No-Reversal Constraint and beyond: Word-internal language mixing in Anatolia

Faruk Akkuş

University of Massachusetts Amherst

(To appear at *Linguistic Variation*)

Abstract

This study investigates various un(der)studied word-internal language mixing patterns among Turkish, Anatolian Arabic and Northern Kurdish, in the context of both verbal and nominal domains. The examination of these patterns reveals various theoretical implications. First, head-directionality may change as a result of language contact. Second, in some instances, certain functional categories are borrowed as semantically vacuous heads, and are identical to their bare counterparts (cf. Marantz 2013; Anagnostopoulou and Samioti 2014). Therefore, such semantically empty heads are ignored for meaning. Moreover, informed by the rarely-discussed trilingual language-mixing contexts, the study demonstrates that various formal approaches to code-switching which rely on either a distinction between functional vs lexical categories or phasehood as the defining constraint on code-switching are not tenable (e.g., Poplack 1981; Belazi et al. 1994; López et al. 2017). This study demonstrates language mixing is more permissive for the languages in question than would be predicted by these approaches, and proposes the *No-Reversal Constraint*, whose governing restriction is that code-switching does not allow a switch back to a language that has already been externalized earlier in the derivation.

Keywords: word-internal language mixing, multilingualism, No-Reversal Constraint, code-switching, light verbs, Anatolia

1 Introduction

This paper discusses various processes of verbal and nominal borrowings in the language contact situations, with Turkish (TK), Anatolian Arabic (AA) and Northern Kurdish (NK) forming the empirical focus. The speakers of these languages have been co-existing for centuries, which has resulted in extensive contact among these languages. The influence of contact can be observed in many domains of the grammar, including syntax (e.g., Akkuş 2016; Akkuş and Benmamoun 2018; Haig 2007; Lahdo 2009; Öpengin 2012).

The types of verbal predicate formations of interest are exemplified in (1). The examples in (1a)-(1b) represent the *light-verb construction* (LVC), a domain in which the influence

of contact is clearly manifested. For instance, in (1a), the nonverbal element *qazan* 'win' is drawn from Turkish, whereas the light verb *sawa* 'do' is from Anatolian Arabic.[1] (1c) illustrates the *templatic pattern*, in which the Turkish loan-verb *kapat* 'close' is incorporated into the verbal pattern of the recipient Semitic language, Anatolian Arabic.

(1) a. qazan_{TK} sawa_{AA}
win do
'to win' (AA) cf. Turkish *kazan-* 'win'

b. $\textit{say}_{\text{TK}}$ $\textit{kirin}_{\text{NK}}$
count do
'to count' (NK) cf. Turkish *say* 'count'

c. *īqappət* 'to close' (AA) cf. Turkish *kapat-* 'close'

Word-internal mixing is another domain where language contact has resulted in (ongoing) language change. Word-internal switches in the nominal domain, which mainly involve the *affixal pattern*, are illustrated in (2). They primarily represent instances in which a root is chosen from language A, whereas an affix from language B.

(2) a. $\textit{xatan}_{\text{AA}}$-$\textit{lar}_{\text{TK}}$ 'son-in-law-PL'

b. $\textit{milyar}_{\text{TK}}$-$\bar{\textit{a}}\textit{t}_{\text{AA}}$ 'billion-PL'

c. $\textit{kıleb}_{\text{AA}}$-$\textit{ler}_{\text{TK}}$ 'dog.PL-PL'

Investigation of the LVC across Anatolian Arabic varieties shows that the head-directionality of a construction may change as a result of language contact.[2] The LVC in some AA dialects has adopted a mainly head-final directionality following the source languages, and differently from the Arabic head-initial pattern. As such, the *light verb +*

[1] I use the *subscripting* notation in the examples, when necessary, to indicate which language a word/lexeme or morpheme comes from, e.g., the subscript $_{\text{AA}}$ means the morpheme is from Anatolian Arabic. This is different from Sankoff and Poplack's (1981) *superscripting* ("bilingual tagging") mechanism that is used to restrict lexical insertion rules so that the grammar contributing the phrase structure rule would also be the grammar from which lexical insertion rules would be drawn.

[2] Through the examination of these case studies, this work addresses the implications of word-internal borrowing and verbal formation for generative theories such as the Minimalist Program, coupled with Distributed Morphology, and thus is also intended to further the work in the non-lexicalist theories regarding word-internal mixing. See also Tracy 2000; González-Vilbazo and López 2011; Grimstad and Åfarli 2014; Bandi-Rao and den Dikken 2014; Alexiadou et al. 2015b; Alexiadou 2017; Gündoğdu and Akkuş 2022, a.o. In so doing, it capitalizes on our current understanding of word-formation, verb-building and argument structure including complex predicates to explain the observed patterns. On the flip side, patterns of word-formation in language mixing contexts are used to inform our theories of the building blocks of verbal meaning as well as language change.

nonverbal element order is now replaced by *nonverbal element + light verb*. The study also examines the LVC borrowings in NK and AA that involve instances with the perfective suffix *-mIş* from Turkish, as shown in (3).[3] It turns out that while *-mIş* is semantically vacuous in NK and many dialects of AA, it still bears semantic function for some speakers of AA, including Sason Arabic. I take this to mean that certain functional categories are borrowed as semantically vacuous heads, and are identical to their bare counterparts (cf. Marantz 2013; Anagnostopoulou and Samioti 2014). Accordingly, semantically empty heads are ignored for meaning.

(3) dinle-miş$_{TK}$ sawa$_{AA}$
listen-PFV do
'to listen' (AA) cf. Turkish *dinle-miş* 'to have listened'

Furthermore, the investigation of language-mixing in the verbal and nominal domains reveals that prior approaches fail to capture the full range of attested patterns. The well-known study by Poplack (1980, 1981) argues that code-switching could only target two free morphemes, and it could not apply between morphemes within the same word. This approach effectively imposes a ban on code-switching in head movement and word-internal contexts (Poplack 1981). Later studies have shown that this statement is too strong, and such patterns indeed exist. Belazi et al. (1994) instead proposes the Functional Head Constraint (FHC), which distinguishes between lexical and functional categories. According to the FHC, a code-switch may not occur between a functional head and its complement, while code-switching between a lexical head and its complement proceeds unimpeded.

An interesting recent analysis by López et al. (2017) argues that code-switching is governed by *phase*-theoretic considerations. Building on Alexiadou et al. 2015b and Alexiadou 2017, López et al. (2017) make the natural assumption that bilinguals have multiple externalization systems or PFs, and suggest that when a structure is transferred, it is transferred in one block to one of the PFs. From this, they argue that "code-switching may take place at phase boundaries but not within the phase" (López et al. 2017:5). This is because code-switching within the phase would entail transferring some material to one externalization system while simultaneously transferring some other material to another externalization system. As will be elaborated in §4, this paper argues that even these two approaches by Belazi et al. (1994) and López et al. (2017) face challenges with a number of code-switching patterns, including those such as (2c). Instead, I propose a constraint that

[3]Following the Turkological convention, I indicate segments that undergo assimilation processes in capital letters. Such vowels undergo vowel harmony for backness and rounding, while consonants undergo assimilation in voicing.

prohibits switching back to a language that has already contributed a morpheme earlier in the derivation (in a domain), which I call *No-Reversal Constraint*. The merits of this constraint become clear especially in the context of rarely-studied trilingual code-switching. We observe that in a contact situation involving three languages, word-internal code-switching allows various patterns of morpheme insertion from the contact languages, (4a) through (4c), but not the pattern in (4d). For example, (4a) indicates that a Morph(eme) from Lang(uage) A can be followed by a morpheme from Language B, which in turn is followed by another morpheme from Language B. (4c) shows that each morpheme can belong to a different language. In the ungrammatical version in (4d), the first morpheme is from Language A, followed by the second morpheme from Language B, and crucially, the third morpheme is drawn back from Language A.

(4) **Morph1** **Morph2** **Morph3**

a. Lang A $\prec$ Lang B $\prec$ Lang B

b. Lang A $\prec$ Lang A $\prec$ Lang B

c. Lang A $\prec$ Lang B $\prec$ Lang C

d. *Lang A $\prec$ Lang B $\prec$ Lang A

I argue that the patterns (4) are capturable only via the *No-Reversal Constraint*. I also illustrate instances in which more morphemes are available in a single morphological word, so long as the patterns in (4) are obeyed (in a way not predicted by the FHC or the phase-based accounts).

The paper is organized as follows: §2 introduces the sociolinguistic situation of the contact languages under investigation, as language contact has both linguistic and extra-linguistic components. §3 discusses the verbal predicate formation processes, i.e., the light verb constructions and the templatic patterns. It highlights the variation across languages and varieties as to the status of the perfective suffix *-mIş*, which is shown to be semantically empty in some varieties. §4 analyzes the affixal and templatic patterns in both the verbal and nominal domains, allowing for a comparison between various models of language-mixing. It argues for the *No-Reversal Constraint* over the alternative hypotheses. §5 summarizes and concludes the paper.

2 The Sociolinguistic and Theoretical Background

This section lays out a brief sociolinguistic background to the contact languages in order to contextualize the potential language-external dynamics that might play a role in the

structural changes. This is because language change is probably an internal and an external process (Heine and Nomachi 2013), and 'structural factors assert themselves against the nonlinguistic factors present in the contact situation' (Johanson 2002:1). Therefore, the sociolinguistic aspect of the contact situation, i.e., the strong influence of the dominant languages (Turkish, Kurdish) on Anatolian Arabic, should be taken into consideration as well.

Moreover, this section also introduces the theoretical background for the phenomena that are the focus of the paper, and analyzes the contact-induced changes by making use of our current understanding of verb-building and argument structure.

2.1 The Linguistic Situation and Multilingualism

Anatolia, especially the (south)eastern part, has been home to many distinct linguistic groups (as well as ethnic and religious groups). Until the beginning of the 20th century, speakers of the largest Anatolian languages – Kurdish, Zazaki, Armenian, Aramaic, Arabic and Turkish – had been co-existing for almost a thousand years. This has naturally resulted in extensive contact among these languages. Contact influence on Anatolian Arabic and Kurdish has arisen mainly through long-term bi- and multi-lingualism rather than through language shift (in which speakers of other languages shifted to Arabic or Kurdish, Thomason 2001). As a result, when applicable, the changes seem to be through borrowing, rather than imposition (*à la* Coetsem 1988, 2000).

Turkish (TK), as the official language of Turkey, currently dominates public life in most Arabic-speaking and some Kurdish-speaking areas. However, as noted by Haig (2014) this was not always the case:

> "The current omnipresent influence of Turkish in the region is in fact a relatively recent phenomenon, fueled by compulsory Turkish-language state education, the mass-media, and large-scale military operations carried out by the Turkish army in the conflict against militant Kurdish groups. But prior to the twentieth century, the influence of Turkish in many parts of rural east Anatolia was negligible." (Haig 2014:14)

Northern Kurdish (Kurmanji, NK, of the Indo-Iranian branch) is spoken by the Kurds of Turkey, Syria and the northwestern perimeter of Iraq, in the province of West Azerbaijan in northwestern Iran and in pockets in the west of Armenia.[4] In Turkey, especially in Mardin and Siirt provinces, Kurds have been in contact with Arabic-speaking communities, but as the

[4] cf. Haig and Öpengin (2014) for a discussion on defining "Kurdish".

lingua franca of the communities of cultural–historical Kurdistan, Kurdish must have been the dominant language of interaction between these communities, and it is indeed possible to observe important influences from Kurdish on the local Arabic dialects (cf. Jastrow 2011).

Anatolian Arabic (AA) is part of the so-called *qəltu*-dialect branch of the larger Mesopotamian Arabic, and essentially refers to most of the Arabic dialects spoken in Turkey. AA dialects are characterized by (i) separation from the Arabic speaking world, (ii) contact with regional languages, which has affected them strongly, and (iii) multilingualism of speakers. These dialects are spoken as minority languages by speakers belonging to different ethnic or religious groups.

Moreover, not all of the Anatolian Arabic varieties are spoken *in situ*, and in fact some may no longer be spoken at all. Today the *in situ* Anatolian Arabic dialects are predominantly spoken by Muslims, and are also subject to a constant linguistic pressure from Turkish (the official language) and Kurdish (the dominant regional language), and social pressure for assimilation. The following quote from Grigore (2007a) succintly summarizes the overall context of Anatolian Arabic:

> *il se situe dans un microcontexte kurde, situé à son tour dans un macrocontexte turc, étant isolé de la sorte de la grande masse des dialectes arabes contemporains* (Grigore 2007a:27)[5]

Still, given that both Kurmanji and Anatolian Arabic speakers have been under Ottoman and Turkish dominance for centuries, it is not surprising that AA and NK show many traces of Turkish influence. The contact between NK-AA and TK has been mostly unidirectional, inasmuch as significant impact of Turkish has been shown on NK (Dorleijn 1996; Bulut 2006) and AA (Akkuş 2020). Particularly, Anatolian Arabic dialects exhibit interesting patterns due to language contact with Turkish and NK in every linguistic aspect.[6] This paper focuses on the verbal predicate formation (periphrastic) and word-internal language-mixing patterns (both affixal and templatic) in turn. Therefore, I first lay out the theoretical tools I employ in the next section.

The data for this study come from my own fieldwork mainly through acceptability judgments (as well as published sources) with four consultants. Fieldwork involved direct

[5] "It is situated in a Kurdish microcontext, which is in turn situated in a Turkish macrocontext, thus being isolated from the vast majority of contemporary Arabic dialects."

As Grigore puts it, in the whole region Kurds constitute the majority, and intermarriage is becoming more and more common, further increasing the influence of Kurdish. The majority of Arabs, especially those in Siirt, Batman, Mardin, Muş, also speak Kurmanji.

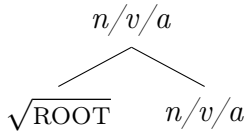
[6] See Jastrow 2011; Haig 2007; Haig and Öpengin 2014; Akkuş 2020; Öpengin 2012, a.o., for other contact-induced changes.

elicitation of (i) judgements about truth in particular contexts and (ii) judgements about felicity in particular contexts. All four consultants are from the Mutki region, Bitlis, Turkey. Two of the consultants are trilingual among a subvariety of Mutki-Sason dialect of Anatolian Arabic, Turkish and Northern Kurdish. Another consultant and the author of the study are bilinguals in the Arabic dialect and Turkish (with the author also being a heritage speaker of Mutki Zazaki). The two trilingual speakers are 54 and 60 years old, whereas the Arabic-Turkish bilinguals are 29 and 33 years old. The consultants regularly use each of the languages in their lives though not necessarily in the same settings (see the above discussion about the sociolinguistic situation). Although consultants report that they are *fully* competent in each language, they do note that some language(s) is 'weaker' than the other(s). For example, the 60-year old trilingual speaker notes that their Turkish is not 'as strong as' their Kurdish and Arabic. I have tried to take care to report any discrepancies or variations among the consultants (see also fn. 10).

2.2 The Theoretical Background to Word-formation

I adopt the DM-view that words are built out of the combination of an uncategorized root with functional material, which acts as the categorizer of this root. As such, a root becomes a noun, a verb, or an adjective once it is merged with a head bearing a category feature, *n*, *v* or *a*, respectively, (5), (see e.g., Arad 2003).

(5)

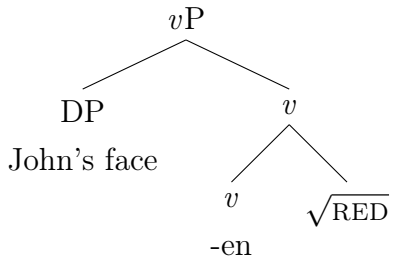


Therefore, a string such as *John's face redd-en-ed* would have the structure in (6b), in which the Root incorporates into *v*.

(6) *Affixal pattern*

a. John's face redd-en-ed.

b.



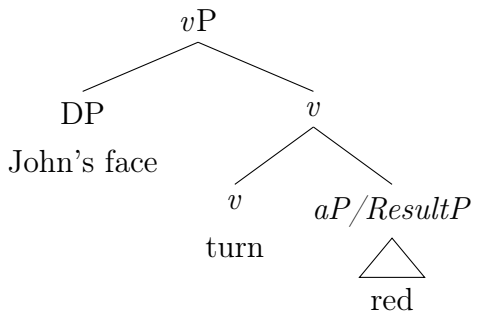
On the other hand, an analytical expression with the same type of Root will have the structure in (7b), in which *turn* is treated as a light-verb, i.e., the spell-out of *v*, similar to

the affixal *-en* (see Embick 2004 for discussion).[7]

(7) *LVC with particle verbs*

a. John's face turned red.

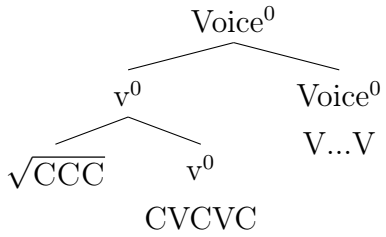
b.



When v's complement is an *aP* (or *ResultP* following e.g., Ramchand 2008; Alexiadou 2017), root incorporation does not take place, and v can be realized as a light verb.

For the syntax of non-concatenation in Arabic, I follow Arad 2003, 2005, according to which the root in Semitic is simply the consonantal root. The categorizing head v has the phonological instantiation of the template. This vP is in turn selected by Voice head that hosts the vocalic melody (in Hebrew, Arad 2005:190-1). Successive-cyclic head movement of the verb up through this structure yields the following head structure at Spell-Out, (8):

(8)



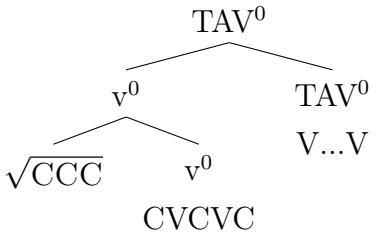
McCarthy (1979, 1981) and Tucker (2011) assume that in Arabic vocalic melody expresses tense, aspect and voice (thus TAV head), (9), presumably via a PF-Fusion rule for Arabic.[8]

[7] The fact that an *aP* appears in the case with the light-verb *turn* is clear from the fact that *red* can be modified – *John's face turned completely red*.

Note that although in English there may be an overlap between Roots that allows the v-Root vs the light-verb cases (similar to Turkish, e.g., *hasta-lan-* vs. *hasta ol-* 'to become sick'), in a language like Hindi, there is no overlap between these two sets of Roots.

[8] Whether this is true for all varieties of Arabic is debatable (see e.g., Benmamoun 2000 for Moroccan Arabic). In some cases, the vocalic melody is lost, but this does not lead to the loss of expressing tense/aspect, unexpected from a theory in which the vocalic melodies carry tense. As such, some varieties might be like Hebrew in which the vocalic melody expresses Voice *only*.

(9)



Besides these points, I also take code-switching / language-mixing to be an I-language phenomenon, i.e. an expression of a type of linguistic competence.[9] The speakers of the languages in question are highly competent bilinguals or trilinguals (rather than advanced second language speakers), by which I mean that their grammaticality judgments in each language do not differ from those of monolingual native speakers.[10]

Against this background, the next section discusses the verbal predicate formation in the contact languages, and demonstrates that they are theoretically informative for (at least) two respects: (i) head-directionality may switch as a result of language contact, and (ii) certain functional categories may be borrowed as semantically vacuous heads.

3 Verbal predicate formation

In the verbal domain, I investigate the borrowing strategy of the light-verb construction (LVC), leaving the discussion of the *templatic pattern* to section 4.

LVC is commonly attested in language mixing contexts (Muysken 2000; Myers-Scotton 2002; Edwards and Gardner-Chloros 2007; Wohlgemuth 2009; Bandi-Rao and den Dikken

[9] Code-switchers are able to produce consistent grammaticality judgments on sentences or phrases reported here, which reveals an underlying linguistic system.

It is worth noting that in the language contact literature, sometimes a distinction is drawn between code-switching and borrowing, which are taken to be specific kinds of language mixing. For some researchers, code-switching and borrowing are part of a continuum, whereas for others, they constitute two distinct processes. Borrowing is assumed to involve the full phonological and morphological integration of a word from one language (say, English *type*) into another (as Spanish *taipiar*), code-switching involves the mixing of phonologically distinctive elements into a single utterance, as illustrated in (i).

(i) ***Mi hija*** *cleaned the house.*
'My daughter ...'

Unless it matters at certain points of the papers, I ignore this potential distinction (see Muysken 2000; Matras 2009 for overviews).

[10] This should not be taken to mean these speakers are just like monolinguals in every aspect. Depending on the degree of competence a speaker has for each language, they might have subtle but perceivable features (e.g., pronunciation of certain sounds) in their speech that differentiate them from monolingual speakers.

As Balam et al. (2020) note, a fine-grained understanding of speakers across varieties could also be very informative, since there might be a community specific distribution for a pattern (thanks to María del Carmen Parafita Couto, p.c., for bringing this to my attention). The current study mainly focuses on the grammars of speakers belonging to the same community, leaving aside the significant issue Balam et al. (2020) raise to future investigation.

2014; Bağrıaçık et al. 2015; Alexiadou 2017, a.m.o.). In the German/Spanish code-switching pair in (10), the light verb *hacer* 'do' is taken from Spanish, while the lexical verb in the infinitival form *nähen* 'sew' is from German.

(10) Juan $hace_{[Spa]}$ **nähen**$_{[Ger]}$ das Hemd.
Juan does sew the shirt
'Juan sews the shirt.' (González-Vilbazo and López 2011:(1))

González-Vilbazo and López (2011) note various pervasive patterns regarding the LVC in language-mixing contexts. One of them is called *asymmetry*, which refers to the fact that the light verb can only be realized in one language (Spanish in (10)) and not in the other (German, e.g., *tun* or *machen*), thus the ungrammaticality of (11).

(11) *Hans $tut_{[Ger]}$ **coser**$_{[Spa]}$ la camisa.
Hans does sew the shirt
'Hans sews the shirt.' (González-Vilbazo and López 2011:(15))

We will observe that the asymmetry holds in the languages discussed in this paper as well; yet the explanation given for Spanish/German based on *conjugation class* in Spanish cannot extend to the situation of Turkish-Arabic-Kurdish (§3.4). While González-Vilbazo and López (2011) emphasize the view that this asymmetry in Spanish/German context does not correlate with the issue of 'language dominance', this may be a plausible explanation in the case of TK-AA-NK.

3.1 Light verb constructions in Turkish and Northern Kurdish

LVCs feature a prominent role in the grammar of Turkish and Northern Kurdish (NK). As in most Iranian and Indo-Aryan languages, both diachronically and synchronically, the LVC is one of the main strategies to create new verb meanings in NK.[11] NK employs a large number of light verb constructions (also called *complex predicates*, e.g., Gündoğdu 2018:76). LVCs consist of a nonverbal element and a light verb to form a single predicate. The verbal element in these predicates ranges over a number of typical simplex verbs such as *bûn* 'be/become', *kirin* 'do', *dan* 'give',[12] whereas the nonverbal element ranges over a number of categories

[11] For example, complex verbs have gradually replaced simple verbs in Persian since the 13th century (Folli et al. 2005). The tendency to form complex verbs has resulted in the existence of two sets of verbs, simple and complex, for a number of verbal concepts. In many cases, the application of the simple verb is restricted to the written and elevated language. (Folli et al. 2005:5).

[12] This type of grammatical borrowing is seen in an indirect causative construction, 'give'-causatives, in Sason Arabic (Gündoğdu and Akkuş 2022).

such as nouns (*av* 'water'), adjectives (*acis* 'bored'), particles (*ra*) and PPs (*ji bîr* 'from the mind'). e.g., *pacî kirin* (kiss do) 'to kiss'.

Turkish also has a considerably large number of complex predicates consisting of a nonverbal element followed by a light verb. The most common light verbs are *eD-*, which forms transitive and unergative predicates, and *ol-*, which forms unaccusative predicates (Kornfilt 1997; Öztürk 2005; Göksel and Kerslake 2005; Key 2013).

(12)

Transitive	Intransitive	
yok eD-	yok ol-	'to destroy/disappear'
hasta eD-	hasta ol-	'to get sick'

The majority of the bases are loanwords from Arabic and Persian, as shown in (13).[13]

(13)

kahr eD-	'to damn'
seyr eD-	'to watch'
kayd eD-	'to register'

3.2 LVC borrowing between Kurdish-Arabic and Turkish

The light verb construction is one of the domains in which the influence of contact between Anatolian Arabic, Turkish and Northern Kurdish is clearly manifested. The predominant strategy in both AA and NK is to resort to the LVC when borrowing verbs from Turkish, rather than adapting a foreign verb directly to their verbal morphology.

As the dominant languages in the region, Turkish and Northern Kurdish usually serve as the donor languages, while Anatolian Arabic is the recipient language. Therefore, there are a large number of compound verbs constructed with Arabic *sāwa – ysawi* 'to do, make' and a nominal borrowed from Turkish, (14a), or Kurdish, (14b).

(14) (*Tillo Arabic*, Lahdo 2009:202)

[13] Turkish itself also usually resorts to periphrastic construction when borrowing verbs from foreign languages, rather than adapting a foreign verb directly to Turkish verbal morphology (Akkuş 2015). Some examples are based on loanwords from English and French are given in (i).

(i)

motive eD-	'to motivate'
dizayn eD-	'to design'
kamufle eD-	'to camouflage'

This pattern is also reported in bilingual settings, e.g., Turkish/Dutch (Myers-Scotton and Jake 2009).

(ii) O diyor ben *uitmaken* yap-tı-m di-yor-du kızınam.
he says I finish.INF do-PST-1SG say-PROG-PST.3SG girl.with
'He says 'I broke up with a girl."

a. *sawa yārdım* 'to help' cf. Turkish *yardım eD-*
ysawaw dawām 'they continue ...' cf. Turkish *devam eD-*
nsayy qaḥwaltə 'we have breakfast' cf. Turkish *kahvaltı eD-*

b. *sawa brīndār* 'to injure' cf. Kurdish *brîndar kirin*
sawa ǧāmērtīye 'to act generously' cf. Kurdish *camêrtî kirin*
sawa ǧōt 'to mow' cf. Kurdish *ǧôt kirin*

The LVC has no obvious analog in other Arabic varieties, and thus, it is unlikely for the LVC to be a genetically inherited Semitic feature. In the majority of the cases, the construction is a complete *calque* of its Turkish or Kurdish counterparts (see e.g., Versteegh 1997; Lahdo 2009; Grigore 2007b; Talay 2007; Jastrow 2006; Akkuş 2016; Akkuş and Benmamoun 2018; Biţună 2016 for more examples). Therefore, the most plausible explanation for the AA light verb construction is that it appeared as a replication of the LVC in Turkish (and Kurdish).

The nonverbal part can be borrowed from Turkish as in (15), including reborrowing of Turkish loans that are themselves of Arabic origin, (15b), or Kurdish as in (16).[14]

(15) *Turkish*

a. qazan sawa
win do.PFV.3SG.M
'to win'

b. işāret sawa
sign do.PFV.3SG.M
'to sign'

(16) *Kurdish*

brīndār sir-tu.
injured become.PFV-1SG

'I became injured.'

The construction has spread beyond the loan vocabulary they had initially been part of and are now able – even if to a more limited extent – to attach to native lexical stock, (17). Thus, it has become an integral part of the morphological system of the recipient language (Arkadiev and Kozhanov 2021), as a result of a high degree of integration into the linguistic system.

[14] A similar strategy is seen in Neo-Aramaic as well, where the nonverbal element is from languages such as Persian, Kurdish or Turkish while the light verb is Neo-Aramaic. For example, Sorani Kurdish *seyr kirdin* 'to look' has been borrowed into Jewish Sulemaniyya Neo-Aramaic as *sayr ʔwl* 'to look', lit. 'to make sayr' (Stem I): the noun is borrowed but the light verb is Aramaic (Khan 2007: 209–210).

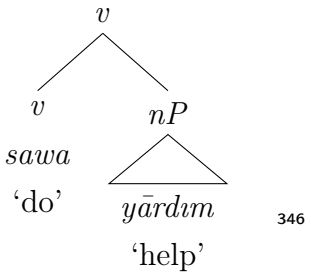
(17) *Sason Arabic*

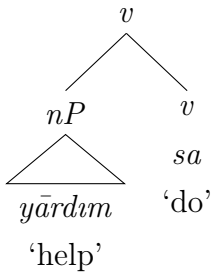
a. gerre/hās sawa
noise/sound do.PFV.3SG.M
'to make noise/sound'

b. şəɣle lā tə-s-i, aməl si!
talk NEG 2-do.IPFV-SG.M work do.IPFV.2M
'Don't talk, do work'

An important aspect of this mode of borrowing in the LVC concerns head-directionality. Whereas in most AA dialects, the Arabic head-initial pattern is maintained (cf. (14) for Tillo, Siirt), as schematized in (18), the default order in Sason Arabic, especially among younger speakers, have reversed. As such, it mainly manifests a head final order, (19), undoubtedly due to contact with Turkish and Kurdish, which are head-final languages.[15] Akkuş and Benmamoun (2018) argue that copular clauses in Sason Arabic have also adopted a head-final pattern (unlike the head-initial verbal clauses) due to language contact. Thus, this study adds to the range of constructions that have undergone this shift in Anatolian Arabic.

(18) Tillo: head-initial ⇒ (19) Sason: head-final





Note that NK also incorporates borrowings from Turkish in the LVC strategy, as in (20).

(20) *say*$_{\text{TK}}$ *kirin*$_{\text{NK}}$
count do.INFV
'to count' (NK) cf. Turkish *say* 'count'

However, as noted above, the LVC is a solid part of the NK grammar, unlike AA, which has developed this strategy as a result of the neighboring languages. Regardless of the status of the LVC in AA and NK, their loan adaptation strategy from Turkish is parallel.

[15] Uzbekistan Arabic underwent a similar head-directionality change due to Uzbek influence. It is also worth noting that Arabic does allow OV pattern, but the Object is in the left periphery either as a focused element, or topicalized constituent triggering resumption. The LVC differs from those instances. Thanks to Enoch Aboch, p.c., for raising this point.

Before we proceed with the next section, it is worth noting that the LVC in AA (at least Sason Arabic) behaves similarly to its Kurdish and Turkish counterparts in other respects too. First, some of the complex predicates can take a direct object besides the nonverbal element. Consider (21). Note that while the order is [NVE-LV-DO] in (21), the positioning of the DO is more flexible and the order [DO-NVE-LV] is also possible (see Akkuş 2014; Akkuş and Benmamoun 2016; Akkuş and Öztürk 2017 for more discussion as to the order of direct objects).

(21) qazan a-si ali lope.
win 1SG-do.IPFV this game
'I will win this game.'

Moreover, the nonverbal element itself shows properties of a phrasal element, again similar to Turkish (Öztürk 2005) and Kurdish (Gündoğdu 2015) (see also Dabir-Moghaddam 1997; Folli et al. 2005; Karimi 1997 for Persian). A focus particle can separate the LV and the nonverbal element, (22a); additionally, two nonverbal elements can be coordinated, (22b).[16]

(22) a. qazan *dāhi* a-si ali lope.
win even 1SG-do.IPFV this game
'I will even win this game.'

b. ser u qazan a-si ali lope.
watch and win 1SG-do.IPFV this game
'I will watch and win this game.'

3.3 Patterns of LVC borrowing with *-mIş* into Kurdish and Arabic

This section examines the LVC borrowings in NK and AA that involve instances with the perfective suffix *-mIş* from Turkish, and demonstrates that it has a non-uniform status across varieties. The investigation reveals that whereas *-mIş* is semantically vacuous in NK and many dialects of AA, it still bears semantic function for some speakers of AA, including Sason Arabic.

[16] While it is not a crucial part of this paper, for the AA, I tentatively adopt the structure proposed by Öztürk (2005) for Turkish LVCs. According to that analysis, the light verb and the nonverbal element, nP, form a complex predicate, which in turn combines with the syntactic object. This analysis is appealing from the perspective of semantic composition too. For object incorporation in Chamorro, Chung and Ladusaw (2003) argue that the incorporated element *restricts* (i.e., modifies) the argument position, which then gets *saturated* by the direct object.

3.3.1 *-mIş* forms in NK

A striking regularity in Kurdish is that many of the items borrowed into the LVC are Turkish verb forms with the suffix *-mIş*, as illustrated in (23).[17]

(23) *annamiş kirin* 'to understand' cf. Turkish *anlamış*
dinlemiş kirin 'to listen' cf. Turkish *dinlemiş*
sevinmiş bûn 'to be happy' cf. Turkish *sevinmiş*

As noted in Haig (2007:175), *-mIş* verb forms used in Kurdish have lost their perfective participial sense, and certainly do not have any sense of evidentiality. In effect, they are a tense-neutral kind of action nominal. An example is provided in (24).

(24) ... mere pera je kazan-miş di-kî
we money.PL from earn-*mIş* IND-do.PL
'... we earn money' (Çabuk 2019, p. 863)

In addition to the well-known borrowing of *-mIş* verb forms, NK has also borrowed Turkish bare verb stems. A few examples are in (25) (in addition to (20)):

(25) *inan kirin* 'to believe' cf. Turkish *inan-*
bekle kirin 'to wait' cf. Turkish *bekle-*
kapat kirin 'to close' cf. Turkish *kapat-*
(Haig 2007:175)

Therefore, the presence or absence of the *-mIş* suffix appears to have no impact on meaning; and even when *-mIş* is present, it is semantically vacuous in Kurmanji. The forms with this suffix can be used in all tense-aspects, including those in non-completive aspect or future interpretations.

3.3.2 *-mIş* forms in Anatolian/Sason Arabic

The pattern of light verb borrowing is also attested in AA, where the complex predicate comprises *sawa* 'do, make' and Turkish verbal form with *-mIş*, rather than the bare form of the verb for the most part.

[17] The borrowing of Turkic *-mIş* verb forms into Iranian languages is a well established phenomenon, found also for example in Tadjik (through contact with Uzbek) or Tati (via contact with Azeri).
This strategy is not limited to Iranian languages and AA, and is attested in other languages in contact with Turkish as well. As reported in Bağrıaçık et al. (2015), it is observed in various Armenian dialects (Vaux 2001) including Hemşin, Aslanbeg; Kabardian (spoken in Uzunyayla, Alagozlu 2017).

(26) *sawa gaçınmış* 'to manage, get along' cf. Turkish *geçinmiş*
başlamış sawa 'to begin' cf. Turkish *başlamış*

Interestingly, while the *-mIş* suffix is optional in some borrowed verb-forms, its presence or absence does lead to a meaning difference, unlike NK. For example, the root *qazan* 'to win, earn' (cf. Turkish *kazan*) can be used with or without *-mIş* in the light verb strategy, (27). Crucially, whereas the bare form can be used in all tense/aspects, the form with the *-mIş* suffix strongly implicates a completive event for many speakers of Sason Arabic.

(27) a. ɣade qazan a-si.
tomorrow win 1SG-do.IPFV
'Tomorrow I will win.'

b. ɣade qazan-mış a-si.
tomorrow win-*mIş* 1SG-do.IPFV
'Tomorrow I will have won.'

The following contexts in (28) further highlight the semantic import of *-mIş*.

(28) a. çax le tı-çi, qazan a-si.
time that 2M-come win 1SG-do.IPFV
'When you come, I will (start to) win.' ⇝ I will win only after you arrive.

b. çax le tı-çi, qazan-mış a-si.
time that 2M-come win-*mIş* 1SG-do.IPFV
'When you come, I will have won.' ⇝ I will have won by the time you arrive.

Another example is provided in (29). The adverbial phrase headed by *mı* 'from' is only compatible with non-culminated events, as such with the bare form of the borrowed verb (*yaşa* 'live'), as in (29a). On the other hand, its counterpart lacking the preposition is used with completed events as in (29b).

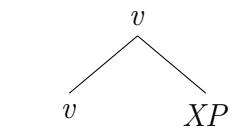
(29) a. fı Istanbul yaşa i-si *(mı) aşrin sınin.
in Istanbul live 3M-do.IPFV from twenty years
'He has been living in Istanbul for 20 years.' ⇝ He still lives in Istanbul.

b. fı Istanbul yaşa-mış sa (*mı) aşrin sınin.
in Istanbul live-*mIş* do.PFV.3M from twenty years
'He lived in Istanbul for 20 years.' ⇝ He no longer lives in Istanbul after having lived there for 20 years.

3.4 An analysis of the *-mIş* patterns

In order to explain the LVC borrowings in NK and AA, I adopt the structure in (30) (cf. §2.2), in which the complement of v is an XP, and v is realized as a light-verb since no incorporation takes place (see Embick 2004; Öztürk 2005; Folli et al. 2005; Ramchand 2008; Alexiadou 2017, a.o.).[18]

(30)



light-verb

Starting with the example in (31a), this structure is interpreted as a nominal as it merges with an n head, thus an LVC with the suffix has the structure in (31b). In this structure, Asp head is both morphologically and semantically contentful. The whole complex serves as the complement of the little-v that is realized as a light verb.[19]

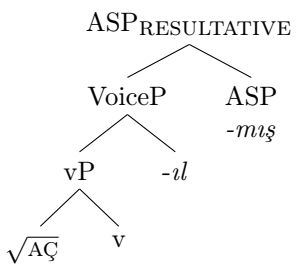
(31) a. ɣade qazan-mış a-si.
tomorrow win-*mIş* 1SG-do.IPFV
'Tomorrow I will have won.' (Sason Arabic)

[18] Key (2021) treats the light verbs *eD-* and *ol-* as the fusion of v/Voice heads, with the former having an active Voice whereas the latter a non-active Voice. This implementation is also compatible with the facts.

[19] This structure is very similar to that proposed for adjectival participles or nominalizations involving resultative aspect crosslinguistically, especially for the complement of n. For concreteness, I adopt the one proposed for Turkish by Gürer 2014. Consider (i).

(i) a. aç-ıl-mış
open-VOICE-PTCP
'opened'

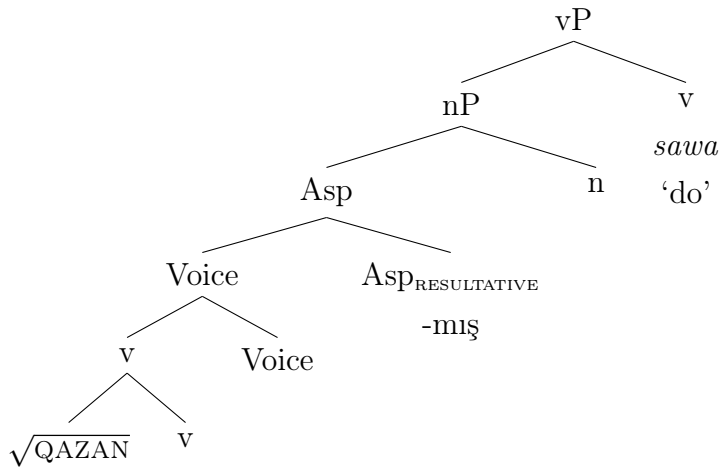
b.



'open'

(Gürer 2014:7)

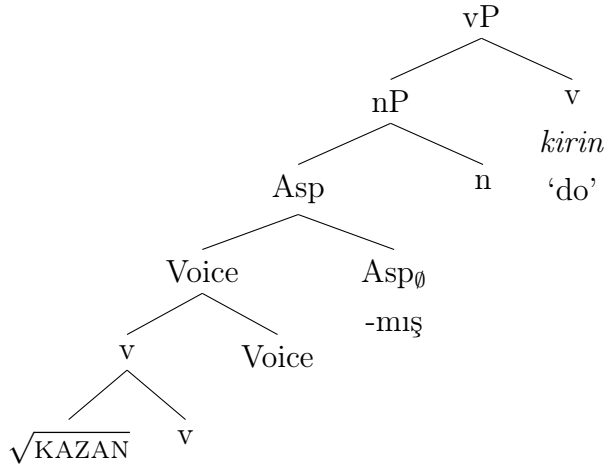
b.



In the case of NK, on the other hand, examples with the *-mIş* suffix are semantically vacuous (cf. Marantz 2013). Thus, they can be regarded as the zero semantic counterparts of phonologically empty *v* heads in cases of allomorphy (Embick 2010). These semantically empty heads are identical to their bare counterparts in NK. "Semantically zero" allosemes or semantically ∅ *v* have been proposed for Greek participles by Anagnostopoulou and Samioti (2014).[20]

(32) a. *kazan-mış kirin* 'to win' (Kurmanji)

b.



Recall that in NK, certain nonverbal elements bear the *-mIş* suffix even though it is semantically expletive, while others lack this morpheme. I implement this by proposing that these lack an Aspect projection in parallel to Alexiadou et al.'s (2015a) treatment of marked versus unmarked anticausatives. On their analysis, marked anticausatives have a semantically vacuous (expletive) Voice head with overt marking, while unmarked

[20]For allosemy of other heads, such as *Voice, Appl, p*, etc., see Wood (2015, 2016); Marantz 2013; Myler (2016).

anticausatives have no Voice layer at all.[21]

Recall that for many speakers of SA, the suffix is associated with the perfective/completive aspect, on the other hand, the null version is underspecified. As such, we can give the following Vocabulary Items in (33).[22]

(33) Vocabulary Items for Aspect in the LVC of Sason Arabic

a. $\text{Asp}_{\text{PERFECTIVE}} \leftrightarrow$ -*mış*

b. Asp $\leftrightarrow$ -$\emptyset$

Thus far, this section has shown that in the LVCs, head-directionality may change as a result of language contact, and that certain functional heads can be borrowed as semantically vacuous. Next, I turn to the generalization noted at the beginning of section 3.

3.5 A solely intra-linguistic/structural explanation for the asymmetry?

At the beginning of §3, we noted the asymmetry that in language-mixing contexts that involve the LVC, the light verb consistently comes from language A, whereas the nonverbal part from language B, e.g., nonverbal part from German, whereas the light-verb from

[21] For the unmarked anticausatives, they reject the option of a Voice layer that is both semantically and phonologically empty on the grounds that such a structure is unlearnable: "Since expletive Voice has no semantic contribution, language learners can only acquire it, if there is morphological evidence. Otherwise, they will not have any reason to assume an expletive projection" (Alexiadou et al. 2015a:109). Thanks to an anonymous reviewer for reminding me of this discussion.

That said, besides the learnability, I am not aware of a convincing argument that rules out the (notwithstanding less economical) alternative, which is to say that Aspect head has multiple Vocabulary Items. The winner is determined by a morpheme local to the morpheme undergoing insertion, resulting in contextual allomorphy, as illustrated in (i). For example, in the context of the root $\sqrt{\text{DINLE}}$ 'listen', the -*mIş* allomorph of Asp is inserted.

(i) Vocabulary Items for Aspect in the LVC of NK

a. $\text{Asp}_{\emptyset} \leftrightarrow$ -*mış* / $\{\sqrt{\text{DINLE}}, \sqrt{\text{SEVIN}}, ...\}$_

b. $\text{Asp}_{\emptyset} \leftrightarrow$ -$\emptyset$ / $\{\sqrt{\text{INAN}}, \sqrt{\text{BEKLE}}, ...\}$_

This would be essentially the same idea standardly applied for the past Tense head in English, (ii), which is contextually realized depending on a particular Root.

(ii) Vocabulary Items for past tense in English (T[past]) (Embick and Marantz 2008:4)

a. T[past] $\leftrightarrow$ -t / $\{\sqrt{\text{LEAVE}}, \sqrt{\text{BEND}}, ...\}$_

b. T[past] $\leftrightarrow$ $\emptyset$ / $\{\sqrt{\text{HIT}}, \sqrt{\text{QUIT}}, ...\}$_

c. T[past] $\leftrightarrow$ -ed

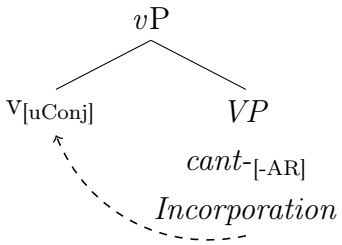
[22] From a diachronic perspective, the comparison of NK and AA reveals that the replication of this structure into NK precedes its counterpart in AA. Most likely, over time the perfective semantics of *mIş* is bleached and lost, and NK is on a later stage of this path.

Spanish, as repeated in (34), with the reverse pattern being ungrammatical.

(34) Juan *hace*$_{[\text{Spa}]}$ nähen$_{[\text{Ger}]}$ das Hemd.
Juan does sew the shirt
'Juan sews the shirt.' (González-Vilbazo and López 2011:(1))

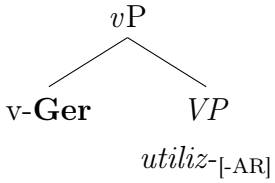
González-Vilbazo and López (2011) emphasize the view that this asymmetry in Spanish/German context does not correlate with the issue of 'language dominance'. Instead, they propose that the LVC strategy is a *last resort* operation, making use of the contrast between Spanish and German in terms of the *conjugation class*. Spanish verbs belong to one of three conjugation classes (the *-ar, -er,* and *-ir* classes), e.g., *cant-**ar*** 'sing', *beb-**er*** 'drink', *viv-**ir*** 'live'. They argue that v in Spanish bears unvalued features for conjugation class. In order to value this feature V-to-v movement takes place, (35).

(35)



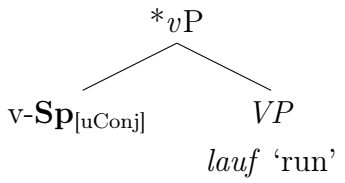
A Spanish verbal root can be embedded and incorporate into a German v, which is unspecified for conjugation class, (36).

(36) ✓ Spanish lexical verb + German inflection

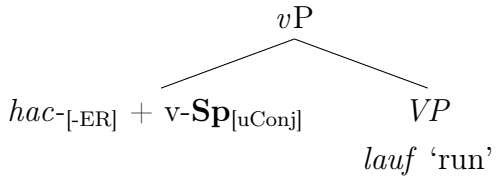


However, German verbal roots have no conjugation class. Since they do not have a conjugation class, they cannot satisfy the [uConj] feature of v in Spanish, leading to ungrammaticality, as shown in (37a). As such, only the insertion of a light verb like *hacer* 'do', that bears the class feature, can make this configuration grammatical, (37b).

(37) a. *German lexical verb + Spanish inflection



b. ✓ German lexical verb + Spanish *hacer*



This analysis relies on the Spanish little v establishing a syntactic dependency with a verbal root with a specification for conjugation class. Thus, it is an Agree relation between v, a probe, and the verbal root, which qualifies as a goal with matching features.

Leaving aside the details of the analysis, it is evident that due to its language-specific nature, it cannot carry over to the contact situation involving languages like Turkish, Arabic and Kurdish. None of these languages are specified for conjugation class, thus an analysis based on conjugation class cannot be the explanation for the asymmetry.[23] As an alternative, without qualifying it much, I hypothesize that the sociolinguistic aspect of the contact situation should be taken into consideration as well, i.e., the degree of contact between the donor and the recipient languages, and the strong influence of the dominant languages (Turkish, Kurdish) on Anatolian Arabic.[24] As such, the recipient language is the one which provides the light verb, which is always Anatolian Arabic. Therefore, when AA borrows from NK or TK, the light verb is consistently of Arabic origin. By the same logic, when NK borrows from Turkish, the nonverbal element is from Turkish, whereas the light verb is Kurdish.[25] It is also plausible that a combined account that involves both intra- and extra-linguistic factors or developments could provide a wide enough perspective for crosslinguistic patterns (cf. Melissaropoulou 2016).

3.6 Section Summary

This section has investigated LVCs as part of the verbal predicate formation in the language-contact situations of Anatolian Arabic, Northern Kurdish and Turkish.

This construction is of theoretical interest for (at least) two reasons: (i) It provides a case-study in which the head-directionality within a specific construction changes along the lines of the source language. As such, in some Arabic dialects, a head-final pattern is taking over following Turkish and Kurdish patterns. (ii) It also provides an illustration in which a functional category may be borrowed as semantically vacuous in some varieties/languages,

[23] This analysis would also face challenges in a language contact situation that involves two languages with conjugation classes, e.g., Portuguese and Spanish. Thanks to Jaklin Kornfilt (p.c.) for this point.

[24] See Bağrıaçık et al. 2015 regarding the need to describe this factor quantificationally.

[25] The same statement can be extended to Turkish as well; when borrowing verbs from foreign languages, Turkish also uses the light verb strategy, as shown in footnote 13, with the light verb coming from Turkish.

while it maintains its semantic function in others.

4 The templatic and affixal borrowing

This section investigates the word-internal language-mixing as a result of templatic and affixal modes of borrowing. Particularly, in the templatic pattern, loans are incorporated into the verbal and nominal vocalic patterns of the recipient Semitic language.

Most of code-switching studies, for obvious reasons, focus on bilingual speakers. This section provides naturally-produced examples from trilingual speakers, which includes language-mixing of three contact languages. The examination of these patterns reveals that neither the Functional Head Constraint of Belazi et al. (1994) nor the phase-based account of López et al. (2017) capture the full range of data. Alternatively, I propose *No-Reversal Constraint*, which is elaborated further in this section. I first introduce the patterns of language-mixing in the verbal and nominal domain, and then provide the structures to be associated with those patterns, along with their implications for the above-mentioned models of language-mixing.

4.1 The templatic pattern in the verbal domain

Besides incorporating verbal forms via the light verb construction (cf. §3), some borrowed verbal forms have been totally assimilated to the Arabic verb system.

In such cases, when a verb lexeme is borrowed into a Semitic language, a tri- or quadriradical root must first be identified and extracted. Then the root is assigned to one of the patterns/derivations. These verbs are in major part formed according to the II. or III. verbal stem, as in (38), where the perfective and imperfective forms of borrowed verbs are shown, respectively. For example, the Turkish verb *kapat* 'close' is borrowed into Anatolian Arabic in Template/Stem II, as such its perfective form is *qappat*, whereas its imperfective form is realized as *īqappət*.[26]

(38) (Āzəx variety, Talay 2007)

Stem II	*qappat – īqappət* 'to close'	cf. Turkish *kapat-*
Stem II	*qayyad – īqayyəd* 'to register'	cf. Turkish *kayıt eD-*
Stem III	*ḍāyan – īḍāyən* 'to be patient, to bear up'	cf. Turkish *dayan-*

Consider the verbal form *karış* 'to meddle, mix', which is also incorporated as Stem II

[26] The second example involving 'to register' is most likely a *reborrowing*, in that although Standard Arabic has a Stem II form *qayyada/yuqayyidu*, the speakers of AA are not familiar with that form. The Standard Arabic form was borrowed into Turkish, which was then in turn borrowed into AA varieties.

in Sason Arabic. This verb is incorporated into the phonology of the language, where velar stop becomes uvular stop $k > q$. Additionally, it is used in different templates, inflected for person-number and different aspects. Therefore, it is fully integrated into the language, both phonologically and morphologically. (39) illustrates the verb inflected for feminine gender in both aspects, and (40) does the same for the masculine, in Stem II.

(39) *feminine; imperfective-perfective*

a. kul çax tı-qarreş fı şıɣlı-di
every time 3F-meddle.IPFV in work-my
'She always meddles in my work.'

b. ams qarş-e fı şıɣlı-di
yesterday meddle.PFV-3F in work-my
'She yesterday meddled in my work.'

(40) *masculine; imperfective-perfective*

a. kul çax i-qarreş fı şıɣlı-di
every time 3M-meddle.IPFV in work-my
'He always meddles in my work.'

b. ams qaraş fı şıɣlı-di
yesterday meddle.PFV.3M in work-my
'He yesterday meddled in my work.'

Thus, some Turkish loan-verbs are borrowed into the Arabic system in a way that only contains the consonantal tier, which in turn is combined with the template. Consonants of a Turkish loanword are retained; yet, the template and the vocalic melody (if any) come from Arabic.[27]

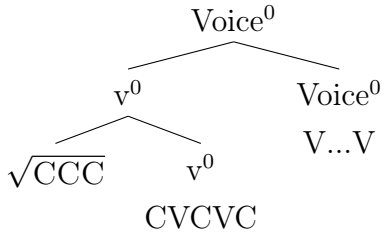
Let us now consider these patterns against the backdrop of the syntax Arad (2003, 2005) proposes for non-concatenation. Arad proposes that the root in Semitic is simply the consonantal root. The categorizing head v has the phonological instantiation of the template. This vP is in turn selected by Voice head that hosts the vocalic melody (in Hebrew, Arad

[27]Similar issue arises in borrowings into other Semitic languages as well. For example, in the Telkepe dialect of Neo-Aramaic spoken in Detroit, the English word *charge* has been borrowed as *çrj* 'to charge' in Stem II.

(i) kə-mçarj-i-lə.
IND-charge.PRS-3PL-OBJ.3MS
'They charge it up.' (Coghill 2015:85,(1))

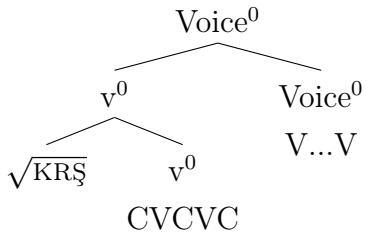
2005:190-1).[28] Successive-cyclic head movement of the verb up through this structure yields the following head structure at Spell-Out, (41):

(41)



The templatic borrowings can be made sense of given this structure: 'karış' would be borrowed as a Root, as in (42), with only the consonantal tier KRŞ preserved, which in turn combines with the template in *v* and the vocalic melody in Voice head.[29] This state of affairs is not surprising given that Turkish lacks the root-and-pattern morphology, and as such the lexical item is borrowed (or re-analyzed) into a node that could have its counterpart, which in this case is the Root.

(42)



Let us now turn to the templatic and affixal borrowings in the nominal domain.

4.2 The templatic pattern in the nominal domain

As in the verbal domain, borrowing via the templatic pattern is attested in the nominal domain as well. This can be seen most clearly in plurals, which can be realized in particular templates known as broken plurals, as well as regular plurals which are realized as a suffix. For this reason, I first introduce the types of plurals and the structures associated with them in (Anatolian) Arabic, before proceeding with the discussion of modes of borrowing attested in the nominal domain.

[28] I assume that prosodic constraints result in the interleaving of the vowels within the consonants such that the syllabic template is emergent/epiphenomenal (see Tucker 2011 for a DM–based analysis in this vein, and many references therein).

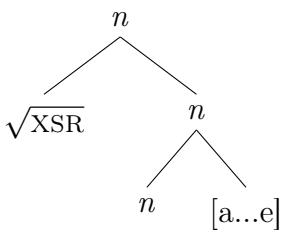
[29] As mentioned earlier, the vocalic melody in Arabic possibly expresses tense, aspect and voice (see McCarthy (1979, 1981) and Tucker (2011)), which according to Tucker would be a TAV head realized via a PF-Fusion rule.

4.2.1 Irregular, broken plurals

Starting with the irregular, broken plurals, we observe that older loans into AA from Turkish have been integrated into the language, both phonologically and morphologically, e.g., by acquiring their own broken plurals, e.g. (Mḥallami, Āxez) *panṭur/panēṭīr* 'trousers' (< Turkish *potur*) (Wittrich 2001). This is the same pattern observed for certain native items that have broken plurals, e.g., *sınnor* - *sınēnīr* 'cat(s)'. In this case, there can be no doubt that we are dealing with a real loan.[30]

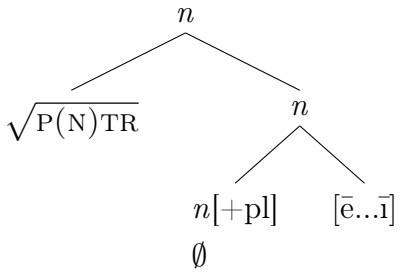
I assume (with Arad 2005 for Hebrew, and Kramer 2016:531 for Amharic) that nominal vocalic patterns for Semitic consonantal roots are inserted at PF. They are inserted at a node adjoined to n (cf. Arad's analysis of verbal patterns (inserted at v, as in §4.1) and DM-approaches to declension classes), although other implementations are also possible. As such, a noun like *xaser* 'yoghurt' consists of the root $\sqrt{\text{XSR}}$ and a null n to which is adjoined the vocalic pattern [a...e]. Consider (43). Similar to the verbal domain, again, prosodic constraints result in the interleaving of the vowels within the consonants.

(43)



Forms involved in language-mixing would be part of a situation in which Root comes from language A, and n (and higher heads) realized in language B. In the case of the loanword 'trousers', the root of the lexeme *po(n)tur* containing the consonantal tier $\sqrt{\text{P(N)TR}}$ comes from Turkish, whereas the n and the vocalic melody come from Arabic. Since nouns can take different patterns in the plural, each plural is taken to be the allomorph of the pattern itself conditioned by plurality, and in the case of 'cat' or 'pants', that n[+pl] is null. Consider (44).

(44)



[30] Armostis and Karyolemou (2023:8-9) report a similar pattern of borrowing for Cypriot (Maronite) Arabic both in the verbal and nominal domain. As such, only the consonantal tier of loanwords are retained, and they are integrated into the relevant template and vocalic melody of Arabic.

Note that in AA, the realization of the irregular plural is not limited to broken plurals; it can also be in the form of the idiosyncratic suffix *-in*, (45). For these instances, I will simply assume that we have an allomorph of *n*[+pl] realized as *-in*, and no adjunction of vocalic melody. (45) also shows that the irregular plural can also attach to Turkish and Kurdish roots (see fn. 32).

(45)

refiq 'friend'	*refq-in* 'friends' (Tillo Arabic)
angir 'friend'	*angir-in* 'friends'
genc 'young one' (cf. Turkish *genç*)	*genc-in* 'young ones'
piroz 'blessed, holy' (cf. Kurdish *pîroz*)	*piroz-in* 'the blessed, holy (ones)'
delal 'dear, lovable' (cf. Kurdish *delal*)	*delal-in* 'lovable (pl.)'[31]

4.2.2 Regular and double plurals

In addition to the broken (irregular) plurals, Arabic (including Anatolian Arabic varieties) also has the regular plural and double plurals.

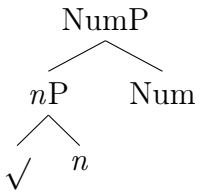
Many nouns are pluralized with the regular plural suffix *-ad*, *-at*, (or *-að*), as in (46).

(46)

horti 'calf'	*horti-yad* 'calves'
kartol 'potato'	*kartol-ad* 'potatoes'
badıncan 'tomato'	*badıncan-ad* 'tomatoes'

Following Kramer (2016), I take it that regular plurals are the realization of Num (thus, *-ad* is on Num head), whereas irregular plurals are the realization of *n*. The combination results in a structure like (47).

(47)



In this split analysis of plurality, regular plurals are formed via a Num[+pl] combining with a nP, whereas irregular plurals are formed by combining a *n*[+pl] with the root. In addition to capturing the difference between regular and irregular plurals, this analysis also correctly predicts (i) the presence of irregular and regular plural morphology on the same nominal without any change on meaning (cf. Arabic (Zabbal 2002), Amharic (Kramer 2016), Breton (Trépos 1957, as cited in Kramer 2016) and other languages), as well as (ii) their

[31] In Tillo Arabic, it is also found as *dalel-in* (Lahdo 2009:93).

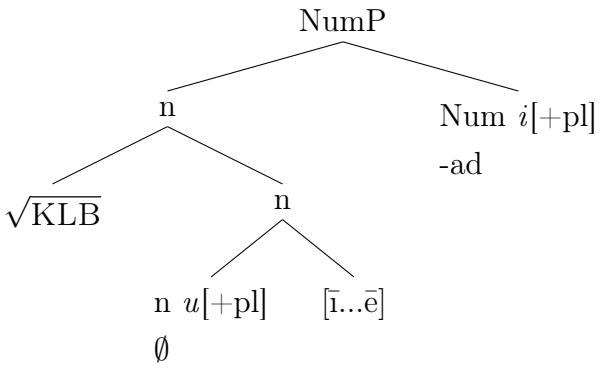
various properties, including morpheme ordering, with the only possibility being $\sqrt{}$-*Irreg-Reg*. The regular plural morpheme *-ad* can be attached to the broken plurals, as in (48a), or to those already bearing the irregular plural morpheme *-in*, as in (48b).

(48) a. *kelp* 'dog' - *kɯleb* 'dogs' - *kɯleb-ad* 'dogs'

b. *genc* 'young one' - *genc-in* 'young ones' - *genc-in-ad* 'young ones'[32]

Intuitively, we are dealing with two plural exponents, but only one interpretation of plurality. For these kinds of instances, I adopt Kramer's (2016) analysis, in which only one of the plural features is interpretable (see also Paparounas and Akkuş (2023) for the same analysis for the so-called Default Triggering Nominals in Turkish such as *biz-ler* 'we-PL'). As such in a double plural, the plural feature on *n* must be uninterpretable, while the one on *Num* is interpretable. This is illustrated in (49).

(49)



Against this backdrop, let us now turn to the patterns of word-internal language mixing, in which we observe that various morphemes can come from different languages. Then we undertake the task of explaining the restrictions on these patterns.

4.2.3 Word-internal Language-mixing and *No-Reversal Constraint*

The plural patterns noted thus far in this section reflect the grammars of monolingual speakers, in which morphemes come from a single language. However, we observe that the primary difference between monolingual grammars versus language-mixing contexts (be

[32] Another way to show that *-in* in forms like (45) and (48b) is an irregular plural morpheme (besides the morpheme order) is that it can never follow broken plurals unlike the regular *-ad*, thus **kɯleb-in*, which would be expected from a regular plural.

As noted by an anonymous reviewer, a more correct characterization of the broken plurals and the *-in* plural - which are in complementary distribution, would be *low-attaching plural*, while the *-ad* plural would be *high-attaching plural*. Acknowledging this, I still maintain the use of regular vs irregular hoping that they are more theory-neutral, although their productive/unproductive nature serves as one of the motivations for low versus high-attaching plural. See Kramer 2016 and references therein for more discussion.

it bilinguals/trilinguals or borrowing) is that in language-mixing contexts, a root can be chosen from language A, whereas the realization of the categorizer or a higher functional head from language B. This is illustrated in (50) for a German-Spanish language-mixing situation, in which the root $\sqrt{\text{UTILISIER}}$ 'use' comes from Spanish, whereas the (dissociated) morpheme corresponding to 1pl *-en* comes from German.

(50) Wir *utilisieren* spanische Wörter, die dann *alemanisiert* werden y hacen klingen
We use Spanish words that then Germanized are and do sound
un poco raro.
a bit strange
'We use Spanish words, that are then Germanized and sound a bit strange.'
(González-Vilbazo and López 2011:(2))

A similar situation is observed in the context of Turkish, Anatolian Arabic and Northern Kurdish. (51) provides a commonly-attested pattern in which the Arabic regular plural *-ad* is attached to Turkish stems.

(51) a. *mılyar-āt* 'billion-PL' (Tillo Arabic)

b. *damad-ad* 'son-in-law-PL'

c. *soba-d* 'stove-PL'

The reverse pattern is also attested: it is possible to replace the Arabic plural morpheme realized on the Num head with its Turkish counterpart *-lAr* in spontaneous speech, which even obeys the vowel harmony.[33] Note that this is attested in both regular, (52a), and double plural forms, (52b), where *kıleb* 'dogs' has the broken plural pattern, and is attached with the Turkish plural suffix.

(52) a. *xatan-lar* 'son-in-law-PL'

b. *kıleb-ler* 'dog.PL-PL'

These examples show that the root (and maybe root + a higher functional head) might come from language A, whereas the top node, Num, comes from language B.

Attempts to add outer morphemes are informative regarding potential restrictions that may apply to word-internal language-mixing. An attempt to add the possessive marker to the form in (51) is allowed only if the possessive morpheme is Arabic, as illustrated in (53)

[33]It is interesting to note that some speakers produce such forms in natural-speech, but when asked about them explicitly, they disfavor some of the forms.

(morphemes are added the subscript notation that shows which language a morpheme comes from).

(53) a. soba-d-i
stove$_{\text{TK}}$-PL$_{\text{AA}}$-1SG.POSS$_{\text{AA}}$
'my stoves'

b. cf. *$soba_{TK}$-d_{AA}-$ım_{TK}$

A similar restriction is observed in (52) involving the Turkish plural suffix as the outer-most morpheme. Any further morpheme can only be drawn from the Turkish inventory, and not Arabic. Consider (54).

(54) a. xatan-lar-ım
son-in-law$_{\text{AA}}$-PL$_{\text{TK}}$-1SG.POSS$_{\text{TK}}$
'my sons-in-law'

b. cf. *$xatan_{AA}$-lar_{TK}-i_{AA}

Looking at the patterns in (53) and (54), one plausible explanation is that whichever language the highest (functional) node is from, any additional morpheme has to be from that language (under the assumption that roots do not enter into the calculus). In the examples in (51), the highest morpheme is from Arabic, as such an additional node also needs to be from Arabic, and its Turkish counterpart is disallowed, (53). By the same token, it would be the case that in (52), since the highest morpheme is from Turkish, any additional morpheme must also come from Turkish, and thus (54).

Another pattern of word-internal borrowing that would be in line with this statement comes from examples involving derivational suffixes. These are examples in which a nominal/adjectival root can be incorporated into the phonological system of the language A, and then attached with a derivational suffix from language B. Consider (55) and (56), in both of which, the Root comes from one language and the categorizer n from the other (noting that the forms in (56) are not as salient for speakers as (55) - I leave aside the source of this non-categorical asymmetry in preference).

(55) *salak* 'stupid' (Turkish root) + *-tiye* 'state of' (Arabic suffix) = *salaq-tiye* 'stupidity'

(56) *hamar* 'donkey' (Arabic root) + *-lık* 'state of' (Turkish suffix) = *hamar-lık* 'asininity, stupidity'

This pattern also obeys the restriction noted above. The addition of a plural morpheme needs to be in the language exponed at the outer-most node. Therefore, the addition of

the plural morpheme to (57) would need to be the Arabic *-ad*, and not the Turkish *-lAr*. Conversely, the plural added to the form *hamar-lık* 'asininity, stupidity' in (58) has to be the Turkish *-lAr*. Examples in (57c) and (58c) show that another morpheme, such as the ablative case *-DAn*, is also possible as long as it follows the restriction noted thus far. (58d) shows that replacing the Turkish ablative case with the Arabic preposition (as Arabic lacks overt case) is grammatical.

(57) a. hamar-lık-lar
donkey$_{AA}$-DER$_{TK}$-PL$_{TK}$
'stupidities'

b. cf. *hamar$_{AA}$-lık$_{TK}$-ad$_{AA}$

c. hamar-lık-tan
donkey$_{AA}$-DER$_{TK}$-ABL$_{TK}$
'from the stupidity'

(58) a. salaq-tiy-ad
stupid$_{TK}$-DER$_{AA}$-PL$_{AA}$
'stupidities'

b. cf. *salaq$_{TK}$-tiye$_{AA}$-ler$_{TK}$

c. *salaq-tiye-den
stupid$_{TK}$-DER$_{AA}$-ABL$_{TK}$
'from the stupidity'

d. cf. mı salaq-tiye
from$_{AA}$ stupid$_{TK}$-DER$_{AA}$
'from the stupidity'

Language-mixing patterns of trilingual speakers are crucial in demonstrating that the above restriction cannot be correct.[34] In all three examples in (59), the root *xatan* 'son-in-law' comes from Arabic. Yet, the examples differ regarding the nature of the ensuing morphemes. While in (59a) and (59b), the plural morpheme is from Arabic just like the root, and the vocative case marker *-no* that follows the plural is from NK in (59a), whereas the possessive marker *-Im* in (59b) is from Turkish. The example in (59c) shows that it is even possible to have three morphemes each belonging to a different language. The root is from AA, the plural is from Turkish, and the vocative is from NK.

[34] As such, some of the following examples are not found in the grammars of bilinguals.

(59) a. xatan-ad-no
son-in-law$_{AA}$-PL$_{AA}$-VOC$_{NK}$
'Sons-in-law!'

b. xatan-ad-ım
son-in-law$_{AA}$-PL$_{AA}$-1SG.POSS$_{TK}$
'my sons-in-law'

c. xatan-lar-no
son-in-law$_{AA}$-PL$_{TK}$-VOC$_{NK}$
'Sons-in-law!'

Examples in (60) illustrate the same pattern: The possibility of (60a) which involves the outer-most possessor morpheme demonstrates that the grammatical examples are not limited to those involving the vocative suffix.[35] (60d) confirms that when the root is NK as in *piroz* 'blessed one', then it is ruled out.

(60) a. $^{(?)}$genc-in-ān me
young.one$_{TK}$-IRR.PL$_{AA}$-EZ.PL$_{NK}$ our.OBL$_{NK}$
'our young ones'[36]

b. ode-yán-i
room$_{TK}$-PL$_{NK}$-1POSS$_{AA}$
'my rooms'

c. %hamar-lık-án
donkey$_{AA}$-DER$_{TK}$-PL$_{NK}$
'stupidities'[37]

[35]Thanks to an anonymous reviewer for suggesting to test such examples.

[36]Interestingly, the minimal pair in (i) is disallowed although it is identical in terms of the ordering of morphemes and their origins. I take this as further support for the irregular nature of the *-in* plural vs the regular *-ad* plural in Arabic. Accordingly, only the latter is incompatible with the Ezafe - a linker morpheme that introduces dependents of the noun including attributive adjectives, possessors in Iranian languages. Crucially the Ezafe marker varies in ϕ-features (gender and number). The incompatibility between the regular plural *-ad* and the Ezafe presumably is because they occupy the same position.

(i) *soba-d-ān me
stove$_{TK}$-PL$_{AA}$-EZ.PL$_{NK}$ our.OBL$_{NK}$
'our stoves'

[37]This example was not accepted by one of the trilingual consultants and some of the speakers they consulted. At this point, I must leave it as open question what constrains such asymmetries that may hold in some circumstances, but not others. A similar point is made in Kerslake 1998 for Middle Ottoman

d. *piroz-in-ān me
blessed.one$_{NK}$-IRR.PL$_{AA}$–EZ.PL$_{NK}$ our.OBL$_{NK}$
'our blessed ones'

Examples in (61) further show that not all instances containing the vocative suffix are grammatical. As seen in (61c), it is ungrammatical to have the vocative morpheme as the outermost morpheme when a plural morpheme from AA intervenes.

(61) a. heval-no
friend$_{NK}$-VOC$_{NK}$
'(hey) friends!'

b. heval-ad
friend$_{NK}$-PL$_{AA}$
'friends'

c. *heval-ad-no
friend$_{NK}$-PL$_{AA}$-VOC$_{NK}$
'(hey) friends!'

It is also possible to find acceptable forms with a longer sequence of morphemes. These are exemplified in (62).

(62) a. xatan-lar-ım-no
son-in-law$_{AA}$-PL$_{TK}$-1SG.POSS$_{TK}$-VOC$_{NK}$
'My sons-in-law!'

b. xatan-ad-ım-no
son-in-law$_{AA}$-PL$_{AA}$-1SG.POSS$_{TK}$-VOC$_{NK}$
'My sons-in-law!'

Word-internal trilingual language mixing can be found in other languages as well, although very rarely. Still, the reported example appears to conform to the generalizations

Turkish, which was developed by an 'effectively trilingual elite' (Kerslake 1998:180) trained in Turkish, Persian, and Arabic. Middle Ottoman Turkish incorporated numerous lexical items alongside fully productive morpho-syntactic structures from both Persian and Arabic which were used alongside pre-existing Turkish equivalents. In spite of the trilingualism of Ottoman authors, the division of labour between Perso-Arabic and Turkish material is highly constrained, and not all expected patterns were available.

A similar restriction was already reported earlier in the context of the light verbs, where we observed that systematically the light verb comes from language A, while the nonverbal counterpart from language B. Crucially, this holds even in the languages that are the focus of this study.

that we have observed so far from AA-NK-TK. For example, in the Corfiot example in (63) the root comes from Hebrew, the verbalizer from Greek and the infinitival from Romance.

(63) (Corfiot; Vardakis 2023:6)[38]

dibur-efs-ár
$\text{talk}_{\text{Hebrew}}$-$\text{VRBZ}_{\text{Greek}}$-$\text{INF}_{\text{Romance}}$

'talk'[39]

Overall, we can summarize the attested and unattested patterns as in (64). In a contact situation involving two or three languages, code-switching allows various patterns of morpheme insertion in (64a) through (64c), but not the pattern in (64d).[40] For example,

[38]Lefteris Paparounas, p.c., notes that the morpheme *efs-* is most likely the combination of the modern Greek verbalizer *-ev-* and the perfective aspect morpheme *-s-* (along with some phonological processes). If that is the case, then the Corfiot example would be closer to the sequence in (63).

[39]Another potential example comes from Ottoman Turkish, (i), where the root and the irregular broken plural come from Arabic, then the Ezafe and the modifier come from Persian, and the case marking comes from Turkish. This follows from the syntax of Ezafe in Toosarvandani and Van Urk 2014; Akkuş et al. 2023.

(i) [küffâr-ı hâksâr]-ı ...
[unbeliever.$\text{PL}_{\text{Arabic}}$-$\text{EZ}_{\text{Persian}}$ $\text{base}_{\text{Persian}}$]-$\text{ACC}_{\text{Turkish}}$ ...
'the base unbelievers ...' (Middle Ottoman Turkish; Woodhead 1983:138)

However, an anonymous reviewer brings to my attention a couple of examples which seem to not obey the generalization, e.g., (ii).

(ii) [sergi-i umûmi]-nin küşad-ın-dan evvel
[$\text{exhibition}_{\text{Turkish}}$-$\text{EZ}_{\text{Persian}}$ $\text{public}_{\text{Arabic}}$]-$\text{GEN}_{\text{Turkish}}$ opening-3SG.POSS-ABL before
'before the opening of the World's Fair' (Ahmet Mithat 1889, *Avrupa'da bir Cevelan*, (473))

Various possibilities seem conceivable regarding how to analyze (ii). One option is that Ottoman Turkish is even more permissive than the No-Reversal Constraint would suggest. Alternatively, other factors, some of which are kindly raised by the reviewer, could be at work. For instance, the first term in the Ezafe may be hierarchically adjacent to the Turkish suffix, so the apparent violation may simply be a matter of the surface string. Moreover, the reviewer notes that as a general convention, the only elements allowed in an Ezafe construction in Ottoman Turkish are Persian and Arabic, like (i). Therefore, (ii) and the other example *ordu-yu hümâyûn* 'Imperial Army' are 'rare exceptions', possibly due to them being formulaic official titles, as the reviewer acknowledges. If one chooses this option, then examples like (i) would still be valid. Yet another option is to take it that following the suggestion in fn. 44, the Ezafe construction itself demarcates a domain, as such the No-Reversal constraint is to be obeyed inside this domain. Thus, the relevant domain would be the content within the square brackets in (i) and (ii). Without an in-depth investigation and understanding of Ottoman Turkish, I have to put aside which of these options is to be chosen.

The reviewer also notes that in Ottoman, the frame language was always Turkish. This differs from the word-internal language mixing patterns in this section where there is more flexibility. However, it is more like the pattern in light verb constructions, where the light verb has to come from a specific language. See fn. 45 for a similar issue between Spanish-German vs Greek-German language-mixing. It is yet another aspect of language-mixing that deserves further investigation as to why languages or constructions are more flexible. The No-Reversal Constraint has no insight on this aspect.

[40]The pattern *Lang A* $\prec$ *Lang A* $\prec$ *Lang A* is not included since it involves no code-switching.

(64a) indicates that a morpheme (*Morph*) from Language A can be followed by a morpheme from language B, which in turn is followed by another morpheme from Language B. (64c) shows that each morpheme can belong to a different language, while in the ungrammatical (64d), the first morpheme is from language A, followed by the second morpheme from language B. Crucially, the third morpheme is drawn back from the language A.

(64) **Morph1** **Morph2** **Morph3**

a. Lang A ≺ Lang B ≺ Lang B

b. Lang A ≺ Lang A ≺ Lang B

c. Lang A ≺ Lang B ≺ Lang C

d. *Lang A ≺ Lang B ≺ Lang A

In order to capture the attested patterns and rule out the unavailable pattern, I propose a constraint that prohibits switching *back* to a language that has already contributed a morpheme (via Vocabulary Insertion) earlier in the derivation, called *No-Reversal Constraint* (the constraint is also compatible with Aboh's (2015) *Recombination* approach). This constraint is at work both for the bilingual and trilingual speakers. This constraint is also compatible with examples that contain more than three morphemes in a single morphological word, as in (63). Those examples also follow from the No-Reversal Constraint.[41]

The comparison of the No-Reversal Constraint with a few prominent alternative proposals reveals that prior approaches fail to capture the full range of the attested patterns.

The well-known study by Poplack (1980, 1981) argues that code-switching could only target two free morphemes, and it could not apply between morphemes within the same word. This approach effectively imposes a prohibition on code-switching in head movement and word-internal contexts (Poplack 1981). I do not dwell on this proposal since studies have shown that this statement is too strong, and such patterns indeed are attested. This paper itself provides further examples to that effect, which shows that language-mixing is not restricted to free morphemes.

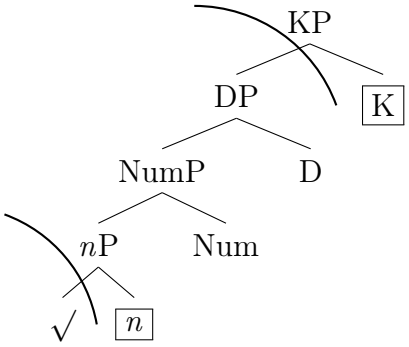
Another influential analysis comes from Belazi et al. (1994), who proposes the Functional Head Constraint (FHC), which distinguishes between lexical and functional categories. According to the FHC, a code-switch may not occur between a functional head and its complement, while code-switching between a lexical head and its complement proceeds

[41] On the surface of it, the patterns in (64) resemble the the so-called *ABA pattern (Bobaljik 2012), which effectively blocks syncretism between two non-adjacent cells in morphological paradigms. However, the containment relation assumed for such patterns is not available in the language-contact examples.

unimpeded.[42] This proposal is also too restrictive, in that it would incorrectly rule out examples in (55) and (56), where the Root and the categorizer n are of different languages. Similarly, it would fail to capture the possibility of templatic borrowing in the verbal domain discussed in section 4.1. Those examples involve the root coming from Turkish, whereas the categorizer v and Voice from Arabic.[43] Perhaps more challenging are the examples in (59), which involves functional categories within the nominal phrase that belong to different languages.

A recent analysis by López et al. (2017) is built on the phase-theory and argues that code-switching is governed by phase-theoretic considerations. López et al. (2017) make the natural and insightful assumption that bilinguals have multiple externalization systems or PFs, and suggest that when a structure is transferred, it is transferred in one block to one of the PFs. From these, it would follow that "code-switching may take place at phase boundaries but not within the phase" (López et al. 2017:5). This is because code-switching within the phase would entail transferring some material to one externalization system while simultaneously transferring some other material to another externalization system. López et al. (2017:(7)) assume the standard structure in (65) for the nominal phrases (which also straightforwardly extends to Arabic/Semitic, Kurdish and Turkish, as evidenced also by the morpheme ordering), where the nominal phrase is headed by K (=case), with the following hypotheses: in a nominal phrase, a root is selected by the categorizer n, which is also a phase head. n is selected by Number, which is itself selected by D. D is the complement of K, another phase head in the nominal domain (see Citko 2014 for the same phasehood properties in the nominal domain).

(65)



[42] Abstracting away from the details, this analysis is built on the view that 'language' is a feature [uninterpretable, in today's terms] a functional head bears, and as such the functional head requires its complement to match its own corresponding 'language' feature. This feature-matching of 'language' is missing between a lexical head and its complement.

[43] As noted by an anonymous reviewer, given that Belazi et al.'s study predates the treatment of v and n as functional categories, it is always a possibility that they could treat them as lexical categories, as such not subject to the FHC. Acknowledging that possibility, which would still not be a standard move, their approach would still have no explanation for examples such as (59), (39).

According to this formulation, the complement of a phase head and the phase head itself are transferred in different phrases. In (65), phase heads are placed in a box, and the nodes that are within the same arc are transferred together. For example, n and its complement $\sqrt{}$ belong to different phases in Spell-Out. Moreover, D would be transferred with Num and n, while K is transferred with the higher phase that contains it, i.e., vP or pP.

The proposal by López et al. (2017) is able to capture a large set of facts, mainly those that involve a Root, and functional nodes that belong to the next phase. These include (45), (51), (53), (54), (55), (56), (57), (58). For example, in (57), repeated here as (66), the Root is Arabic, and is transferred once the categorizer n is merged. The n and Num heads are transferred together, as such they can be a different language than Arabic, as long as both are from the same language, which is Turkish in the grammatical (66a). On the other hand, in the unacceptable form in (66b), the n and D heads belong to distinct languages, which violates code-switching between two functional heads that belong in the same extended projection, and more specifically in the same phase domain.

(66) a. hamar-lık-lar
donkey$_{\text{AA}}$-NMZ$_{\text{TK}}$-PL$_{\text{TK}}$
'stupidities'

b. cf. *hamar$_{\text{AA}}$-lık$_{\text{TK}}$-ad$_{\text{AA}}$

Despite its appeal, the phase-based account by López et al. (2017) rules out some of the attested patterns in the language-mixing situation at hand, e.g., those in (52b), (59b), (60) and (62b). For example, in (59b), both the plural on Num head and the possessive morpheme on D/Poss would be within the same Spell-out domain, therefore should be transferred together. Therefore, they are predicted to be coming from the same language, which is not the case. Similarly, in (52b), the categorizer n head and the plural morpheme are within the same phase, thus code-switching between the two should be disallowed. This is also not correct. For these reasons, I take the No-Reversal Constraint to be a more suitable analysis for code-switching involving Turkish, Anatolian Arabic and Kurdish.[44]

[44] Although the phase-account does not demarcate the correct domains for the languages in question for word-internal language-mixing, an open question is the nature of the exact boundary for this domain. It is clear that No-Reversal Constraint is obviated when larger constituents beyond morphological words are taken into consideration. Thus a more exact formulation is needed to determine at what stage of a derivation it is deactivated, and becomes active again. For example, (i.b) involves combination of Spanish roots/lexical item plus Quechua inflection/grammatical properties (that are boldfaced).

(i) a. xatan-lar-ım ma-co.
son.in.law$_{\text{AA}}$-PL$_{\text{TK}}$-1POSS$_{\text{TK}}$ NEG$_{\text{AA}}$-came$_{\text{AA}}$
'My sons-in-law didn't come.'

This account makes no reference to the phase-theory, but whether code-switching leads to switching back to an earlier language.[45]

5 Conclusions

This study has investigated the patterns of verbal and nominal formations in language mixing contexts of Turkish, Anatolian Arabic and Northern Kurdish. The examination has revealed various properties of theoretical interest. It shows that head-directionality may change as a result of level of contact. In particular, a change from head-initial to head-final pattern is an ongoing process, whereby *light verb + NV* > *NV + light verb* is attested. Moreover, certain functional categories can be borrowed as semantically vacuous heads, and are identical to their bare counterparts (cf. Marantz 2013; Anagnostopoulou and Samioti 2014). Such semantically empty heads are ignored for meaning in some dialects of Anatolian Arabic.

Furthermore, informed by trilingual language-mixing contexts, the study demonstrates that various formal approaches to code-switching, which rely on either a distinction between functional vs lexical categories or phasehood as the defining constraint on code-switching Belazi et al. 1994; Poplack 1981; López et al. 2017), are too restrictive for the languages in question. Instead, a restriction called *No-Reversal Constraint* is proposed, which places

b. (Media Lengua, Muysken 1997)

unu fabur-**ta** *pidi*-**nga-bu** *bini*-**xu-ni**.
one favor-ACC ask-NOM-BEN come-PROG-1SG

'I come to ask a favor.'

On a related note, Enoch Aboh, p.c., notes that the No-Reversal Constraint could be attributed to an effect of a ban against countercyclicity, which prevents looking further inside the derivation.

[45]The No-Reversal Constraint might turn out to be too weak when more crosslinguistic patterns are investigated (in that it might over-generate in other languages). However, I believe it is more straightforward to find ways to restrict this constrain, than to make the other constraints more flexible.

Note that a similar crosslinguistic issue arises in López et al. (2017). The authors note that a Spanish inflectional morpheme has been added to the German word *stuhl* in (i), which leads to ungrammaticality.

(i) *auf einem $\text{Stuhl}_{\text{Ger}}$-$o_{\text{Spa}}$
on a chair.M
(López et al. 2017:(32c))

However, as they note (without an explanation), this kind of code-switching is attested in the case of Greek-German, as in (ii).

(ii) to $\text{matratz}_{\text{Ger}}$-$i_{\text{Greek}}$
the mattress-NEUT-ACC
(López et al. 2017:fn. 7-ii)

Therefore, it is likely that any theory would need to be flexible enough to allow language-particular (even community-particular) properties to play a role, as also hinted at in Balam et al. (2020).

a ban on reverting back to a previous language for morpheme insertion in code-switching. The study leaves aside exactly how this constraint could be further restricted to apply straightforwardly to other languages, with the hope that it will generate a fruitful area of future research.

6 Acknowledgments

Many thanks to Deniz Tat for organizing the *Formal Approaches to Contact in/with Turkish (FACT)* at the Arctic University of Norway in Tromsø, where I presented this study and received very valuable feedback from the participants. Thanks also to Beste Kamali, Jaklin Kornfilt, Ümit Atlamaz, Enoch Aboh, María del Carmen Parafita Couto for their suggestions, and three anonymous reviewers of *Linguistic Variation* whose detailed and constructive comments greatly improved the paper.

References

Aboh, Enoch Oladé. 2015. *The emergence of hybrid grammars: Language contact and change*. Cambridge University Press.

Akkuş, Faruk. 2014. The Functional Categories and Phrase Structure of Sason Arabic. *MA thesis, Boğaziçi University* .

Akkuş, Faruk. 2015. Light Verb Constructions in Turkish: A Case for DP Predication and Blocking. In *Proceedings of the 9th Workshop on Altaic Formal Linguistics (MITWPL 76)*, ed. Andrew Joseph and Esra Predolac, 133–145. Cambridge: MIT Press.

Akkuş, Faruk. 2016. The Arabic Dialect of Mutki-Sason Areas. In *Arabic Varieties: Far and Wide. Proceedings of the 11th Conference of AIDA*, 29–41. Editura Universității din București.

Akkuş, Faruk. 2020. Anatolian Arabic. In *Arabic and contact-induced change*, ed. Christopher Lucas and Stefano Manfredi, 135–158. Language Science Press.

Akkuş, Faruk, and Elabbas Benmamoun. 2016. Clause Structure in Contact Contexts: the Case of Sason Arabic. *Perspectives on Arabic Linguistics XXVIII* 153–172.

Akkuş, Faruk, and Elabbas Benmamoun. 2018. Syntactic Outcomes of Contact in Sason Arabic. In *Arabic in Contact*, ed. Manfredi Stefano, Mauro Tosco, and Giorgio Banti, 38–52. John Benjamins.

Akkuş, Faruk, David Embick, and Mohammed Salih. 2023. Case and the syntax of argument indexation: An analysis of Sorani Kurdish. *Ms., University of Massachusetts Amherst and University of Pennsylvania* URL https://lingbuzz.net/lingbuzz/007378.

Akkuş, Faruk, and Balkız Öztürk. 2017. On cognate objects in Sason Arabic. *University of Pennsylvania Working Papers in Linguistics* 23:2.

Alagozlu, Nuray. 2017. Code-switching between Kabardian and Turkish. *Paper presented at the 6th International Symposium on Bilingualism* URL http://www.isb6.org/static/proceedings/alagozlu.pdf.

Alexiadou, Artemis. 2017. Building verbs in language mixing varieties. *Zeitschrift für Sprachwissenschaft* 36:165–192.

Alexiadou, Artemis, Elena Anagnostopoulou, and Florian Schäfer. 2015a. *External arguments in transitivity alternations: A layering approach*, volume 55. Oxford University Press.

Alexiadou, Artemis, Terje Lohndal, Tor A Åfarli, and Maren Berg Grimstad. 2015b. Language mixing: a distributed morphology approach. In *Proceedings of NELS 45*, ed. Thuy Bui and Deniz Özyıldız, volume 45, 25–38. Amherst, MA: GLSA.

Anagnostopoulou, Elena, and Yota Samioti. 2014. Domains within words and their meanings: A case study. In *The syntax of roots and the roots of syntax*, ed. Artemis Alexiadou, Hagit Borer, and Florian Schäfer, 81–111. Oxford University Press.

Arad, Maya. 2003. Locality constraints on the interpretation of roots: The case of Hebrew denominal verbs. *Natural Language & Linguistic Theory* 21:737–778.

Arad, Maya. 2005. *Roots and patterns: Hebrew morpho-syntax*. Springer.

Arkadiev, Peter, and Kirill Kozhanov. 2021. Borrowing of Morphology. In *The Wiley Blackwell Companion to Morphology*, ed. Peter Ackema, Sabrina Bendjaballah, Eulalia Bonet, and Antonio Fábregas. Blackwell.

Armostis, Spyros, and Marilena Karyolemou. 2023. Contact-Induced Change in an Endangered Language: The Case of Cypriot Arabic. *Languages* 8:10.

Bağrıaçık, Metin, Angela Ralli, and Dimitra Melissaropoulou. 2015. Borrowing verbs from Oghuz Turkic: two linguistic areas. In *Borrowed morphology*, ed. Francesco Gardani, Peter Arkadiev, and Nino Amiridze, 109–136. Mouton de Gruyter.

Balam, Osmer, María del Carmen Parafita Couto, and Hans Stadthagen-González. 2020. Bilingual verbs in three spanish/english code-switching communities. *International Journal of Bilingualism* 24:952–967.

Bandi-Rao, Shoba, and Marcel den Dikken. 2014. Light switches. In *Grammatical theory and bilingual codeswitching*, ed. Jeff MacSwan, 161–184. Cambridge, MA: MIT Press.

Belazi, Hedi M, Edward J Rubin, and Almeida Jacqueline Toribio. 1994. Code switching and X-bar theory: The functional head constraint. *Linguistic Inquiry* 221–237.

Benmamoun, Elabbas. 2000. *The feature structure of functional categories: A comparative study of Arabic dialects*. Oxford University Press.

Biţună, Gabriel. 2016. Morfo-Sintaxa Dialectului Arab Nord-Mesopotamian Din Siirt, Turcia. Doctoral Dissertation, Universitatea din Bucureşti.

Bobaljik, Jonathan David. 2012. *Universals in comparative morphology: Suppletion, superlatives, and the structure of words*. MIT Press.

Bulut, Christiane. 2006. Turkish elements in spoken Kurmanji. In *Turkic languages in contact*, ed. Hendrik Boeschoten and Lars Johanson, 95–121. Berlin: Harrassowitz Verlag.

Çabuk, Sakine. 2019. A note on the contact between Kurmanji Kurdish and Turkish at lexical and morphological level. *International Journal of Bilingualism* 23:861–864.

Chung, Sandra, and William Ladusaw. 2003. *Restriction and saturation*. MIT Press.

Citko, Barbara. 2014. *Phase theory: An introduction*. Cambridge University Press.

Coetsem, Frans van. 1988. *Loan Phonology and the Two Transfer Types in Language Contact*. Foris.

Coetsem, Frans van. 2000. *A General and Unified Theory of the Transmission Process in Language Contact*. Winter.

Coghill, Eleanor. 2015. Borrowing of verbal derivational morphology between Semitic languages: the case of Arabic verb derivations in Neo-Aramaic. In *Borrowed morphology*, ed. Francesco Gardani, Peter Arkadiev, and Nino Amiridze, 83–108. Mouton de Gruyter.

Dabir-Moghaddam, Mohammad. 1997. Compound verbs in Persian. *Studies in the Linguistic Sciences* 27:25–59.

Dorleijn, Margreet. 1996. *The decay of ergativity in Kurdish*. Tilburg University Press.

Edwards, Malcolm, and Penelope Gardner-Chloros. 2007. Compound verbs in codeswitching: Bilinguals making do? *International Journal of Bilingualism* 11:73–91.

Embick, David. 2004. On the structure of resultative participles in English. *Linguistic Inquiry* 35:355–392.

Embick, David. 2010. *Localism versus globalism in morphology and phonology*, volume 60. MIT Press.

Embick, David, and Alec Marantz. 2008. Architecture and blocking. *Linguistic Inquiry* 39:1–53.

Folli, Raffaella, Heidi Harley, and Simin Karimi. 2005. Determinants of event type in Persian complex predicates. *Lingua* 115:1365–1401.

Göksel, Aslı, and Celia Kerslake. 2005. *Turkish: A comprehensive grammar*. Routledge.

González-Vilbazo, Kay, and Luis López. 2011. Some properties of light verbs in code-switching. *Lingua* 121:832–850.

Grigore, George. 2007a. *L'arabe parlé à Mardin – monographie d'un parler arabe périphérique*. Editura Universităţii din Bucureşti.

Grigore, George. 2007b. L'énoncé non verbal dans l'arabe parlé à Mardin. *Romano-Arabica* 51–63.

Grimstad, Maren Berg, and Tor A Åfarli. 2014. Language mixing and exoskeletal theory: A case study of word-internal mixing in American Norwegian. *Nordlyd* 41:213–217.

Gündoğdu, Songül. 2015. Noun-Verb Complex Predicates In Kurmanji Kurdish. *Proceedings of ConSOLE XXIII* 279–301.

Gündoğdu, Songül. 2018. Argument-Adjunct distinction in Kurmanji Kurdish. Doctoral Dissertation, Boğaziçi University. Istanbul.

Gündoğdu, Songül, and Faruk Akkuş. 2022. Syntax of causatives in language contact: Insights from Kurdish and Arabic. In *Poster presented at LSA Annual Meeting 2022*.

Gürer, Aslı. 2014. Adjectival participles in Turkish. *Lingua* 149:166–187.

Haig, Geoffrey. 2007. Grammatical borrowing in Kurdish (Northern group). In *Grammatical borrowing and cross-linguistic perspective*, ed. Yaron Matras and J. Sekkal, 165–184. Paris, France: Mouton de Gruyter.

Haig, Geoffrey, and Ergin Öpengin. 2014. Kurdish: A critical research overview. *Kurdish Studies* 2:99–122.

Haig, Geofrey. 2014. East Anatolia as a linguistic area? Conceptual and empirical issues. In *Bamberger Orientstudien*, ed. Lale Behzadi, Patrick Franke, Geofrey Haig, Christoph Herzog, Birgitt Hofmann, Lorenz Korn, and Susanne Talabardon, 13–36. University of Bamberg Press.

Heine, Bernd, and Motoki Nomachi. 2013. Contact-induced replication: Some diagnostics. In *Shared Grammaticalization: With special focus on the Transeurasian languages*, ed. Martine Robbeets and Hubert Cuyckens, 67–100. John Benjamins.

Jastrow, Otto. 2006. Anatolian Arabic. *Encyclopedia of Arabic Language and Linguistics* 1:87–96.

Jastrow, Otto. 2011. Turkish and Kurdish influences in the Arabic Dialects of Anatolia. *Türk Dilleri Araştırmaları* 21:83–94.

Johanson, Lars. 2002. *Structural Factors in Turkic Language Contacts*. Routledge.

Karimi, Simin. 1997. Persian complex verbs: Idiomatic or compositional. *Lexicology-Berlin* 3:273–318.

Kerslake, Celia. 1998. Ottoman Turkish. In *The Turkic Languages*, ed. Lars Johanson and Éva Á. Csató, 174–194. Routledge.

Key, Gregory. 2013. The morphosyntax of the Turkish causative construction. Doctoral Dissertation, University of Arizona.

Key, Gregory. 2021. Passive in Turkish: Re-thinking *u*-syncretism. *Manuscript* .

Khan, Geoffrey. 2007. North Eastern Neo-Aramaic. In *Grammatical borrowing and cross-linguistic perspective*, ed. Yaron Matras and J. Sekkal, 197–214. Paris, France: Mouton de Gruyter.

Kornfilt, Jaklin. 1997. *Turkish*. London/New York: Routledge.

Kramer, Ruth. 2016. A split analysis of plurality: Number in Amharic. *Linguistic Inquiry* 47:527–559.

Lahdo, Ablahad. 2009. *The Arabic Dialect of Tillo in the Region of Siirt (south-eastern Turkey)*. Acta Universitatis Upsaliensis.

López, Luis, Artemis Alexiadou, and Tonjes Veenstra. 2017. Code-switching by phase. *Languages* 2:9.

Marantz, Alec. 2013. Locality domains for contextual allomorphy across the interfaces. In *Distributed Morphology Today: Morphemes for Morris Halle*, ed. Ora Matushansky and Alec Marantz, 95–115. MIT Press.

Matras, Yaron. 2009. *Language contact*. Cambridge University Press.

McCarthy, John. 1979. Formal problems in semitic phonology and morphology. Doctoral Dissertation, MIT.

McCarthy, John. 1981. A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry* 12:373–418.

Melissaropoulou, Dimitra. 2016. Variation in word formation in situations of language contact: The case of Cappadocian Greek. *Language Sciences* 55:55–67.

Muysken, Pieter. 1997. Media Lengua. In *Contact languages: A wider perspective*, ed. Sarah Grey Thomason, 365–426. John Benjamins.

Muysken, Pieter. 2000. *Bilingual speech: A typology of code-mixing*. Cambridge University Press.

Myers-Scotton, Carol. 2002. *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford University Press.

Myers-Scotton, Carol, and Janice Jake. 2009. A universal model of code-switching and bilingual language processing and production. In *The Cambridge Handbook of Linguistic Code-switching*, ed. B. Bullock and J.A Toribio, 336–357. Cambridge University Press.

Myler, Neil. 2016. *Building and interpreting possession sentences*. MIT Press.

Öpengin, Ergin. 2012. Sociolinguistic situation of Kurdish in Turkey: Sociopolitical factors and language use patterns. *International Journal of the Sociology of Language* 151–180.

Öztürk, Balkız. 2005. *Case, referentiality and phrase structure*. Amsterdam/Philadelphia: John Benjamins.

Paparounas, Lefteris, and Faruk Akkuş. 2023. Anaphora and agreement in the Turkish DP: Delimiting binding-through-Agree. *Natural Language & Linguistic Theory* .

Poplack, Shana. 1980. Sometimes I'll start a sentence in Spanish *y termino en español*: toward a typology of code-switching. *Linguistics* 18:581–618.

Poplack, Shana. 1981. Syntactic structure and social function of code-switching. In *Latino language and communicative behavior*, ed. R. P. Durán, 169–184. Norwood, NJ: Ablex.

Ramchand, Gillian. 2008. *Verb meaning and the lexicon: A first-phase syntax*. Cambridge: Cambridge University Press.

Sankoff, David, and Shana Poplack. 1981. A formal grammar for code-switching. *Paper in Linguistics* 14:3–45.

Talay, Shabo. 2007. The influence of Turkish, Kurdish and other neighbouring languages on Anatolian Arabic. *Romano-Arabica* 179–189.

Thomason, Sarah G. 2001. *Language Contact. An Introduction*. Edinburgh University Press.

Toosarvandani, Maziar, and Coppe Van Urk. 2014. Agreement in Zazaki and the nature of nominal concord. *Ms., UC-Santa Cruz and MIT* .

Tracy, Rosemarie. 2000. Language mixing as a challenge for linguistics. In *Crosslinguistic structures in simultaneous bilingualism*, ed. Susanne Döpke, 11–36. Amsterdam: Benjamins.

Trépos, Pierre. 1957. *Le pluriel breton*. Brest: Emgleo Breiz.

Tucker, Matthew A. 2011. The Morphosyntax of the Arabic Verb: Toward a Unified Syntax-Prosody. In *Morphology at Santa Cruz: Papers in Honor of Jorge Hankamer*. URL https://escholarship.org/uc/item/0wx0s7qw.

Vardakis, Georgios. 2023. Language Contact in the Written Sources of the Corfiot Jews. In *Talk Delivered at the 43th Annual Meeting of the Department of Linguistics, School of Philology, Aristotle University of Thessaloniki*.

Vaux, Bert. 2001. Hemshinli: The Forgotten Black Sea Armenians. *Journal of Armenian Studies* 6:47–71.

Versteegh, Kees. 1997. *The Arabic Language*. New York: Columbia University Press.

Wittrich, Michaela. 2001. *Der arabische Dialekt von Āzex*, volume 25. Otto Harrassowitz Verlag.

Wohlgemuth, Jan. 2009. *A typology of verbal borrowings*. De Gruyter Mouton.

Wood, Jim. 2015. *Icelandic morphosyntax and argument structure*, volume 90. Springer.

Wood, Jim. 2016. How roots do and don't constrain the interpretation of voice. *Working Papers in Scandinavian Syntax* 96:1.

Woodhead, Christine. 1983. *Talikizade's Şehname-i Hümayun: A history of the Ottoman campaign into Hungary 1593-1594*. Klaus Schwarz Verlag.

Zabbal, Youri. 2002. The semantics of number in the Arabic number phrase. *Master's thesis, University of Alberta* .

Contact info:

fakkus@umass.edu

Department of Linguistics, University of Massachusetts Amherst, Integrative Learning Center, 650 North Pleasant Street, Amherst, MA 01003, USA