ledgeway, Adam. 2019. The causative construction in the dialects of Southern Italy and the phonology-syntax interface. In Giulia Bellucci, Ludovico Franco and Paolo Lorusso (eds.), *Linguistic Variation: Structure and Interpretation*, 371–400. Berlin: De Gruyter.

Ledgeway, Adam, Norma Schifano and Giuseppina Silvestri. 2018. Il contatto tra il greco e le varietà romanze nella Calabria meridionale. *Lingue antiche e moderne* 7: 95–133.

Ledgeway, Adam, Norma Schifano and Giuseppina Silvestri. 2020a. Microvariation in dative-marking in the Romance and Greek varieties of southern Italy. In Anna Pineda and Jaume Mateu (eds.), *Dative Structures and Beyond*, 317–349. Berlin: Language Science Press.

Ledgeway, Adam, Norma Schifano and Giuseppina Silvestri. 2020b. I possessivi in italo-greco e italo-romanzo: parallelismi strutturali in un'area di contatto. *Archivio Glottologico Italiano* CV (I): 85–114.

Ledgeway, Adam, Norma Schifano and Giuseppina Silvestri. In prep. *The Fading Voices of Southern Italy: A Case Study in Greek-Romance Language Contact in Magna Graecia*. Oxford: Oxford University Press.

Loporcaro, Michele. 1986. L'infinito coniugato nell'Italia centro-meridionale: ipotesi genetica e ricostruzione storica. *Italia Dialettale* 49: 173–240.

Loporcaro, Michele. 1995. Recessività sintattica dell'infinito e vitalità morfologica della desinenza infinitivale in alcuni dialetti del reggino. In Roberto Ajello and Saverio Sani (eds.), *Scritti linguistici e filologici: in onore di Tristano Bolelli*, 331–358. Ospedaletto, Pisa: Pacini.

Maiden, Martin. 2006. On Romanian imperatives. *Philologica Jassyensia* II: 47–59.

Maiden, Martin, Adina Dragomirescu, Gabriela Pană Dindelegan, Oana Uță, and Rodica Zafiu. forthcoming. *The Oxford History of Romanian Morphology*. Oxford: Oxford University Press.

Manzini, Maria Rita and Leonardo Savoia. 2005. *I dialetti italiani e romanci: Morfosintassi generativa. Volume III*. Alessandria: Edizioni dell'Orso.

Mării, Ion. 1969. Note etimologice și morfologice. *Cercetări de lingvistică* 14: 54–56.

Martino, Paolo. 1980. L'isola grecanica dell'Aspromonte. Aspetti sociolinguistici. In Federico Albano Leoni (ed), *I dialetti e le lingue delle minoranze di fronte all'italiano. Atti dell'XI congresso internazionale di studi. Cagliari 27–30 maggio 1977*, 305–341. Rome: Bulzoni.

Matras, Yaron and Jeanette Sakel. 2007. Investigating the mechanisms of pattern replication in language convergence. *Studies in Language* 31: 829–865.

Maurer, Theodoro. 1968. *O infinito flexionado português. Estudo histórico-descritivo*. São Paulo: Editora Nacional.

McGregor, William. 1990. *A Functional Grammar of Gooniyandi*. Amsterdam: Benjamins.

Monod-Becquelin, Aurore. 1975. *La Pratique linguistique des indiens trumai* (*Haut-Xingu, Mato Grosso, Brésil*). Paris: Centre National de la Recherche Scientifique.

Morosi, Giuseppe. 1878. I dialetti romaici del Mandamento di Bova in Calabria. *Archivio Glottologico Italiano* 4: 1–110.

Ouali, Hamid. 2008. On C-to-T φ-feature transfer: The nature of agreement and anti-agreement in Berber. In Roberta D'Alessandro, Gunnar Hrafnbjargarson and Susann Fischer (eds.), *Agreement Restrictions*, 159–180. Berlin: Mouton.

Papageorgiadis, Ianni. N.D. *Grammatica degli idiomi del Sud Italia parlati in Grecìa salentina* (*Puglia*) *e in Calabria*. From the Grammar by A. Karanastasis published by the Academy of Athens in 1997. Web version in Italian. (Available at http://www.grecosuditalia.it, accessed 18/09/20).

Poletto, Cecilia and Raffaella Zanuttini. 2003. Making imperatives: Evidence from Central Rhaetoromance. In Christina Tortora (ed), *The Syntax of Italian Dialects*, 175–206. Oxford: Oxford University Press.

Portner, Paul and Raffaella Zanuttini. 2003. Decomposing imperatives. Paper presented at *IX Giornata di Dialettologia*. University of Padua, 26 June 2003.

Puyau, Jean-Marie. 2013. *Comprendre, parler, lire, écrire le béarnais*. Cressé: Éditions des Régionalismes.

Ralli, Angela. 2021. Contrasting Romance and Turkish as donor languages: Evidence from borrowing verbs in Modern Greek dialects. *Journal of Language Contact* 14(1).

Rivero, María Luisa. 1994a. Clause structure and V-movement in the languages of the Balkans. *Natural Language and Linguistic Theory* 12: 63–120.

Rivero, María Luisa. 1994b. Negation, imperatives and Wackernagel effects. *Rivista di Linguistica* 6: 39–66.

Rivero, María Luisa and Arhonto Terzi. 1995. Imperatives, V-movement and logical mood. *Journal of Linguistics* 31: 301–332.

Rohlfs, Gerhard. 1968. *Grammatica storica della lingua italiana e dei suoi dialetti. Morfologia*. Turin: Einaudi.

Rohlfs, Gerhard. 1977. *Grammatica storica dei dialetti italogreci*. Munich: C.H. Beck.

Silva-Villar, Luis. 1998. Morphology and syntax of Romance imperatives: An incomplete history. *Probus* 10: 115–138.

Smith, John Charles. 2005. How to do things without junk: The refunctionalization of a pronominal subsystem between Latin and Romance. In Jean-Pierre Montreuil (ed), *New Perspectives on Romance Linguistics: Selected Papers from the 35th Linguistic*

Symposium on Romance Languages (LSRL), Austin, Texas, February 2005. II: Phonetics, Phonology and Dialectology, 183–205. Amsterdam: Benjamins.

Smith, John Charles. 2011. Change and continuity in form-function relationships. In, Maiden Martin, John Charles Smith, and Adam Ledgeway (eds.), *The Cambridge History of the Romance Languages. Volume 1: Structures*, 268–317. Cambridge: Cambridge University Press.

Squillaci, Maria Olimpia. 2017. *When Greek Meets Romance: A Morphosyntactic Analysis of Language Contact in Aspromonte*. PhD dissertation, University of Cambridge.

Tekavčić, Pavao. 1972. *Grammatica storica dell'italiano*. Bologna: il Mulino.

Tomić, Olga M. 1999. Negation and imperatives. *Catalan Working Papers in Linguistics* 7: 199–206.

Tomić, Olga M. 2006. *Balkan Sprachbund Morpho-syntactic Features*. Dordrecht: Springer.

Tortora, Christina. 2014. *A Comparative Grammar of Borgomanerese*. Oxford, New York: Oxford University Press.

Tuite, Kevin. 1997. *Svan*. Munich: Lincom Europa.

Tuite, Kevin. 2018. The Svan Language. Unpublished ms., University of Montreal.

Violi, Filippo. 2004. *Grammatica sintassi e lessico della lingua grecocalabra*. Reggio Calabria: Iiriti Editore.

Zamfir, Dana-Mihaela. 2005–2007. *Morfologia verbului în dacoromâna veche (secolele al XVI-lea–al XVII-lea), I, II*. Bucharest: Editura Academiei Române.

Zanuttini, Raffaella. 1994. Speculations on negative imperatives. *Rivista di Linguistica* 6: 67–89.

Zanuttini, Raffaella. 1997. *Negation and Clausal Structure*. Oxford: Oxford University Press.

Zeijlstra, Hedde. 2006. The ban on true negative imperatives. In Oliver Bonami and Patricia Cabredo (eds.), *Empirical Issues in Syntax and Semantics* 6, 405–424.

BRILL

Contrasting Romance and Turkish as Source Languages: Evidence from Borrowing Verbs in Modern Greek Dialects

Angela Ralli
Emeritus Professor of Linguistics, Department of Philology, School of
Humanities, University of Patras, Patras, Greece
ralli@upatras.gr

Abstract

In this paper, I deal with verb borrowing in a language-contact situation involving Greek as target and Romance and Turkish as source languages. More particularly, I discuss the reasons and techniques that make verbs of typologically and genetically different languages to be accommodated in a uniform way within the same linguistic system, and verbs of the same donor to be integrated in a different manner within the same recipient. I try to provide an explanation for the observed divergences and similarities by appealing to an interplay of linguistic and extra-linguistic factors. For the purposes of this study, evidence is drawn from both written and oral sources from five Greek dialectal varieties: Grekanico, Heptanesian, Pontic, Aivaliot and Cypriot.

Keywords

verb borrowing – integration – Modern Greek dialects – Romance – Turkish

1   Introduction

Contact between languages, triggered by socio-political, economic and cultural relations, is, among other things, the cause of linguistic innovations due to matter and pattern replication.[1] Contact phenomena have attracted the interest of scholars in recent years, who try to describe and analyze them from different perspectives and various theoretical frameworks, as well as to establish the principles constraining the influence of the source on the target languages.

Greek,[2] throughout its long history, shows an interesting diversity of these phenomena, which is particularly witnessed on the dialectal level, since the current official language, that is, Standard Modern Greek, is not the real outcome of the evolution of Ancient Greek: Standard Modern Greek has been primarily based on Peloponnesian, the dialect spoken at the time of the formation of the first Greek state (1827), contains few elements of the prestigious dialects of the Ionian islands (see Fig. 1) and Constantinople (today's Istanbul), and during the last two centuries, it has been enriched with words and expressions
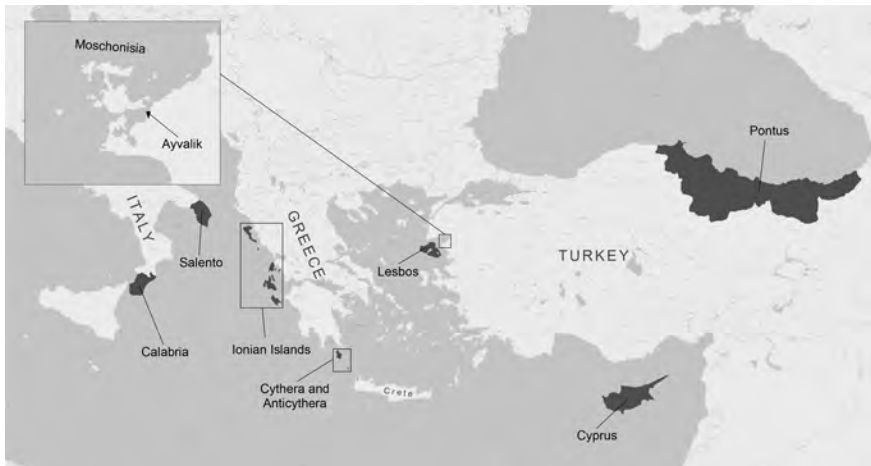
FIGURE 1:   Areas where the dialects examined in this article are spoken.

---

1   The terms "matter" and "pattern replication" for lexical and structural borrowing are due to Sakel (2007). See Gardani (2020a) for a typology on matter and pattern borrowing which goes beyond morphology.

2   In this article, "Greek" is used as a roof term referring to the language throughout its history. The other terms, "Ancient Greek" (fifth – fourth century BCE), "Hellenistic Koine" (ca. third century BCE – third century CE), "Modern Greek" (after the fifteenth century CE) and "Standard Modern Greek" (the current official language) refer to the linguistic systems of specific periods.

from Katharevousa, a high-style written language form, created for political purposes in the nineteenth century. Contact phenomena emerge in all Modern Greek dialects, the origin of which is usually considered to be the Hellenistic Koine,[3] and are observed on multiple linguistic levels – lexical, grammatical, phonological, semantic – and in heterogeneous communicative contexts. They are detected in the existing dialectal texts and in various oral narratives that have been collected in the last two centuries. Significantly, in the dialect geography of Modern Greek, two antithetical developments, divergence and convergence, produce changes in areas which have been under the influence of different languages.

In this article, I investigate five Greek dialects, Grekanico, Heptanesian, Pontic, Aivaliot and Cypriot, with the aim to determine phenomena and unexplored paths produced by contact between Greek and two genetically and typologically different languages, Romance on one side and Turkish on the other, depending on the case.

My purpose is to examine why in a contact situation involving Greek as target language and two different source languages, both genetically and typologically (Romance is Indo-European and semi-fusional while Turkish is Altaic and agglutinative), can produce not only divergent but also similar effects in Greek, which are mostly shown on lexical borrowing, identified as the commonest and most frequent type of transfer in contact situations (Haspelmath, 2009), and the morphological structure of the borrowed elements. More precisely, I focus on the domain of verbal loans that are usually considered to be among the most difficult items to be transferred from one language to another, due to the rich information they carry (see Wohlgemuth, 2009 for details). In contrast, nouns are generally seen as the easiest borrowable grammatical category (Whitney, 1881; Hock and Joseph, 1996),[4] due to their referential properties as Matras (2009: 168) sustains.[5]

In this work, I restrict my attention to the transfer of verbs, excluding other parts of speech (e.g., nouns), which have become verbs in the Greek dialects with the addition of a verbalizer. I try to investigate four general questions

---

3   Tsakonian is an exception, since for many scholars (see for instance Deffner, 1881), it derives from the Ancient Doric dialect.

4   A stronger view is expressed by Moravcsik (1978: 111), who claims that "verbs cannot be borrowed as such but must be borrowed as nouns and 'reverbalized' in the borrowing language".

5   Another explanation could be drawn from Seifart et al. (2018), who have observed a cross-linguistic tendency for slower speech before nouns, compared to speech occurring before verbs, defined in terms of pauses and slower articulation. While these authors do not account for lexical borrowing, one could argue that this property makes nouns easier to isolate and as a consequence to borrow. I owe this remark to Francesco Gardani (p.c.).

referring to: (a) the role of the source and the target languages; (b) the mechanisms and paths involved in verb borrowing and adaptation; (c) the factors and principles regulating the degree and type of verb integration; (d) the constraints applying to the occurrence of an integrating element.

The examined data are drawn from both written and oral sources. Among the written sources, there are available dictionaries, glossaries and grammars, such as: Karanastasis (1984; 1997) and Rohlfs (1933; 1977) for Grekanico, Soldatos (1967), Katsouda (2016) and Simiris (2017) for Heptanesian, Papadopoulos (1955: 1958–1961) and Oikonomidis (1958) for Pontic, Sakkaris (1940; 1948) and Ralli (2017) for Aivaliot, and Chatziioannou (1936) and Chatzipieris and Kapatas (2015) for Cypriot. The oral data are drawn from a digitized corpus consisting of about 200 hours of oral narratives, stored at the *Laboratory of Modern Greek Dialects* (LMGD, https://www.lmgd.philology.upatras.gr) of the University of Patras.

The article is structured as follows: after the introduction, Section 2 contains a brief overview of Grekanico, Heptanesian, Pontic and Aivaliot, accompanied by an indicative list of loan verbs and a number of remarks on the properties shown by these loans. A discussion of the four research questions mentioned above is provided in Section 3, where tentative answers are proposed, and the interplay of endogenous and exogenous factors in verb integration is pointed out. The Cypriot data provided in Section 4 serves as a testing bed to hypotheses and proposals, put forward in Section 3. The article ends with the conclusions (Section 5) and the relevant bibliography.

2   Dialects and Dialectal Data

2.1   *Grekaniko*

The Greek speaking dialectal enclaves in South Italy are located in Puglia (area of Salento, the so-called "Grecia Salentina") and South Calabria (Bovese area), as depicted in Fig. 1. In this article, the Greek dialect of South Italy will be called "Grekanico", used as a roof term for Greko, the Greek variety of Calabria, and Griko, the Greek variety of Salento. Other roof terms found in the literature are "Italiot" (e.g., Ralli, 2016) or "South Italian Greek" (e.g., Manolessou and Ralli, 2020).

Due to the long-term Italo-Romance rule, Grekanico has been affected by the Southern Italo-Romance dialects (Salentino or Calabrese, depending on the case), the official Italian (mainly from the second half of the twentieth century onwards), and a form of Regional Italian (Martino, 1979; Profili, 1983; Katsoyannou, 1995; 1999; Fanciullo, 2001; Manolessou, 2005; Squillaci,

2016).[6] Nowadays, Greko presents a rapid decrease and Katsoyannou (1995) reported that in mid-nineties there were no more than 500 native speakers left in Calabria, while several villages were deserted. In contrast, Griko seems to be resisting, although native competence has been rather confined to elderly people. According to Profili (1983), there were about nine Griko-speaking villages in the early eighties, where speakers communicated in the dialect mostly in family. Details about the socio-linguistic situation in the Greek-speaking areas of South Italy are given, among others, in Profili (1983), Telmon (1992), Katsoyannou (1995), Manolessou (2005) and Squillaci (2016).

The origin of Grekanico is a debated issue, and arguments pertaining to the "archaism" or "byzantinism" of this dialect are of historical and linguistic nature (Fanciullo, 2001; Manolessou, 2005). There are three different views: (a) the dialect is of Byzantine origin (among others, Morosi, 1870; Parlangeli, 1953); (b) Grekanico originates from the Ancient Greek of Magna Graecia (among others, Rohlfs, 1933; 1977; 1997; Caratzas, 1958; and Karanastasis, 1984); (c) it descends from the Hellenistic Koine, while it has been enriched with byzantinisms, especially Griko, due to Byzantine settlers in the area (Horrocks, 1997; Ledgeway, 1998; Fanciullo, 2001; Browning, 2004; Manolessou, 2005).

Griko and Greko display a number of differences (see, among others, Rohlfs, 1933; Caracausi, 1979; Karanastasis, 1997; Squillaci, 2016). However, these differences are not significant in order to consider Griko and Greko as different dialectal systems. Crucially, there is no divergence in the way the two varieties adopt Italo-Romance verbs.

In this section, I list some indicative Griko and Greko examples under (1).[7] All these examples as well as those of the other dialects are given in the citation form, that is, in the first person singular of the present tense (overtly realized infinitival forms have disappeared from Greek during the Hellenistic period, see Horrocks, 1997), and are transcribed in a broad phonological transcription. For clarity reasons, hyphens separate the stems from the verbalizer and inflection.

(1)   a.   Griko   Salentino
           bbund-e-o   'to abound'   *bbunn(d)-are*
           bbamp-e-o   'to go red'   *bbamp-are*

---

6   For the contact between Greko and Italo-Romance, see also the article by Ledgeway, Schifano and Silvestri (2021) on the formation of negative imperative forms in Southern Calabrian.

7   The Griko examples are drawn from the oral narratives stored at LMGD, which were recorded in 2000, during a research expedition to Salento, under the direction of Angela Ralli (European Project INTERREG II). I owe the Greko examples to Maria Olimpia Squillaci.

ffrunt-e-o	'to confront'	*ffrunt-are*
kunt-e-o	'to narrate'	*kunt-are*
mbest-e-o	'to guess'	*mbišt-ire*
nat-e-o	'to swim'	*nat-are*
skupr-e-o	'to discover'	*skupr-ire*

(1)   b.

	Greko		Calabrian
	bbamp-eggu-o	'to go red'	*bbamp-ari*
	nnat-eggu-o	'to swim'	*nnat-ari*
	pass-eggu-o	'to pass'	*pass-ari*
	pens-eggu-o	'to think'	*pens-ari*
	spend-eggu-o	'to spend'	*spend-iri*

As shown in (1), only the Italo-Romance stems are adopted by the Griko and Greko speakers, who hellenicize them with the use of an integrating element, the Greek verbalizer *-e(v)-*. In Griko, /v/ is phonologically deleted in the intervocalic position (for this, see Karanastasis, 1997), while in Greko, *-ev-* appears as *-eggu-*, originating from *-evɣ-* with the insertion of a /ɣ/, well known in other Modern Greek dialects too, e.g., in Cretan and Lesbian (Contossopoulos, 2001).

For Wichmann and Wohlgemuth (2008) and Wohlgemuth (2009), the accommodation of adopted verbs follows three insertion strategies, direct, indirect and light verb. For the direct strategy, the loan verb is plugged in the target language as it is, or with a slight phonological modification. The indirect strategy involves the presence of an integrating element – in the Grekanico case, the verbalizer *-e(v)-* –, while a light verb is employed for the implementation of the third strategy.[8]

According to what is commonly assumed in the literature, for lexical borrowing, languages borrow entire words (Thomason, 2001; Winford, 2003; Matras, 2009). As shown in Section 3, I suppose that word forms are transferred to Grekanico. Then, the speakers analyze them, retain only the stem, add the Greek verbalizer *-e(v)-* in order to produce a more hellenicized form, and the structure [stem + verbalizer] is further combined with the Greek inflectional endings.

2.2   *Heptanesian*

Heptanesian is the dialect of the Ionian islands including Corfu, Paxi, Cephalonia, Ithaca, Zante and the smaller islands of Othoni, Herikusa, Mathraki,

---

8   There is no application of the light-verb strategy to the integration of verbal loans in the dialects investigated in this article.

Antipaxi, as well as Cythera and Anticythera south of the Peloponnese (Fig. 1. See, among others, Salvanos, 1918; Soldatos, 1967; Hitiris, 1987; Konomos, 2003; Katsouda, 2016; Simiris, 2017).[9] Heptanesian displays several traits of Italo-Romance, that is, Venetian and Italian, due to the long-term domination of the islands by the Republic of Venice: the islands remained under Venetian rule from approximately the mid-fourteenth century (the exact dates vary, depending on the island[10]) until the turn of the nineteenth century, when they passed to British control, and ultimately to the Greek state in 1864.

There are some linguistic differences from island to island, but not substantial enough to challenge the linguistic unity of Heptanesian, which can be seen as an umbrella term capturing the features that are shared by all varieties.

As far as the existing written sources are concerned, beside the literary works, one can locate a plethora of glossaries, excerpts of old magazines, manuscripts, fairytales, personal stories and memoirs, recipes, folk songs, proverbs, administrative documents and correspondence, collections of civil acts, ecclesiastical and educational records (see Makri, 2020 for a detailed list). There are several works reporting contact between Heptanesian and Italo-Romance, but with some exceptions (for example Ralli, 2012a; Katsouda, 2016; Simiris, 2017), they focus on nouns (see, among others, Hitiris, 1987; Kahane and Kahane, 1982; Korosidou-Karra, 2003; Ralli et al., 2015; Krimpas, 2018; Makri, 2020).

In what follows, I provide a number of Italo-Romance verbal loans (a hyphen separates stems from inflectional endings), taken from Ralli (2012a), Katsouda (2016) and Simiris (2017).[11] It is worth mentioning that it is often difficult to discern whether a verb has been transferred from Venetian or Italian, because, as noted by Fanciullo (2008), there was a kind of diglossia in the Republic of Venice from the sixteenth century onwards, where Venetian was mainly employed in the everyday communication, while Italian for administrative purposes. This diglossic situation was also transferred to areas under Venetian rule, among which, the Ionian islands (Ralli, 2019a).

---

9   The inhabitants of Leukada, another big Ionian island, speak a dialect which shares
     similarities with the group of Northern Greek dialects, mainly due to the proximity of the
     island to the Greek mainland, but also because Leukada was under Venetian rule for a
     shorter period of time, as Venice conquered the island in 1664 (Contossopoulos, 2001: 67).
     For example, it displays the phenomena of high-vowel deletion and mid-vowel raising in
     unstressed position (see footnote [17] for details).

10   Corfu underwent the Venetian domination as early as in 1387, while the other Heptanesian-
     speaking islands were taken by Venice during the 15th century.

11   In (2), the It(alian) origin is noted when a Venetian form is not attested in Boerio's (1856) and
     Cortelazzo's (2017) dictionaries of the Venetian dialect.

(2)   Heptanesian   Venetian
        abandoner-o   'to abandon'   *abandon-ar*
        akompaɲar-o   'to accompany'   *acompagn-ar*
        amolar-o   'to free'   *amol-ir*
        arivar-o   'to arrive'   *ariv-ar*
        γoðer-o   'to enjoy'   *god-er*
        (i)bitsilir-o   'to render imbe-   It. *imbecill-ire*
                             cile'
        (i)mitar-o   'to imitate'   It. *imit-are*
        kojonar-o   'to make fun of'   *cogion-ar*
        krepar-o   'to crack'   *crep-ar*
        lustrar-o   'to shine'   *lustr-ar*
        tratar-o   'to serve'   *trat-ar*   'to deal with, to
                                                                   treat'
        vatsinar-o[12]   'to vaccinate'   *vacin-ar*

A first examination of the Heptanesian verbal loanwords reveals that they are based on the entire Italo-Romance infinitival forms -they involve the infinitival marker *-ar(e)-* and appear to be fully integrated in the Heptanesian system, being regularly inflected with the addition of the Greek inflectional endings.[13] This distinguishes them from the Grekanico loans, where no infinitival mark is part of the adopted Italo-Romance verb. Having adopted full-word forms, Heptanesian does not employ any specific integrating element, since, as stated by Wohlgemuth (2009: 87–92), the inflectional endings do not count as such, being compulsory in the recipient's system, as is the case with Greek morphology. Italo-Romance verbal loans in Heptanesian obey the direct insertion strategy, contrary to those in Grekanico (1), which follow the indirect insertion strategy. Thus, an interesting question which arises is why verbs of the same source language are integrated differently in the same target language.

2.3   *Pontic*

Pontic belongs to the inner Asia Minor dialectal group (Manolessou, 2019). It was spoken in a geographical area of about 400 kilometers (from Inepolis to Colchis) in the northeast part of Asia Minor (Black Sea coast of Turkey), and in

---

12   Francesco Gardani (p.c.) has pointed out that this must have been one of the latest Venetian loans, since *vaccinus* in its current meaning is not attested before the first decade of the nineteenth century.

13   Note that the same integration of Italo-Romance verbs is observed in other Modern Greek Dialects that have been in contact with Venetian, as for instance in the dialect of Crete (Chairetakis, 2020).

parts of the inland of Asia Minor, 100 kilometers from the coast (Oikonomidis, 1958), as shown in Fig. 1. In the nineteenth century, a number of Pontic communities settled in Ukraine and Georgia, and a massive movement of Pontics, principally to mainland Greece, took place under the Lausanne Treaty (July 1923), which led to an exchange of Muslim and Christian Orthodox people between Greece and Turkey. Since then, the dialect has been used by second and/or third generation refugees in Greece (according to Drettas, 1997: 15, there are about 300,000 speakers), but can also be found in Georgia, Ukraine (Mariupol area) and in Rostov-on-Don of the Russian Federation (see, among others, Berikashvili, 2017). Moreover, a variety of Pontic is still spoken in certain dialectal enclaves in Turkey, namely in the Western part of Trebizond (Tonya, Sürmene and the valley of Of), by Muslim Pontics, who were exempted from the population exchange for religious reasons. Their dialect is usually called "Muslim Pontic" or simply "Romeyka" (Mackridge, 1987; Sitaridou, 2013; Özkan, 2013).

The Pontus area came under the Ottoman rule in 1461, but the Greek-speaking people resisted to massive islamization and shifting to Turkish (among others, Vryonis, 1971; Bryer, 1975; Kitromilides and Alexandris, 1984). According to Manolessou (2019: 34), there is a sufficient number of written sources in Pontic (see Henrich, 1990; 2011 for the Medieval Pontic texts), compared to the existing sources of other Modern Greek dialects.

Pontic is divided into Western and Eastern, the two varieties displaying some differences but also sharing striking similarities (see Papadopoulos, 1955 and Oikonomidis, 1958 for a detailed account of this division). The examples listed in (3) are taken from Papadopoulos (1958–1961) and from 30 hours oral sources, stored at LMGD.[14] The Turkish examples are given in the *-mAk* infinitival form, while the loan stems in Pontic are separated from the Greek/Pontic verbalizer *-ev-* and the personal ending *-o* (first person singular of the present tense).

(3)   Pontic   Turkish
       xazirla-ev-o   'to prepare'   *hazırla-mak*   'to get ready'[15]
       tokun-ev-o   'to insult'   *dokun-mak*
       ta(γ)ut-ev-o   'to scatter/disperse'   *dağıt-mak*
       γazan-ev-o   'to earn'   *kazan-mak*
       axtar-ev-o   'to overturn/transfer'   *aktar-mak*

---

14   The oral material was collected under the direction of Angela Ralli, within the framework of the THALIS project (2012–2015), funded by ESF and the Greek Ministry of Education.

15   The Turkish original word is translated only when its meaning (slightly) differs from that in Pontic.

pašla-ev-o	'to begin'	*başla-mak*
γurtar-ev-o	'to free/save'	*kurtar-mak*
tajan-ev-o	'to be patient'	*dayan-mak*
konuš-ev-o	'to talk'	*konuş-mak*
pekle-ev-o	'to wait'	*bekle-mek*
šašir-ev-o	'to be surprised'	*şaşır-mak*
γantur-ev-o	'to trick'	*kandır-mak*

These examples show that for adopting Turkish verbs, Pontic retains only the stem and applies the indirect insertion strategy with the use of the derivational suffix -*ev*-. It is worth stressing that Italo-Romance loans in Grekanico (1) follow the same path as Turkish loans in Pontic (3): both dialects adopt the stem from the source language (Italo-Romance or Turkish, depending on the dialect), and employ the same integrator, that is, the Greek verbalizer -*ev*-. Given the difference in the source language, that is, Turkish for Pontic and Italo-Romance for Grekanico, this is an interesting point that should be explored.

2.4   *Aivaliot*

Aivaliot, another Asia Minor dialect (Ralli, 2019b), was spoken until 1922, in the town of Kydonies or commonly Aivali (present day Ayvalık), on the Western coast of Turkey (Edremit gulf, four to five miles from the Aegean Greek island of Lesbos, see Fig. 1). A slightly different variety, Moschonisiot, was the language of the inhabitants of the nearby islands of Moschonisia (nowadays Cunda), while a variant of Aivaliot is still found on the Aegean Turkish island Bozcaada (Imbros in Greek). Before the First World War, Aivali had about 30,000 Greek-speaking residents, while Moschonisia counted circa 15,000 people. Moschonisia and Aivali were deserted from their Greek speaking population in September 1922, after the end of the Greek-Turkish war and several months before the Lausanne treaty. Aivaliots and Moschonisiots who escaped killing flew to Greece, principally to Lesbos, while a number of them moved to other countries as well, e.g., France, North America and Australia.

The dialect has originally emerged following a settlement of colonists from the island of Lesbos, around the late sixteenth or early seventeenth centuries. Today, there is an estimated number of few hundreds of Aivaliot speakers, descendants from first-generation refugees, most of them living in Lesbos, who often mix their own dialectal variety with the parent Lesbian.[16]

Aivaliot and Moschonisiot do not display significant differences to constitute different dialects. Thus, in this article, the term "Aivaliot" will be used for

---

16   See Sakkaris (1920) and Ralli (2019b) for a detailed history of the area and the dialect.

both varieties. The dialect belongs to the group of Northern Greek dialects[17] and has been heavily influenced by Turkish, mainly on the lexical and morphological levels (see Ralli, 2012a; 2016; 2019b; Ralli et al., 2015 for a detailed presentation). It has adopted many lexical items from Turkish, especially nouns and verbs. For verbs, as shown in (4) with the examples drawn from Ralli (2012b), the transferred item is not the infinitival form ending in -*mAk*, but the third person singular of the Turkish past tense ending in -*dI* (given in parenthesis).[18] Hyphens separate the loan items from an optional Greek verbalizer -*iz*- and the compulsory inflectional suffix -*o* or -*u* (unstressed –*o*, as explained in footnote [17]).

(4)   Aivaliot[19]   Turkish

burd-íz-u	'to twist'	*bur-mak*	(*bur-du*)	
dald-íz-u	'to be absent-minded'	*dal-mak*	(*dal-dı*)	'to dive, plunge'
kudurd-íz-u	'to be very active' (pej.)	*kudur-mak*	(*kudur-du*)	'to go mad'
kazad-íz-u / *kazad-ó*	'to earn, become rich'	*kazan-mak*	(*kazan-dı*)	
furlad-íz-u	'to burn from anger'	*fırla-mak*	(*fırla-dı*)	'to dash, pop up'
zurlad-íz-u	'to force, stretch'	*zorla-mak*	(*zorladı*)	
katsird-íz-u / *katsird-ó*	'to escape'	*kaçır-mak*	(*kaçırdı*)	'take away, kidnap'
axtard-íz-u / *axtard-ó*	'to throw, overturn'	*aktar-mak*	(*aktar-dı*)	'to transfer, mix'

---

17   The Modern Greek dialects are divided into northern and southern on the basis of a high-vowel loss and a mid-vowel raising in unstressed position (Chatzidakis, 1905–1907; Newton, 1972; Contossopoulos, 2001; Trudgill, 2003). For example, /fegári/ 'moon' becomes [figár], /xoráfi / 'field' [xuráf] and /kutí/ 'box' [kti].

18   In (4), the infinitival ending and the past tense (aorist) ending are noted with the capital letters A and I, respectively, because Turkish is subject to a vowel-harmony law, according to which in consecutive syllables, the vowel of the second syllable is changed conforming to the vowel of the preceding syllable. Thus, a verb like 'to love' has an infinitival form in -*mek* (*sevmek*), and a third person aorist form in -*di* (*sevdi*), while the corresponding forms of the verb 'to earn, profit' are *kazanma*k and *kazandı*. As a contact phenomenon, vowel harmony is observed in the Asia Minor Cappadocian dialect (Dawkins, 1916). However, it is not a transferred feature in Aivaliot and Pontic.

19   Stress appears on all Aivaliot verbs because of the Northern Greek vocalism which deletes unstressed /i/ and /u/ and raises unstressed /e/ and /o/ (see footnote [17] for details).

sacind-íz-u /	'to stand back/	*sakın-mak*	(*sakın-dı*)	'beware,
sacind-ó	aside'			avoid'
dajad-ó	'to bear, endure'	*dayan-mak*	(*dayan-dı*)	
savurd-ó	'to throw'	*savur-mak*	(*savur-dı*)	
sasird-íz-u /	'to be at a loss'	*şaşır-mak*	(*şaşır-dı*)	'to wonder,
sasird-ó				be at a loss'

Similarly to all Greek native verbs, the loans receive inflection, with the addition of Greek inflectional endings. As mentioned in Section 2.2, the inflectional ending could not be considered as integrator, since its presence is compulsory in the recipient system, and its addition to the Turkish lexeme is done by default.

As opposed to Pontic (3), which also borrows Turkish verbs but retains only the stem, Aivaliot adopts the *-dI* third person singular of the past tense and most of the times adds the Greek verbalizer *-iz-* as integrating element, obeying the indirect insertion strategy. The particular choice of integrator is another difference with Pontic, where, as observed in (3), the integrator is *-ev-*. Moreover, contrary to the other three dialects, Aivaliot adopts Turkish verbs by following two insertion strategies which, in some cases, seem to function in parallel: the indirect strategy with the use of *-iz-*, and the direct one, without the presence of any verbalizer. The question is, thus, why there is such divergence, creating doublets in *-dizu* and *-do* (e.g., *kazadízu / kazadó* 'to earn').[20]

Interestingly, the direct strategy has been attested in Heptanesian (2) with respect to Italo-Romance loans (e.g., *trataro* 'to treat' < Venetian *tratar*). However, the two dialects differ as far as the point of departure of the adopted material is concerned, in that, Heptanesian loans are transferred from the infinitival form, while the Aivaliot ones are based on a finite form (third person singular of the past tense), the latter being ultimately reanalyzed as stem in order to receive the optional Greek verbalizer *-iz-* and the compulsory Greek inflectional suffix. One may wonder whether the type of the source language plays a role in this choice. Such role is debatable though, given the fact that for both Aivaliot and Pontic the source language is the same, that is, Turkish, but Pontic adopts only the verbal stem and hellenicizes it with the verbalizer *-ev-*, while Aivaliot keeps the entire form of the third person singular of the past tense and optionally uses the verbalizer *-iz-*.

---

20   It should be noted that, in Aivaliot, as well as in the other dialects examined in this article, there are verbs of Turkish or Italo-Romance origin which may display a different derivational suffix, for instance *-on-*. However, these loans are derived structures on the basis of other parts of speech, for example nouns (e.g., Aivaliot *batakónu* 'sink in the mud' < Turkish *batak* 'swamp, mud' + Greek verbalizer *-on-* + Greek 1SG), and do not belong to direct verb transfer. Thus, they are excluded from my investigation.

3   Loan Verb Integration

On the basis of what has been seen so far, the overall typology of the source language, that is, semi-fusional or agglutinative, does not seem to determine the strategy and the pattern that are followed for adopting and accommodating verbs, as well as the choice of a particular integrating element. According to the "anything goes" position (among others, Thomason, 2001), anything can be borrowed as long as there is heavy socio-cultural contact. In fact, the long-term Italo-Romance or Turkish domination of the lands where the four dialects are/were spoken could justify the abundance of borrowed verbs. However, the extra-linguistic factors alone are not sufficient to explain the specific way of verb accommodation. Thus, although socio-linguistics play an important role in verb borrowing and integration, additional reasons must be searched in the linguistic factors, more particularly in the linguistic properties of the languages in contact.

As proposed by Ralli (2005; 2015), Greek is a stem-based language, where an inflectional ending is obligatorily added to nominal and verbal stems in order for them to become inflected words. That is, inflected words obey the [[Stem-(Der)]-Infl] structural pattern, where the stem can be derived or non-derived. This property is crucial for the borrowing of nominal and verbal items, since, with the only exception of a small amount of English and French borrowed nouns (see Ralli and Makri, 2020), the adopted elements are treated as stems, undergo small phonological adjustments dictated by the Greek phonological system, and are ultimately combined with overt Greek inflectional suffixes, while some of them include a derivational suffix too. For an illustration, see the examples in (5–8), taken from the dialects under investigation, where only Pontic and Grekanico seem to select the bare stem. All the other items are adopted from Italo-Romance and Turkish as word forms, which are further reanalyzed as stems.

(5)   Grekanico   Italo-Romance
     a.   *masari-(s)*[21]   *massar-o*
            estate.manager-NOM.SG
            'estate manager'

---

21   The final -s of the nominative singular (citation form) is deleted because Grekanico, influenced by Italo-Romance, does not allow closed syllables at the end of words. See Karanastasis (1997) and Manolessou and Ralli (2020) for details.

	b.	*čec-e-o*[22]	*čik-are*
		to blind-DER-1SG	'to blind'
		'I blind'	

(6)		Heptanesian	Italo-Romance (Venetian)
	a.	*pitor-os*	*pitòr*
		painter-NOM.SG	
		'painter'	
	b.	*tratar-o*	*trat-ar*
		serve-1SG	'to treat, deal with'
		'I serve'	

(7)		Pontic	Turkish
	a.	*tsopan-os*	*çoban*
		shepherd-NOM.SG	
		'shepherd'	
	b.	*tokun-ev-o*	*dokun-mak*
		to insult-DER-1SG	'to insult'
		'I insult'	

(8)		Aivaliot	Turkish
	a.	*paralí-s*	*para-lı*
		wealthy.man-NOM.SG	
		'wealthy man'	
	b.	*axtard-íz-u*	*aktar-dı*
		overturn-DER-1SG	'(s)he overturned'
		'I overturn'	

Ralli (2005; 2015) has shown that the stem-based property of Greek morphology affects all word-formation processes in Greek, that is, compounding and derivation, as well as inflection. Given the fact that most borrowed items are molded as stems to fit in the Greek morphological system, they adjust either to a type of stem-based derived material bearing an overt derivational suffix (see verbs in Grekanico (1, 5), Pontic (3, 7) and Aivaliot (4, 8)), or to non-derived stem-based inflected forms (see nouns and verbs in Heptanesian (2, 6) and verbs in Aivaliot (4)). Therefore, the requirement of Greek morphology to have stem-based inflected items, derived or non-derived, is fulfilled.

The question that needs an answer now is why Grekanico and Pontic speakers retain only the stems from Italo-Romance and Turkish adopted verbs, by

---

22 The vowel of the first syllable is changed into [e] because of the neighboring /i/, triggered by a Greek phonological law of vowel assimilation. This law is different from the vowel harmony observed in Turkish (see footnote [18]).

subtracting inflection from the borrowed material and replacing it with the Greek suffixes (see (1) and (3)), whereas Heptanesian and Aivaliot ones reanalyze as stems full word-forms (see (2) and (4)). Turning back to socio-linguistic factors, I believe that this discrepancy could not be interpreted by appealing to linguistic properties alone. In accordance with Ralli (2016), I would like to propose that both the Grekanico and Pontic speakers who were subject to a long-term dominance by Italo-Romance and Turkish, respectively, and generally had a good command of those languages, have reacted against a simple and unelaborated borrowing of verbs by creating more hellenicized forms, those carrying Greek derivation (the -ev- suffix) and inflection.[23] Significant support to this process was also provided by their traditionally known conservatism towards innovation: Grekanico and Pontic are well known for preserving a considerable number of archaic (Ancient Greek or Medieval) features, as shown by Rohlfs (1933; 1977; 1997), Caratzas (1958) and Karanastasis (1997) for Grekanico, Manolessou and Pantelidis (2011) and Sitaridou (2014) for Pontic. These elements have led scholars to consider them as the two Modern Greek dialects closer to Ancient Greek.[24] In fact, the verbalizer -ev-, the presence of which is attested in verbal loans, belongs to the Ancient Greek features: -ev- (-ευ- in Ancient Greek) productively created verb stems in the classical period (fifth-fourth century BCE), as asserted by Chantraine (1945: 244), and is still the most productively employed verbalizer in both Grekanico and Pontic. It is less frequent in Standard Modern Greek and the other Modern Greek dialects, compared to another verbalizer, -iz-, which has gained its vast productivity after the Hellenistic period (Browning, 2004: 92–93). On the basis of these considerations, I would like to conclude that the outcome of verb integration in Grekanico and Pontic has been shaped by the interplay of both intra-linguistic and extra-linguistic factors. On the one hand, the speakers' good command of the source language and their reaction to heavy borrowing facilitate but also constrain the type of transferred material. On the other hand, a decisive role is played by the morphological properties of the target language, since both Grekanico and Pontic retain only the stem from the borrowed words -stems being the base of all native verbs- and select the highly productive derivational suffix -ev- as integrating element.

---

23   Enrique-Arias (2010: 97) has reached a more or less similar conclusion for a contact situation involving Spanish and Catalan in Majorca.

24   For instance, the pronunciation of 'η' as [ɛ:] in Pontic (Manolessou and Pantelidis 2011) and the use of non-finite forms after the verb 'can' (Squillaci 2016) are among the preserved features. Interestingly, Manolessou (2005: 117) claims that the archaisms, shown in Grekanico on all grammatical levels, phonology and vocabulary, may be due to the fact that communication between South Italy and the rest of the Greek-speaking world was interrupted in the Middle Ages, before the thirteenth century CE.

As shown in the following paragraphs, the close interaction of intra-linguistic and extra-linguistic factors is not limited to Grekanico and Pontic but applies to verb borrowing and integration in all dialects under consideration. For instance, the low command of Turkish and the socio-political conditions in the Aivali area can explain why Aivaliot speakers do not analyze the Turkish adopted verbs in order to keep the stem, but reanalyze as stems the entire word forms. At the beginning of the twentieth century, as reported in the literature (Sakkaris, 1920), Aivali was inhabited by an entirely Greek-speaking population, had a kind of socio-political autonomy, granted by the Ottoman authorities, and education was provided in Greek, while Turkish or French were taught at school as second language. Women did not usually speak Turkish and those men who knew Turkish used it in trade and administration. Therefore, living in a mainly Greek-speaking environment, Aivaliot speakers borrowed the entire verbal words from Turkish, without feeling the need, or being able to further analyze their structure.[25] Moreover, they applied the indirect insertion strategy to accommodate the verbal loans with the use of -iz-, instead of that of -ev- observed in Pontic (3). I believe that a possible justification for the choice of the particular integrator can be found in the high frequency of -iz- in this dialect, which surmounts that of -ev-.[26]

However, an explanation is still pending as to why Aivaliot speakers resort to the adoption of the third person singular form of the Turkish past perfective tense.[27] Along the lines of Ralli (2012b; 2016; 2019b) and Bağrıaçık et al.

25　This socio-linguistic explanation could not apply to Cappadocian, where Turkish verbal loans are integrated exactly like the Aivaliot ones (Dawkins, 1916), in spite of the fact that Cappadocians had been under the heaviest contact with Turkish, from all Asia Minor Greek speaking populations, and had the highest command of the Turkish language. A different explanation is needed for Cappadocian verb accommodation which exceeds the limits of this article. Nevertheless, I would like to suggest that Cappadocian speakers do not analyze the Turkish past perfective forms in order to keep only the stem because, contrary to the other Greek speakers, they base word formation (among which loan verb integration) on entire words and not on stems (see Ralli, 2009 for an argumentation and examples on this position).

26　For details on the rise of productivity and the use of -iz- in several dialects, see Chatzidakis (1905–1907) and Browning (2004), although there are no available statistics.

27　The use of a past perfective form seems to be a generalized process across the Balkan languages (see Breu, 1991a; 1991b). However, instead of attributing this feature to a mutual spread in the linguistic area of the Balkans that is, a feature of the so-called "Balkan Sprachbund", I propose that this follows from contact with the Greek language, where the use of the past perfective stem as base for word-formation purposes is already attested in the Hellenistic period (Chatzidakis, 1905–1907), that is, long before the creation of the Balkan Sprachbund. Therefore, I agree with Gardani, Loporcaro, and Giudici (2021) who claim that the processes underlying some basic developments of the Balkan languages do not belong to mutual convergence but to borrowing and reanalysis.

(2015), I assume that among the most salient properties which can constrain lexical borrowing is a pre-existing similarity between the morphological features of the source and those of the target language.[28] In this spirit, the choice of the most productively used -*iz*- in Aivaliot is also facilitated by the fact that there are Greek native verbal stems ending in -*i* in the past perfective tense (aorist), something which renders them close to Turkish verbal forms in -*I*, as the examples in (9) illustrate.[29] These verbs are distributed in two categories/ classes: the first inflection class/conjugation containing verbs with the verbalizer -*i(z)*- (e.g., *cerð-iz-o* 'to win' in Standard Modern Greek, *cirðízu* in Aivaliot with mid-vowel raising, see footnote 17), and the second inflection class/conjugation (e.g., *aɣapó* 'to love') consisting of verbs originating from the Ancient Greek 'contract' ones,[30] any inflectional difference between the two classes being neutralized in the aorist paradigm.

(9)   Past tense (aorist) of the Aivaliot native verbs *cirðízu* 'to gain, win' and *aɣapó* 'to love', and the Turkish verb *sevmek* 'to love'

	a. Aivaliot		b. Turkish
1SG	*cérð(i)-sa*[31]	*aɣáp(i)-sa*	*sevdi-m*
2SG	*cérð(i)-sis*	*aɣáp(i)-sis*	*sevdi-n*
3SG	*cérð(i)-si*	*aɣáp(i)-si*	*sevdi*
1PL	*cirðí-sami*	*aɣapí-sami*	*sevdi-k*
2PL	*cirðí-sati*	*aɣapí-sati*	*sevdi-niz*
3PL	*cirðí-san*	*aɣapí-san*	*sevdi-ler*

---

28   According to Meillet's (1921) "retentionist" position, in a language-contact situation, a transfer of morphological features is feasible if source and target languages share the same morphology (see also Gardani, 2020b and Gardani et al., 2015 for relevant discussion), a claim that has been reformulated as "morphological congruence" by Myers-Scotton (2002) and Field (2002). A weakened view of this position has been put forward by Jakobson (1938) who rejects the idea of "overall identity" and speaks about "morphological tendencies".

29   Note that /I/ in Turkish can change according to the vowel harmony law. For instance, *burdI* '(s)he twisted' becomes *burdu, dayandI* '(s)he endured' *dayandı*, and *sevdI* '(s)he loved' *sevdi*. As already mentioned in footnote [18], vowel harmony does not exist in Aivaliot and /I/ is usually realized as [i].

30   Ancient contract verbs had a stem final vowel /ā/, /ɛ/, or /o/, which was fused with the initial vowel of the inflectional ending by the so-called "contraction" phonological law (e.g., *aɣapā+ō -> aɣapō* 'I love'). In the Hellenistic period, this law had already disappeared from the phonological system of Greek.

31   In the singular number, the unstressed /e/ has become [i] and the stem final vowel /i/ is put in parenthesis because it is deleted also in unstressed position (see footnote [17]). /i/ appears in plural because the stress is shifted on it. Note also that the initial -*s*- of the ending is the marker of the past perfective and -*a, -is, -i, -ame, -ate, -an*, the personal endings indicating the features of past, person and number. For clarity reasons, -*s*- and the personal endings are taken together. For a detailed analysis of Greek inflection, see Ralli (2005).

This form similarity has most probably led the Aivaliot speakers to adopt the Turkish past tense forms, reanalyze them as stems and further combine them with the Greek inflectional endings, as in (10).

(10)   Adoption and integration of the Turkish verb form *kazandı* '(s)he earned'

	Aorist (past perfective tense)
1SG	*kazád(i)-sa*
2SG	*kazád(i)-sis*
3SG	*kazád(i)-si*
1PL	*kazadí-sami*
2PL	*kazadí-sati*
3PL	*kazadí-san*

Moreover, the selection of the third person could be triggered by the fact that, in the Turkish past tense paradigm, this was the only form with no overt inflectional ending (see 9b), and thus, the easiest form to be adopted, being the most unmarked one. Besides, as claimed in the literature, the third person singular is prone to become the base for morphological changes (Joseph, 1998: 368).

Once the formation of the past tense is achieved, the rest of the verbal paradigm, that is, the personal forms of the present tense, imperfect and future tense, is shaped, either with the help of the verbalizer *-iz-* (analogically to the native Greek verbs in *-iz-* e.g., *cerðízo* 'to gain, win' in (9a)) or without it (following the native verbs of the second inflection class, e.g., *aɣapó* 'to love' in (9a)). Thus, the use of two alternative strategies, the direct and the indirect one, for the same loan can be justified, but the particular choice of the first or the second strategy, or even that of both strategies, seems to be at random. For an illustration, see the paradigms of the present tense in (11).

(11)   Aivaliot present tense of the parallel integrated forms of the Turkish verb *aktarmak* 'to transfer' ('overthrow' in Aivaliot)

1SG	*axtardízu*	*axtardó*
2SG	*axtardɟjs*[32]	*axtardás*
3SG	*axtardíz(i)*	*axtardá*
1PL	*axtadízumi*	*axtardúmi*
2PL	*axtardíziti*	*axtardúti*
3PL	*axtardízin*	*axtardún*

---

32   The forms *axtardɟjs* and *axtardíz* phonologically derive from *axtardízis* and *axtardízi* after the deletion of unstressed final /i/. However, unstressed /i/ is kept in the plural number, probably for reasons due to the ease of pronunciation.

Leaving aside Aivaliot, a point that needs to be clarified with respect to the Italo-Romance influence on Greek is why Heptanesian speakers adopt the entire Italo-Romance infinitival forms. Again, socio-linguistic reasons, together with the endogenous Greek morphological properties, could elucidate this behavior. Salvanos (1918) reports that, during the Venetian regime, Venetian and Italian were the dominant linguistic systems of the upper class in the Ionian islands, while people of the lower classes had small command of Italo-Romance and kept communicating in Greek. Crucial support to this observation is the statistic figures of 1849, provided by Soldatos (1967: 100) for Corfu, the Ionian island with the heaviest Italo-Romance influence, where only 6,000 speakers were bilingual, from a total of 200,000 inhabitants who spoke Greek, while only 100 people were reported to speak exclusively Italo-Romance. Compared to the Italo-Romance linguistic skills of Grekanico speakers, those of Heptanesian speakers were definitely lower. In accordance with what I have suggested for the Aivaliot speakers, this may explain why Heptanesians accept the entire word forms without resorting to their internal analysis, while the Grekanico speakers proceed to the analysis of borrowed words. In addition, while both Heptanesian and Grekanico had been under Italo-Romance influence, the socio-linguistic situation around Heptanesian was different from that of Grekanico, in that Heptanesian enjoyed a high prestige within the Greek-speaking world: it was used in literature and became one of the basic dialects for the development of Standard Modern Greek (Ralli, 2013). On the contrary, Grekanico was considered as a linguistic variety spoken by a lower class of peasants, in poor and isolated areas of South Italy (Katsoyannou, 1999). I believe that, as opposed to the Grekanico speakers who felt that their language was endangered and by reaction opted for a high hellenicization of verbal loans, Heptanesians did not probably sense the need to analyze the Italo-Romance words and imported infinitives as a whole.

A further point requiring elucidation is why Aivaliot speakers do not mold their verbal loans on the basis of an infinitival form, as Heptanesian speakers do with respect to Italo-Romance verbs. Assuming that Chatzidakis (1905–1907) is right, in the Hellenistic period, the aorist (past perfective) stem started being used as a base for the formation of verbal derivatives and this became a frequent tendency of Greek morphology across centuries.[33] Since the integration of a verbal loan with the help of a derivational suffix (-*iz*-) could be considered as a kind of derivational process, this may explain why the form of Turkish verbs that is adopted and further reanalyzed as stem is that of the

---

33   Chatzidakis's proposal has been embraced by many linguists dealing with Greek. See, among others, Janse (2004).

Turkish *-dI* past tense. In the same spirit, even the Turkish verbal forms that are borrowed in Pontic (3) could be those in *-dI* of the past perfective tense, and not the infinitival types in *-mAk*, because structurally, there is no formal difference between the Turkish stem in the infinitive and that in the past tense.

(12)	Pontic		Turkish infinitive	Turkish past tense (3SG)
	tokun-ev-o	'to insult'	*dokun-mak*	*dokun-du*
	γazan-ev-o	'to earn'	*kazan-mak*	*kazan-dı*
	axtar-ev-o	'to overturn/transfer'	*aktar-mak*	*aktar-dı*
	γurtar-ev-o	'to free/save'	*kurtar-mak*	*kurtar-dı*
	taǰan-ev-o	'to be patient'	*dayan-mak*	*dayan-dı*
	pekle-ev-o	'to wait'	*bekle-mek*	*bekle-di*
	šašir-ev-o	'to be surprised'	*şaşır-mak*	*şaşır-dı*

The problem, however, remains why Heptanesian speakers do not borrow a past perfective form of Italo-Romance verbs -like the Aivaliot speakers do- but select the infinitival forms. A solution could be found in the matching of features and structures between the source and the target language, that is, between Venetian and Heptanesian. According to Gambino (2007: cv) in old Venetian, the commonly employed form for the past perfective was the so-called "passato remoto". Its predominant use gave gradually place to the periphrastic passato prossimo, which became frequent only around the sixteenth century (Skubic, 1986: 31–43), and ultimately limited passato remoto in narrative contexts, around the nineteenth century (Loporcaro, 2013). Note now that the third person singular of the Venetian passato remoto, which could serve as the base for the integration of verbal loans in Heptanesian, was not uniform: its form varied depending on the verbal inflection class (Stussi, 1965: LXVIII), ending in *-a* (e.g., *cercà* '(s)he searched'), *-o* (e.g., *tochò* '(s)he touched'), *-e* (e.g., *vendè* '(s)he sold'), or *-i* (e.g., *morì* '(s)he died'). I believe that this variance rendered difficult the matching of forms between the Heptanesian past perfective and the Venetian one.[34] The absence of uniformity of the past perfective Venetian forms, as well as full access to infinitives that were productively used in Venetian probably led the Heptanesian speakers to adopt the most unmarked infinitival types.[35]

---

34 In the same spirit, even the periphrastic passato prossimo, built with the auxiliary *avere* 'to have' or *essere* 'to be' and the past participle of the main verb, could not be a suitable model for accommodating the Italo-Romance verbs. For an overview of simple and periphrastic past tenses in Romance languages, see Squartini and Bertinetto (2000).

35 It is worth mentioning that the borrowing of Italo-Romance infinitival forms in Greek had an impact on its morphology: it led to the innovative creation of a derivational suffix

The last question needing an answer concerns the original verbal form transferred in Grekanico: is it the infinitive, like in Heptanesian, or the third person singular of the past tense as in Aivaliot (or even in Pontic)? Assuming that overt infinitives do not exist in Southern Italian, as mentioned in several works (see, among others, Rohlfs, 1977; 1997;[36] Ledgeway, 1998; Squillaci, 2016), and according to the view that the one-word past perfective stems serve as the base for building derived structures in Greek (see above), one may suggest that, before resorting to an analysis in order to retain the stem, the Grekanico speakers adopt the past perfective forms, that is, the one-word forms of passato remoto, since the passato prossimo periphrastic ones have been recently inserted in the dialect, as suggested by Squillaci (2016: 62–74). See the following indicative examples taken from Greko (Squillaci, p.c.).

(13)   Greko   Calabrian (Infinitive)   Calabrian (Passato remoto 3SG)

nnat-eggu-o	'to swim'	nnat-ari	nnat-a-u
spend-eggu-o	'to spend'	spend-iri	spend-i-u
piac-eggu-o	'to like'	piac-iri	piac-i-u
pens-eggu-o	'to think'	pens-ari	pens-a-u
arriv-eggu-o	'to arrive'	rriv-ari	rriv-a-u
pass-eggu-o	'to pass'	pass-ari	pass-a-u
bbamp-eggu-o	'to go red'	bbamp-ari	bbamp-a-u

Interestingly, contrary to Venetian verbs, which considerably vary in the third person singular of the past tense, depending on the verb, the corresponding forms in Calabrian display a certain uniformity: as shown in (13), they consist of a stem, a vowel /a/ or /i/ indicating the inflection class, and they all end in -u, the marker for the third person singular.[37] This form regularity and the absence of infinitival forms may suggest that the adopted verbal material in Grekanico was initially drawn from the past perfective paradigm, before being submitted to an analysis in the purpose of supplying only the stem.

-ar-, based on the infinitival Italo-Romance marker -ar(e). -ar- is exclusively used for the formation of verbs of foreign origin, not necessarily Italo-Romance (e.g., Standard Modern Greek makijaro < French maquiller, Standard Modern Greek filmaro < English to film, etc.). It has been shaped by what Gardani (2016) calls "allogenous exaptation", since a change in function has occurred, from an inflectional marker of the source language to a derivational suffix in the target. See Ralli (2012a; 2016) for details.

36   Rohlfs (1977: §699) "[…] nelle parti più meridionali d'Italia, per influsso greco, l'uso dell'infinito è sconosciuto […]" ['due to Greek influence, in the southernmost parts of Italy, the use of infinitive is unknown', translation by A.R.].

37   However, the endings may change depending on the Calabrian variety (Squillaci, p.c.).

4   The Cypriot Case

As shown in Section 3, extra-linguistic factors interact with the linguistic ones for the adoption and integration of verbs from Romance and Turkish in four Modern Greek dialects, two of them being influenced by Italo-Romance and the other two by Turkish. The investigated data confirm that the overall typology of the source language, semi-fusional or agglutinative, does not exert any specific influence on the accommodation of loans in the target language, and that the same type of loans can be found in varieties which are affected either by Italo-Romance or by Turkish. Nevertheless, the morphological properties of the two systems in contact, and a certain pre-existing similarity in the sub-domain of verbal forms were suggested to play a substantial role on how these forms are integrated in the target system.

I now examine a dialect which has undergone the influence of both Romance and Turkish, that is, Cypriot. The point at issue is to see not only whether Cypriot has borrowed verbs from these languages but how it has accommodated them and whether verb borrowing in Cypriot shows similarities with verb borrowing in the four dialects examined so far.

Cypriot is the Greek-based dialect of circa 700,000 Greek-Cypriot people in Cyprus and of Cypriots living abroad, many of them in Great Britain. It is also the dialect of many aged Turkish Cypriot people who are nowadays confined in the northern part of Cyprus. It is basically a spoken dialect, but it can be found in many literary and non-literary texts.

In its long history, Cyprus has been ruled by different people and has been subject to different civilizations. As a result, Cypriot has entered in contact with languages that left their marks on it, mostly on its lexicon. As early as in 632 CE, Byzantine Cyprus was invaded by the Arabs and was reconquered by the Byzantines only in 964. From 1191 to 1489, the island was governed by the French dynasty of Lusignan, who spoke a form of Old Provençal (Chatziioannou, 1936). In 1489, the Lusignans transferred Cyprus to the Republic of Venice, whose rule ended in 1571, when the island was captured by the Ottoman Turks. In 1878, Great Britain took over the administration, and finally, in 1960, Cyprus became independent. After the Turkish invasion in 1974, the actual Cypriot state has been limited to the southern part of the island, where Cypriot is currently used in everyday communication.

The first dialectal sources appear around the 14th century. They consist of a translation of the French legal text of Assizes, which was followed by two dialectal texts in the fifteenth century, the Chronicles of Leontios Machairas and Georgios Voustronios.[38] Today, there are many lexical loans in the dialect,

---

38   See Beaudoin (1884) and Davy and Panayotou (2000) for details on these texts.

among which French, mainly of Old Provençal origin, Italo-Romance, Turkish, Arab and English (Papapavlou, 1994). Therefore, Cypriot constitutes an excellent case for testing hypotheses about language contact involving one target and several source languages of different origin and typology. In what follows, I will focus on the Romance and Turkish verbal loans, the topic of this article.

(14)   Cypriot (from Chatziioannou, 1936)   Old Provençal (from Anglade, 1921)

avanziazo	'to move forward'	*avançar*
anunsiazo	'to announce'	*anounçar*
ateniazo	'to stick'	*atenir*
finiazo	'to end'	*finir*
kufertiazo	'to confort'	*counfourtar*
kunsentiazo	'to consent'	*counsentir*
manteniazo	'to maintain'	*mantenir*
prezentiazo	'to present'	*presentar*
sufriazo	'to suffer'	*soufrir*
spiazo	'to spy'	*espiar*

(15)   Cypriot (from Chatziioannou, 1936)   Venetian (from Boerio 1856)

vantzaro	'to advance'	*vanzar*
vortaro	'to call on'	*voltar*
kreparo	'to crack'	*crepar*
payaro	'to pay'	*pagar*
rifuðaro	'to refuse'	*refudar*
salvaro	'to save'	*salvar*
saltaro	'to jump'	*saltar*
stimaro	'to estimate'	*stimar*
tratteniro	'to hold back'	*trategnir*
fermaro	'to stop'	*fermar*

(16)   Cypriot (from Chatzipieris and Kapatas, 2015)   Turkish infinitive (3sg)

kaurtízo	'to fry/sizzle'	*kavurmak* (*kavurdu*)
kunuštízo	'to kid around'	*konušmak* (*konuštu*)
kazandízo	'to win (a game)'	*kazanmak* (*kazandı*)
paylatízo/ *paylató*	'to tie'	*bağlamak* (*bağladı*)
peendízo/ *peendó*	'to respect'	*beğenmek* (*beğendi*)
taništízo	'to take advice'	*današmak* (*današtı*)
sajdízo/sajdó	'to appreciate'	*saymak* (*saydı*)
čattízo/čattó	'to match'	*çatmak* (*çatı*)

dajandízo/ *dajandó*	'to endure'	*dayanmak* (*dayandı*)
šastízo	'to get confused'	*šašmak* (*šaštı*)

The adoption of Romance and Turkish verbal loans in Cypriot is particularly intriguing, since it combines all strategies and patterns seen so far for Grekanico, Heptanesian, Pontic and Aivaliot. First, similarly to Grekanico (1) and Pontic (3), Cypriot has retained only the stem from the Old Provençal verbs and has accommodated it via the indirect strategy, that is with the help of an integrating element, the suffix *-iaz-*. The only difference between Cypriot and these two dialects is that whilst Grekanico and Pontic employ the *-ev-* suffix, Cypriot uses *-iaz-*. According to Chatzidakis (1905: 305), verbal derivation in *-iaz-* was particularly productive in Medieval Cyprus, contrary to the other parts of the Greek speaking world, where *-iz-* (see Aivaliot in (4)) or *-ev-* (see Grekanico and Pontic in (1) and (3)) were most frequent. I have already proposed that high productivity constitutes a decisive factor for the selection of an integrating element. Thus, it is not surprising that, during the Lusignan rule (from 1191 to 1489), Cypriot has accommodated Old Provençal verbs by using the most productive *-iaz-*. A crucial question that arises though is whether, before getting analyzed, the Old Provençal verbs had entered Cypriot as infinitival or as past tense forms. Along the lines of what I have suggested for Heptanesian, I am tempted to propose that Cypriots had adopted the infinitival forms, as being the most unmarked and regular ones, because the third person singular of the past perfective (simple past) paradigm was not uniform for all verbs. Following Anglade (1921), Old Provençal verbs in *-ar* (e.g., *avançar* 'to move forward') form their third person singular of the past perfective tense in *-et* (pronounced as [e]), e.g., *avancet*, whereas a suffix *-gu-* appears in the paradigms of verbs in *-ir*, *-er* and *-re* (except for those in *-dre*, e.g., *vendre* 'to sell'). See, for instance, the form *finiguet* of the verb *finir* 'to end'.

Second, like Heptanesians (2) and Aivaliots (4), Cypriots borrow from Venetian the entire infinitival forms and from Turkish the third person singular of the past tense. In the first case, only the direct insertion is used, similarly to Heptanesian, while in the second case, both strategies alternate for most verbs, the indirect one -with *-iz-* as integrating element- and the direct strategy. As seen in Section 3.4, this alternation also occurs in the accommodation of Turkish loans in Aivaliot. Again, I would like to propose that this peculiar situation is due to the interplay of linguistic and sociolinguistic factors. According to Dendias (1923: 157), during the French regime in Cyprus, the degree of

education was relatively high, and French and Greek were taught at schools. As a consequence, when borrowing occurred, loan verbs were heavily hellenicized by the Cypriot speakers, who could analyze them, retain the stems and combine them with the Greek integrating suffix *-iaz-* before adding the Greek inflectional ending. In contrast, during the Venetian and the Ottoman periods, schools were closed, education was poor and the command of the politically dominant language was definitely low. I, thus, assume that educational deficiency and low command of the source language led the speakers borrow the entire Venetian infinitives or the Turkish past forms, reanalyze them as stems and add the inflectional endings without proceeding to a word internal analysis. That the socio-linguistic context is critical for the adoption of a specific strategy in loan verb accommodation gets further support by the fact that, even during the Lusignan rule and because of a flourishing trade with Venice, Venetian verbs had entered the Cypriot vocabulary as verbs in *-iaz-o* and not as verbs in *-ar-o*. The example *siɣur-iaz-o* 'to make sure', adopted from the Venetian *sicurar* according to Chatzidakis (1905: 304), adds substantial proof to this suggestion. However, the Old Provençal counterpart *asegurar* makes its origin disputable.[39]

To partially sum up, the Cypriot case shows that borrowing in a particular linguistic system can be constrained by the interplay of linguistic and extra-linguistic factors. That is, high linguistic skills or educational deficiency can lead to heavy or weak integration of the borrowed material, the form of which is determined by the morphological properties of the target language, while a certain matching of forms between the two languages in contact may also be at play. The situation where linguistic factors interact with extra-linguistic ones for molding the accommodation of loan items is not unknown in other contact settings. See, for instance, Clements and Luís (2015) on how borrowing is affected in Korlai Indo-Portuguese.

5   Conclusions

This article compares Romance and Turkish verbal loans in five Modern Greek dialects, Grekanico, Heptanesian, Pontic, Aivaliot and Cypriot, revealing that there is no consistency in the way loans are adopted and accommodated. The absence of uniformity is observed not only from dialect to dialect, and from source language to source language, but even within the same target system. It is shown that the various integration strategies and patterns are not

---

39   On the difficulty of determining the origin of borrowed words, see Minervini (2019).

distributed according to the particular source or target languages, since the same strategy or pattern can be found in dialects which are in contact with different languages and different strategies or patterns can alternate within the same dialect influenced by the same source language. While borrowing of verbs occurs in a heavy-contact situation, the native morphological properties, principally those of the target language, as well as a certain compatibility between the transferred forms of the source language and the corresponding native forms of the target have been proposed to be the determining linguistic factors for verbal loan integration, together with the extra-linguistic factors referring to the intensity of contact, a profound knowledge of the source language, the degree of education and the speakers' sociolinguistic attitude towards the dominant language. As far as the native morphological characteristics are concerned, those which seem to play a predominant role in the integration of verbal loans is the stem-based property of Greek to build its derived and inflected formations on stems, and the prevalence of the past perfective stem as a base for the creation of derivative structures. It is also demonstrated that a more elaborated accommodation of the transferred verbs results in the adoption of an integrating element, the choice of which is controlled by productivity considerations. Agreeing with Wohlgemuth (2009), this work proves that verbs can be borrowed as such, provided that certain conditions are met, contra Moravcsik (1978: 111), who has suggested that a lexical item whose meaning is verbal cannot be included in the set of borrowed elements without being nominalized.

Finally, this study shows that verb borrowing from Turkish does not seem to belong to the Balkan Sprachbund features, at least as far as Greek is concerned. First, there are Greek varieties which are outside the geographic boundaries of the Balkans, that is, Pontic, Grekanico and Cypriot (even Cappadocian, mentioned in footnotes 18 and 25) but share the same type of loan integration as other dialects which belong to the Balkan area, that is, Heptanesian and Aivaliot. Second, languages which are traditionally considered to belong to the Balkan Sprachbund, as for instance, Bulgarian, do not integrate their Turkish verbal loans directly from Turkish, as is the Greek case, but through the mediation of Greek: Bulgarian adopts the past perfective stems of Turkish loans as they appear in Greek, that is, with the Greek perfective marker -s- (see footnote [31]), adds the Bulgarian verbalizer -va- and the Bulgarian personal ending (e.g., Bulgarian *bastisvam* 'I attack suddenly/stomp/print' < Turkish past perfective *bastı* + Greek perfective marker -s- + Bulgarian verbalizer -va- + Bulgarian 1sg -m). Thus, along the lines of Gardani, Loporcaro, and Giudici (2021) I believe that change in the Balkan languages does not necessarily converge but is due to a pairwise contact over a long period of time.

Acknowledgments

References

Anglade, Joseph. 1921. *Grammaire de l'ancien provençal ou ancienne langue d'oc*. Paris: Klincksieck.

Bağrıaçık, Metin, Angela Ralli, and Dimitra Melissaropoulou. 2015. Borrowing verbs from Oghuz Turkic: two linguistic areas. In Francesco Gardani, Peter Arkadiev and Nino Amiridze (eds.), *Borrowed Morphology*, 109–136. Berlin: De Gruyter Mouton.

Beaudoin, Mondry. 1884. *Étude du dialecte chypriote moderne et médiéval*. Paris: Ernest Thorin.

Berikashvili, Svetlana. 2017. *Morphological Aspects of Pontic Greek Spoken in Georgia*. Munich: LINCOM.

Boerio, Giuseppe. 1856. *Dizionario del dialetto veneziano*. 2nd edn. Venice: Premiata tipografia di Giovanni Cecchini.

Breu, Walter. 1991a. Abweichungen vom phonetischen Prinzip bei der Integration von Lehnwörtern. In Klaus Hartenstein and Helmut Jachnow (eds.), *Slavistische Linguistik 1990*: *Referate des XVI. Konstanzer Slavistischen Arbeitstreffens Bochum/ Löllinghausen 19.-21.9.1990* (Slavistische Beiträge 274), 36–69. Munich: Otto Sagner.

Breu, Walter. 1991b. System und Analogie bei der Integration von Lehnwörtern: Versuch einer Klassifikation. *Incontri Linguistici* 14(1): 13–28.

Browning, Robert. 2004. *Elliniki Glossa, Meseoniki ke Nea* [*Greek Language, Medieval and Modern*]. Athens: Papadimas. [Originally published as *Medieval and Modern Greek*. London: Hutchinson University Library 1969].

Bryer, Anthony. 1975. Greeks and Türkmens: The Pontic exception. *Dumbarton Oaks Papers* 29: 113–148.

Caracausi, Girolamo. 1979. *Testi neogreci di Calabria. Indice lessicale*. Palermo: Istituto Siciliano di Studi Bizantini e Neoellenici.

Caratzas, Stamatis. 1958. L'origine des dialectes néo-grecs de l'Italie méridionale. In *Collection de l'Institut d'Études Byzantines et Néo-helléniques* 18. Paris: Les Belles Lettres.

Chairetakis, George. 2020. *Η μορφολογία της κρητικής διαλέκτου: κλίση και παραγωγή* [*The morphology of the Cretan dialect: inflection and derivation*]. PhD dissertation, University of Patras.

Chantraine, Pierre. 1945. *Morphologie historique du grec*. Paris: Klincksieck.

Chatzidakis, Georgios. 1905. *Meseonika ke Nea Ellinika* [*Medieval and Modern Greek*], vol. 1. Athens: Sakellarios.

Chatziioannou, Kyriakos. 1936. *Peri ton en ti Meseoniki ke Neotera Kypriaki Ksenon Glossikon Stichion* [*About the Foreign Linguistic Elements in Medieval and Modern Cypriot*]. Athens: Xenou.

Chatzipieris, Iakovos and Orhan Kapatas. 2015. *Koino Leksiko tis Ellinokyrpiakis ke Tourkokypriakis Dialektou* [Common Dictionary of the Greek-Cypriot and Turkish-Cypriot Dialect]. Nicosia: Lithographica Ltd.

Clements, Clancy, J. and Ana R. Luís. 2015. Contact intensity and the borrowing of bound morphology in Korlai Indo-Portuguese. In Francesco Gardani, Peter Arkadiev and Nino Amiridze (eds.), *Borrowed Morphology*, 219–240. Berlin: De Gruyter Mouton.

Contosopoulos, Nicolaos 2001. *Dialects and Idioms of Modern Greek*. Athens: Grigoris.

Cortelazzo, Manlio. 2017. *Dizionario veneziano della lingua e della cultura popolare nel XVI secolo*. Limena, Padova: La Linea.

Davy, John and Anna Panayotou. 2000. French loans in Cypriot Greek. In Jean Pouilloux (ed.), *Travaux de la Maison de l'Orient Méditerranéen*, 113–125. Lyon: Maison de l'Orient et de la Méditerranée.

Dawkins, Richard. 1916. *Modern Greek in Asia Minor.* Cambridge: Cambridge University Press.

Deffner, Michael. 1881. *Zakonische Grammatik*. Berlin: Weidmannsche Buchhandlung.

Dendias, Michael. 1923. *Peri ton en ti Kypriaki rimaton ek tis Italikis ke tis Gallikis* [*About the Italian and French verbs in Cypriot*]. *Glossika Simeiomata*, 142–165.

Drettas, Georges. 1997. *Aspects pontiques*. Paris: Arp.

Enrique-Arias, Andrés. 2010. On language contact as an inhibitor of language change: The Spanish of Catalan bilinguals in Majorca. In Anne Breitbarth, Christopher Lucas, Sheila Wattsy, and David Willis (eds.), *Continuity and Change in Grammar*, 97–118. Amsterdam: Benjamins.

Fanciullo, Franco. 2001. On the origins of Greek in Southern Italy. In Angela Ralli, Brian Joseph and Mark Janse (eds.), *Proceedings of the 1st International Conference of Modern Greek Dialects and Linguistic Theory*, 67–78. Patras: University of Patras.

Fanciullo, Franco. 2008. Gli italianismi del neogreco. *L'Italia Dialettale* 69: 1–41.

Field, Fredric. 2002. *Linguistic Borrowing in Bilingual Contexts.* Amsterdam: Benjamins.

Gambino, Francesca (ed.). 2007. *I vangeli in antico veneziano*, manoscritto Marciano it. I, 3 (4889). Rome: Antenore.

Gardani, Francesco. 2016. Allogenous exaptation. In Muriel Norde and Freek van de Velde (eds.), *Exaptation and Language Change*, 227–261. Amsterdam: Benjamins.

Gardani, Francesco. 2020a. Borrowing matter and pattern in morphology. An overview. *Morphology* 30(4): 263–282.

Gardani, Francesco. 2020b. Morphology and contact-induced language change. In Anthony Grant (ed.), *The Oxford Handbook of Language Contact*, 96–122. Oxford: Oxford University Press.

Gardani, Francesco, Peter Arkadiev, and Nino Amiridze. 2015. Borrowed morphology: An overview. In Francesco Gardani, Peter Arkadiev and Nino Amiridze (eds.), *Borrowed Morphology*, 1–26. Berlin: De Gruyter Mouton.

Gardani, Francesco, Michele Loporcaro, and Alberto Giudici. 2021. In and around the Balkans: Romance languages and the making of multi-layered languages. *Journal of Language Contact* 14(1): 1-23.

Haspelmath, Martin. 2009. Lexical borrowing: Concepts and issues. In Martin Haspelmath and Uri Tadmor (eds.), *Loanwords in the World's Languages: A Comparative Handbook*, 35–54. Berlin: Mouton De Gruyter.

Henrich, Günther Steffen. 1990. The Pontic elements in a Greek-language islamic text of the 1500s. In *Festschrift in memoriam of Stamatis Caratzas*, 355–374. Thessaloniki: Aristotle University of Thessaloniki.

Henrich, Günther Steffen. 2011. First appearance of grammatical innovations of the Pontic and Crimean Pontic dialect. *Neoelliniki Dialektologia* 6: 89–109.

Hitiris, Georgios. 1987. *Kerkyraiko Glossario* [*Corfiot Glossary*]. Corfu.

Hock, Hans Henrich and Brian Joseph. 1996. *Language History, Language Change and Language Relationship*. Berlin: Mouton De Gruyter.

Horrocks, Geoffrey. 1997. *Greek: A History of the Language and its Speakers*. London: Longman.

Jakobson, Roman. 1938. Sur la théorie des affinités phonologiques entre les langues. In *Actes du 4e Congrès International des Linguistes*, 48–59. Copenhagen: Einar Munskgaard.

Janse, Mark. 2004. Animacy, definiteness, and case in Cappadocian and other Asia Minor dialects. *Journal of Greek Linguistics* 5: 3–26.

Joseph, Brian. 1998. Diachronic morphology. In Arnold Zwicky and Andrew Spencer (eds.), *The Handbook of Morphology*, 351–373. Oxford: Blackwell.

Kahane, Henry and Renée Kahane. 1982. The Western impact on Byzantium: The linguistic evidence. *Dumbarton Oaks Papers* 36: 127–153.

Karanastasis, Anastasios. 1984. *Istorikon Leksikon ton Ellinikon Idiomaton tis Kato Italias* [*Historical Dictionary of the Greek Dialects of South Italy*]. Athens: Academy of Athens.

Karanastasis, Anastasios. 1997. *Grammatiki ton Ellinikon Idiomaton tis Kato Italias* [*Grammar of the Greek Dialects of South Italy*]. Athens: Academy of Athens.

Katsoyannou, Marianna. 1995. *Le parler greco de Galliciano (Italia): description d'une langue en voie de disparition*. PhD dissertation, Université de Paris VII.

Katsoyannou, Marianna. 1999. To idioma tis Calavrias [The dialect of Calabria]. In Tasos Christidis (ed.), *Dialectal Enclaves of the Greek Language*, 39–45. Thessaloniki: Center of the Greek Language.

Katsouda, Georgia. 2016. *To Idioma ton Kythiron: Perigrafi ke Analysi* [*The Dialect of Cythera: Description and Analysis*]. Athens: Cythera Association.

Kitromilides, Paschalis and Alexis Alexandris. 1984. Ethnic survival, nationalism and forced migration. The historical demography of the Greek community of Asia Minor at the close of the Ottoman era. *Deltio Kentrou Mikrasiatikon Spoudon* 5: 9–44.

Konomos, Dinos. 2003. *Leksiko tis Zakynthou* [Dictionary of Zante]. 2nd edn. Zante: Trimorfo.

Korosidou-Karra, Ermioni. 2003. *Ta Romanika* (*Italika – Gallica*) *Dania sto Sinxrono Idioma tis Zakynthou* [The Romance Borrowings in the Zante dialect]. PhD dissertation, Aristotle University of Thessaloniki.

Krimpas, Panagiotis. 2018. Italian. In George Kanarakis (ed.), *The Legacy of the Greek Language*, 37–73. New York: Peridot.

Ledgeway, Adam. 1998. Variation in the Romance infinitive: The case of Southern Calabrian inflected infinitive. *Transactions of the Philological Society* 96: 1–61.

Ledgeway, Adam, Norma Schifano and Giuseppina Silvestri. 2021. The negative imperative in southern Calabria. *Spirito greco, materia romanza* again? *Journal of Language Contact* 14(1): 184–219.

Loporcaro, Michele. 2013. *Profilo linguistico dei dialetti italiani*. 2nd edn. Rome-Bari: Laterza.

Mackridge, Peter. 1987. Greek speaking Moslems of North-East Turkey: Prolegomena to a study of the Ophitic sub-dialect of Pontic. *Byzantine and Modern Greek Studies* 11: 115–137.

Makri, Vasiliki. 2020. *The Formal Expression of Grammatical Gender in a Modern Greek Dialect Affected by Italo-Romance*. PhD dissertation, University of Patras.

Manolessou, Io. 2005. The Greek dialects of Southern Italy: An overview. *KAMPOS* 13: 103–125.

Manolessou, Io. 2019. The historical background of the Asia Minor dialects. In Angela Ralli (ed.), *The Morphology of Asia Minor Dialects*, 20–65. Leiden: Brill.

Manolessou, Io and Nikolaos Pantelidis. 2011. To archaeo elliniko /ɛː/ stin Pontiaki [Ancient Greek /ɛː/ in Pontic]. *Neoelliniki Dialektologia* 6: 245–271.

Manolessou, Io and Angela Ralli. 2020. On borrowing and integrating Italo-Romance nouns in South Italian Greek. In Patrizia del Puente, Francesca Guazzelli, Lucia Molinu, and Simone Pisano (eds.), *Festschrift to Franco Fanciullo*, 417–434. Alessandria: Edizioni dell'Orso.

Martino, Paolo. 1979. L'isola grecanica dell'Aspromonte; aspetti sociolinguistici. In Federico Albano Leoni (ed.), *I dialetti e le lingue delle minoranze di fronte all'italiano: Atti del XI Congresso Internazionale di Studi, Cagliari, 27-30 Maggio 1977*, 305–341. Rome: Bulzoni.

Matras, Yaron. 2009. *Language Contact*. Cambridge: Cambridge University Press.

Meillet, Antoine. 1921. *Linguistique historique et linguistique générale*. Paris: Champion.

Minervini, Laura. 2019. Veneziano e francese nell'Oriente latino. In Daniele Baglioni (ed.), *Il veneziano «de là da mar». Contesti, testi, dinamiche del contatto linguistico e culturale*, 177–200. Berlin: De Gruyter.

Moravcsik, Edith A. 1978. Language contact. In Joseph H. Greenberg, Charles A. Ferguson, and Edith A. Moravcsik (eds.), *Universals of Human Language: Method & Theory*, 93–122. Stanford: Stanford University Press.

Morosi, Giuseppe. 1870. *Studio sui dialetti greci della Terra di Otranto*. Lecce: Tipografia Editrice Salentina.

Myers-Scotton, Carol. 2002. *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes.* Oxford: Oxford University Press.

Newton, Brian. 1972. *Cypriot Greek: Its Phonology and Inflections*. The Hague: Mouton.

Oikonomidis, Dimosthenis. 1958. *Grammatiki tis Ellinikis Dialektou tou Pontou* [*Grammar of Pontic Greek*]. Athens: Academy of Athens.

Özkan, Hakan. 2013. The Pontic Greek spoken by Muslims in the villages of Beşköy in the province of present-day Trabzon. *Byzantine and Modern Greek Studies* 37(1): 130–150.

Papadopoulos, Anthimos. 1955. *Istoriki Grammatiki tis Pontikis Dialektou* [*Historical Grammar of the Pontic Dialect*]. Athens: Epitropi Pontiakon Meleton.

Papadopoulos, Anthimos. 1958-1961. *Istorikon Leksikon tis Pontikis Dialektou* [*Historical Dictionary of the Pontic Dialect*]. Athens: Epitropi Pontiakon Meleton.

Papapavlou, Andreas. 1994. *Language Contact and Lexical Borrowing in the Greek Cypriot Dialect*. Athens: Grivas.

Parlangeli, Oronzo. 1953. Sui dialetti romanzi e romaici del Salento. *Memorie dell'Instituto Lombardo di Scienze e Lettere* (*Classe di Lettere*) 25: 93–198.

Profili, Olga. 1983. *Le parler grico de Corigliano d'Otranto* (*provence Lecce, Italie*)*: Phénomènes d'interférence entre le parler grec et les parlers romans environnants, ainsi qu'avec l'italien*. PhD dissertation, Université Stendhal, Grenoble.

Ralli, Angela. 2005. *Morfologia* [*Morphology*]. Athens: Patakis. Updated and enriched version published in 2021.

Ralli, Angela. 2009. Morphology meets dialectology: Insights from Modern Greek dialects. *Morphology* 19(1): 87–105.

Ralli, Angela. 2012a. Verbal loanblends in Griko and Heptanesian: A case study of contact morphology. *L'Italia Dialettale* LXXIII: 111–132.

Ralli, Angela. 2012b. Morphology in language contact: Verbal loanblend formation in Asia Minor Greek (Aivaliot). In Martine Vanhove, Thomas Stolz, Aina Urdze, and Hitomi Otsuka (eds.), *Morphologies in Contact*, 185–201. Berlin: Akademie Verlag.

Ralli, Angela. 2013. Greek. *Revue belge de philologie et d'histoire* 90: 939–966.

Ralli, Angela. 2015. Greek. In Peter Müller, Ingeborg Ohnheiser, Susan Olsen, and Franz Rainer (eds.), *Word Formation. An International Handbook of the Languages of Europe*, 3138–3156. Berlin: Mouton De Gruyter.

Ralli, Angela. 2016. Strategies and patterns of loan verb integration in Modern Greek varieties. In Angela Ralli (ed.), *Contact Morphology in Modern Greek Dialects*, 73–108. Newcastle upon Tyne: Cambridge Scholars Publishing.

Ralli, Angela. 2017. *Dictionary of the Dialectal Varieties of Moschonisia, Aivali and Eastern Lesbos*. Athens: Foundation of Historical Studies.

Ralli, Angela. 2019a. Greek in contact with Romance. In Michele Loporcaro and Francesco Gardani (eds.), *The Oxford Encyclopedia of Romance Linguistics*. New York: Oxford University Press. DOI: 10.1093/acrefore/9780199384655.013.422.

Ralli, Angela. 2019b. Affixoids and verb borrowing in Aivaliot morphology. In Angela Ralli (ed.), *The Morphology of Asia Minor Greek. Selected Topics*, 221–254. Leiden: Brill.

Ralli, Angela, Marianna Gkiouleka, and Vasiliki Makri. 2015. Gender and inflection class in loan-noun integration. *SKASE Journal of Theoretical Linguistics* 12(3): 422–460.

Ralli, Angela and Vasiliki Makri. 2020. Examining the integration of borrowed nouns in immigrant speech: The case of Canadian-Greek. In Pius ten Hacken and Renáta Panocová (eds.), *The Interaction of Borrowing and Word Formation*, 237-258. Edinburgh: Edinburgh University Press.

Rohlfs, Gerhard. 1933. *Scavi Linguistici nella Magna Grecia*. Rome: Collezione Meridionale.

Rohlfs, Gerhard. 1977. *Grammatica storica dei dialetti italogreci* (*Calabria, Salento*). Munich: Beck.

Rohlfs, Gerhard. 1997. *Studi e ricerche su lingue e dialetti d'Italia*. 2nd edn. Milan: Sansoni. [1st edn. Florence 1972].

Sakel, Jeanette. 2007. Types of loan: Matter and pattern. In Yaron Matras and Jeanette Sakel (eds.), *Grammatical Borrowing in Cross-Linguistic Perspective*, 15–29. Berlin: Mouton de Gruyter.

Sakkaris, Georgios. 1920. *History of Kydonies*. Kydonies. [Reprinted in 1940 in Athens, by the Academy of Athens].

Sakkaris, Georgios. 1940. On the Kydonies dialect in comparison with Lesbian. *Mikrasiatika Chronika* 3: 74–141.

Sakkaris, Georgios. 1948. On the Kydonies dialect. Vocabulary. *Mikrasiatika Chronika* 4: 49–81.

Salvanos, Georgios. 1918. *Meleti peri tou Glossikou Idiomatos ton en Kerkyra Argyradon* [*Study of the Dialect of Corfiot Argyrades*]. Athens.

Seifart, Frank, Jan Strunk, Swintha Danielsen, Iren Hartmann, Brigitte Pakendorf, Søren Wichmann, Alena Witzlack-Makarevich, Nivja H. de Jong, and Balthasar Bickel.

2018. Nouns slow down speech across structurally and culturally diverse languages. *Proceedings of the National Academy of Sciences* 115(22): 5720–5725.

Simiris, Dimitris. 2017. *To Glossiko Idioma tis Ithakis* [*The Dialect of Ithaca*]. Patras: Laboratory of Modern Greek Dialects and Community of Ithaca.

Sitaridou, Ioanna. 2013. Greek-speaking enclaves in Pontus today: The documentation and revitalization of Romeyka. In Mari C. Jones and Sarah Ogilvie (eds.), *Keeping Languages Alive: Language Endangerment, Documentation, Pedagogy and Revitalization*, 98–112. Cambridge: Cambridge University Press.

Sitaridou, Ioanna. 2014. The Romeyka infinitive: Continuity, contact and change in the Hellenic varieties of Pontus. *Diachronica* 31(1): 23–73.

Skubic, Mitja. 1986. Passato prossimo e passato remoto nei dialetti veneti. In Manlio Cortelazzo (ed.), *Guida ai dialetti veneti V*, 31–43. Padua: Cleup.

Soldatos, Konstantinos. 1967. I elliniki glossa sta Eptanisa [The Greek language in the Heptanesian islands]. *Kerkyraika Chronika* 13-14: 84–104.

Squartini, Mario and Pier Marco Bertinetto. 2000. The simple and compound past in Romance languages. In Östen Dahl (ed.), *Tense and Aspect in the Languages of Europe*, 403–439. Berlin: Mouton De Gruyter.

Squillaci, Maria Olimpia. 2016. *When Greek Meets Romance: A Morphosyntactic Investigation of Language Contact in Aspromonte*. PhD dissertation, Cambridge University.

Stussi, Alfredo (ed.). 1965. *Testi veneziani del Duecento e dei primi del Trecento*. Pisa: Nistri-Lischi.

Telmon, Tullio. 1992. *Le minoranze linguistiche in Italia*. Turin: Edizioni dell'Orso.

Thomason, Sarah. 2001. *Language Contact. An Introduction*. Edinburgh: Edinburgh University Press.

Trudgill, Peter. 2003. Modern Greek dialects: A preliminary classification. *Journal of Greek Linguistics* 4: 45–63.

Vryonis, Speros. 1971. *The Decline of Medieval Hellenism in Asia Minor and the Process of Islamization from the Eleventh through the Fifteenth Century*. Berkeley: University of California Press.

Wichmann, Søren and Jan Wohlgemuth. 2008. Loan verbs in a typological perspective. In Thomas Stolz, Dick Bakker and Rosa Salas Palomo (eds.), *Aspects of Language Contact. New Theoretical, Methodological and Empirical Findings with Special Focus on Romancisation Processes*, 89–121. Berlin: Mouton De Gruyter.

Whitney, William Dwight. 1881. On mixture in language. *Transactions of the American Philosophical Association* 12: 1–26.

Winford, Donald. 2003. *An Introduction to Contact Linguistics*. Oxford: Blackwell.

Wohlgemuth, Jan. 2009. *A Typology of Verbal Borrowings*. Berlin: Mouton De Gruyter.