

Locating code-switching in the grammar: Role of M-words and hierarchy*

Faruk Akkuş · fakkus@umass.edu
University of Massachusetts Amherst

NELS 55
Yale University
October 17-18, 2024

1 Introduction

The long term bi- and multi-lingualism in southeastern Turkey has yielded patterns of intra-word code-switching (CS) among Turkish (TK), Anatolian Arabic (AA), and Northern Kurdish (NK).

I investigate this process in the verbal and the nominal domain, informed by the very rarely-studied *trilingual* language mixing.

- **CS in the verbal domain** In (1), the Turkish lexeme *kapat* ‘close’ is used in Anatolian Arabic, by first identifying and extracting a tri- or quadriradical root $\sqrt{\text{KPT}}$. Then the root is assigned to one of the patterns/derivations.

- (1) Stem II *qappat* – *iqappət* ‘to close’ cf. Turkish *kapat*-
Stem III *ḍāyan* – *iḍāyən* ‘to be patient, to bear up’ cf. Turkish *dayan*- (Talay 2007)

- **CS in the nominal domain** (2) is an initial illustration.¹

- e.g., in (2a), the root comes from AA, the plural from TK, and the case marker from NK.
- not all permutations are licit: (2b) illustrates one unacceptable combination of morphemes.

- (2) a. *xatan-lar-no*
son-in-law_{AA}-PL_{TK}-VOC_{NK}
‘Sons-in-law!’
b. cf. **xatan_{AA}-lar_{TK}-i_{AA}* ‘my sons-in-law’

Generalization: The investigation of trilingual data reveals the patterns in (3):

- (3) **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

*Many thanks to David Embick, Lefteris Paparounas, Enoch Aboh, María del Carmen Parafita Couto, Cem Keskin, the reviewers of NELS 55 for their valuable feedback. The study wouldn’t have been possible without the generosity of my language consultants.

¹I use the *subscripting* notation in the examples, when necessary, to indicate which language a word/lexeme or morpheme comes from, e.g., the subscript _{AA} means the morpheme is from Anatolian Arabic. Some researchers also use the term *language mixing* for CS (cf. Tracy 2000, Alexiadou et al. 2015, Alexiadou 2017).

- (3a) indicates that a *Morph(eme)* from language A can be followed by a morpheme from language B, which in turn is followed by another morpheme from language B.
- In (3b) the first two morphemes are from language A, followed by a morpheme from language B.
- (3c) shows that each morpheme can belong to a different language.
- while in the ungrammatical (3d), 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.

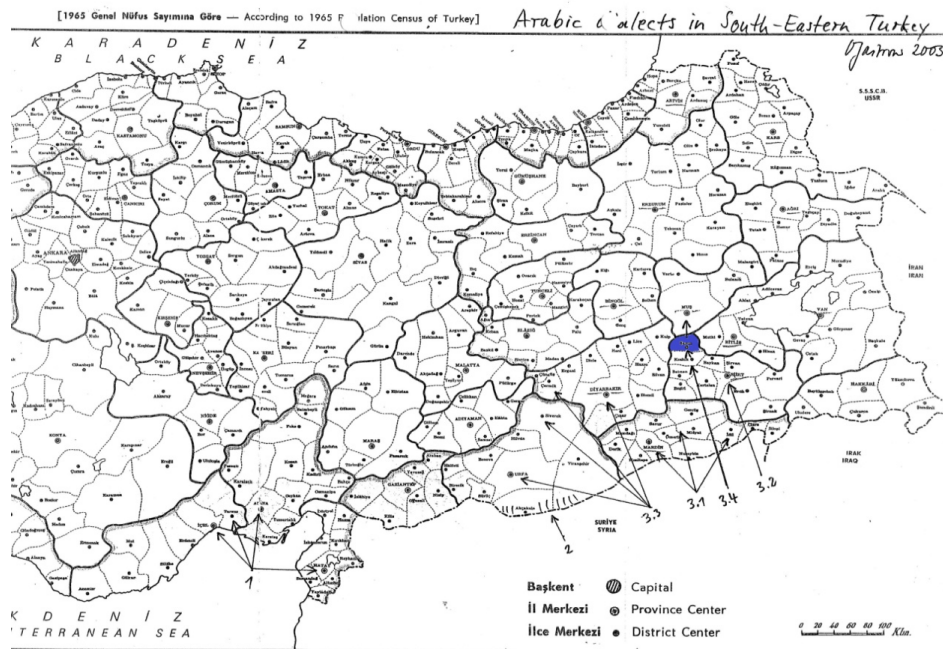
RESULTS PREVIEWED

- The intra-word CS patterns in (3) call for a different explanation than those of prior approaches:
 - e.g., a distinction between functional vs lexical categories (e.g., Belazi, Rubin, and Toribio 1994) or *phasehood* as the defining constraint on CS (e.g., López, Alexiadou, and Veenstra 2017) [See Appendix §5].
- In Akkuş 2024, I called the pattern in (3) the *No-Reversal Constraint (NRC)*,
 - whose governing restriction is that CS does not allow a switch back to a language that has already been externalized earlier in the derivation.

Today's Goal: exploration of the nature of the NRC and where in the grammar it applies.

- This constraint applies to subwords in Morphology prior to (or concomitant with) linearization of morphosyntactic words (MWds), and before Vocabulary Insertion (Embick and Noyer 2001).
 - Moreover, the NRC exhibits a bottom-up effect (i.e., no [[A B] A] structures), rather than simply applying over linearization of terminal nodes, which would impose a left-to-right effect (i.e. no A-B-A strings).

Sason Arabic is an endangered Arabic variety spoken in south-eastern Turkey, partly shown in the map (Jastrow 1978; Akkuş 2017). (Peripheral Arabic > Mesopotamian *qeltu*-Arabic > Anatolian Arabic > Kozluk-Sason-Muş).



2 The (non)-templatic language mixing

- Patterns of language mixing that involve affixal (concatenative) languages are commonly reported,
 - e.g., in (4), the root $\sqrt{\text{UTILISI\text{ER}}}$ ‘use’ comes from Spanish, whereas the (dissociated) morpheme corresponding to 1pl *-en* comes from German.
- (4) Wir *utilisier-en* spanische Wörter, die dann *alemanisiert* werden y hacen klingen un
We use_{Spanish-PLGerman} Spanish words that then Germanized are and do sound a
poco raro.
bit strange
‘We use Spanish words, that are then Germanized and sound a bit strange.’ (González-Vilbazo and López 2011:(2))

What about cases that involve interactions of concatenative and non-concatenative languages?

- For example, what happens when an element from a concatenative language is integrated into a non-concatenative language such as Semitic?
- A common strategy is that such elements are incorporated into the verbal and nominal vocalic patterns of the recipient Semitic language, thus the *templatic pattern*.

2.1 The templatic pattern in the verbal domain

- 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 by the majority formed according to the II. or III. verbal stem, (5):
- (5) (Anatolian Arabic, Āzəx variety, Talay 2007)
- | | | |
|----------|----------------------------------------------------------|-------------------------------|
| Stem II | <i>qappat</i> – <i>iqappət</i> ‘to close’ | cf. Turkish <i>kapat</i> - |
| Stem II | <i>qayyad</i> – <i>iqayyəd</i> ‘to register’ | cf. Turkish <i>kayıt et</i> - |
| Stem III | <i>ḍāyan</i> – <i>iḍāyən</i> ‘to be patient, to bear up’ | cf. Turkish <i>dayan</i> - |

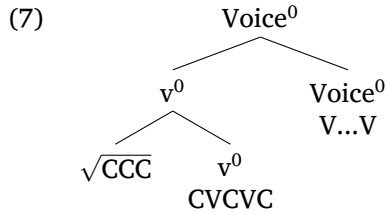
Similar issues arise in borrowings into other Semitic languages as well. For example, in the Telkepe dialect of Neo-Aramaic spoken in Detroit, the English word *charge* has been borrowed as *çrj* ‘to charge’ in Stem II.

- (6) Neo-Aramaic (Telkepe)
kə-mçarj-i-lə.
IND-charge.PRS-3PL-OBJ.3MS
‘They charge it up.’ (Coghill 2015:85,(1))

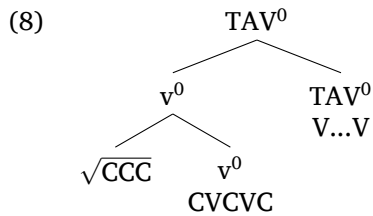
The syntax of verbal non-concatenation:

- According to Arad (2003, 2005), 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).²
- Successive-cyclic head movement of the verb up through this structure yields the following head structure at Spell-Out, (7):

²I 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; Tucker and Kastner 2019 and references therein).



- McCarthy (1979, 1981) and Tucker (2011) assume that in Arabic vocalic melody expresses tense, aspect and voice (thus TAV head), (8), presumably via a PF-Fusion rule for Arabic.³



Against this background, let's turn to language mixing:

- Consider the Turkish verbal form *karış* 'to meddle, mix', also incorporated as Stem III in (Sason) Arabic.
 - It is integrated into the phonology of the language, where velar stop becomes uvular, i.e., $k > q$. Additionally, it is used in different templates, inflected for person-number and different aspects.⁴

(9) *feminine; imperfective-perfective*

- a. kul çax **tı-qareş** fı şıylı-di
every time 3F-meddle.IPFV in work-my
'She always meddles in my work.'
- b. ams **qarş-e** fı şıylı-di
yesterday meddle.PFV-3F in work-my
'She yesterday meddled in my work.'

(10) *masculine; imperfective-perfective*

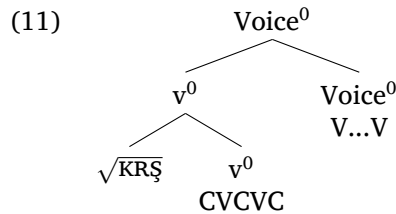
- a. kul çax **i-qareş** fı şıylı-di
every time 3M-meddle.IPFV in work-my
'He always meddles in my work.'
- b. ams **qaraş** fı şıylı-di
yesterday meddle.PFV.3M in work-my
'He yesterday meddled in my work.'

- To repeat, some Turkish/English verbs are borrowed into the Arabic/Neo-Aramaic system in a way that only contains the consonantal tier, which in turn is combined with the template.

³Whether this is true for all varieties of Arabic is debatable (see e.g., Benmamoun 2000 for Moroccan Arabic).

⁴The data for this study come from my own fieldwork mainly through acceptability judgments (as well as published sources) with six consultants. The speakers are able to produce consistent grammaticality judgments on sentences or phrases reported here, which reveal an underlying linguistic system. Fieldwork involved direct elicitation of (i) judgements about truth in particular contexts and (ii) judgements about felicity in particular contexts. All six consultants are from the Mutki region, Bitlis, Turkey. Three of the consultants are trilingual among a subvariety of Mutki-Sason dialects of Anatolian Arabic, Turkish and Northern Kurdish. Two other consultants 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 three trilingual speakers are 54, 43, and 60 years old, whereas the Arabic-Turkish bilinguals are 23, 29 and 33 years old. The consultants regularly use each of the languages in their lives though not necessarily in the same settings. 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.

- Thus consonants of the source loanword is retained; yet, the template and the vocalic melody (if any) come from the recipient language.
- Under the assumption that 'karış' is borrowed as a Root, this state of affairs is not surprising given that Turkish lacks the root-and-pattern morphology, 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, (11).
 - Thus the rest of the system functions as if a native trilateral root is merged, allowing bottom-up head movement (and other operations).



2.2 The templatic pattern in the nominal domain

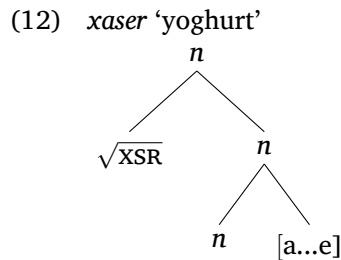
A similar templatic pattern is also attested in the nominal domain, esp. in plurals, which can be realized in particular templates known as broken plurals, as well as regular plurals realized as a suffix.

I first introduce the types of plurals and the structures associated with them in (Anatolian) Arabic, before proceeding with the discussion of observed modes of language mixing in the nominal domain.

2.2.1 Irregular, broken plurals

↪ I assume (with [Arad 2005](#) for Hebrew, and [Kramer 2016](#) 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 §2.1) and DM-approaches to declension classes) – though 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].
 - Prosodic constraints result in the interleaving of the vowels within the consonants.

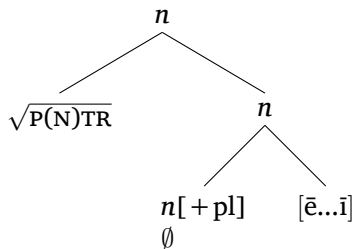


- Some (older) loans into the AA from Turkish have been integrated into the language, both phonologically and morphologically, paralleling the verbal domain.
 - for instance by acquiring their own broken plurals, e.g. (Mḥallami Arabic): *panṭur/panēṭir* 'trousers' (< Turkish *potur*) ([Wittrich 2001](#)).
 - This is the same pattern observed for certain native items that have broken plurals, e.g., *ṣinnor - ṣnēnir* 'cat(s)'.

- In the case of ‘trousers’, the root of the lexeme *po(n)tur* comes from Turkish, the *n* and the vocalic melody come from Arabic.⁵

- 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.

(13) *panētīr* ‘trousers’



The realization of the irregular plural is not limited to broken plurals; it can also be in the idiosyncratic suffix *-in*, (14).

(14)	<i>refiq</i> ‘friend’	<i>refq-in</i> ‘friends’ (Tillo Arabic)
	<i>angir</i> ‘friend’	<i>angir-in</i> ‘friends’
	<i>genc</i> ‘young one’ (cf. Turkish <i>genç</i>)	<i>genc-in</i> ‘young ones’
	<i>piroz</i> ‘blessed, holy’ (cf. Kurdish <i>pîroz</i>)	<i>piroz-in</i> ‘the blessed, holy (ones)’
	<i>delal</i> ‘dear, lovable’ (cf. Kurdish <i>delal</i>)	<i>delal-in</i> ‘lovable (pl.)’ ⁶

- For these instances, I will simply assume that we have an allomorph of *n[+ pl]* realized as *-in*, and no adjunction of vocalic melody.

- (14) also shows that the irregular plural can also attach to Turkish and Kurdish roots.

2.2.2 Regular and double plurals

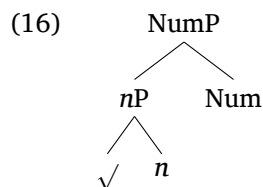
In addition to the broken (irregular) plurals, Arabic (including Anatolian Arabic varieties) also has the regular plural and double plurals.

- Many nouns are pluralized with the regular plural suffix *-ad*, *-at*.

(15)	<i>horti</i> ‘calf’	<i>horti-yad</i> ‘calves’
	<i>kartol</i> ‘potato’	<i>kartol-ad</i> ‘potatoes’
	<i>badıncan</i> ‘tomato’	<i>badıncan-ad</i> ‘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 (16).



⁵Armostis 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.

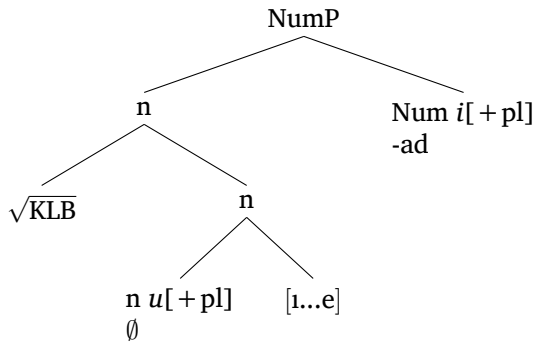
⁶In Tillo Arabic, this word is also found as *dalel-in* (Lahdo 2009:93). This suffix is attached to masculine nouns that denote humans only. Laks 2014 discusses plural formation of loanwords in Palestinian and Jordanian Arabic, and illustrates examples with broken plural and regular plurals.

- 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
 - 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 their various properties, including morpheme ordering, with the only possibility being $\sqrt{-Irreg-Reg}$.

- (17) a. *kelp* ‘dog’ *kileb* ‘dog.PL’ *kileb-ad* ‘dog.PL-PL’
 b. *genc* ‘young one’ *genc-in* ‘young one-PL’ *genc-in-ad* ‘young one-PL-PL’⁷

- Intuitively, we are dealing with two plural exponents, but only one interpretation of plurality.
- For the sake of exposition, I adopt Kramer’s (2016) analysis (see also Harbour 2011 for Kiowa, and Paparounas and Akkuş (2024) for the same analysis to forms like *biz-ler* ‘we-PL’ in Turkish), in which only one of the plural features is interpretable. As such in a double plural, the plural feature on *n* must be uninterpretable, (18).

(18) *kileb-ad* ‘dog.PL-PL’



Against this background, let us now turn to the patterns of word-internal CS in the nominal domain.

3 Word-internal language mixing and *No-Reversal Constraint*

Unlike most of the nominal examples thus far, in multilingual settings, the exponents of morphemes come from different languages. We see the same situation in the context of Turkish, Anatolian Arabic and Northern Kurdish.

3.1 Initial illustrations from Bilinguals

A commonly attested pattern involves addition of the Arabic regular plural morpheme *-ad* to Turkish roots.

- (19) a. *milyar-at* ‘billion_{TK}-PL_{AA}’ (Tillo Arabic, Lahdo 2009:92)
 b. *damad-ad* ‘son-in-law_{TK}-PL_{AA}’
 c. *soba-d* ‘stove_{TK}-PL_{AA}’

- The reverse pattern is also attested: it is possible to replace the Arabic plural morpheme on the Num head with its Turkish counterpart *-lar* in spontaneous speech, which also obeys the vowel harmony.⁸

⁷Another way to show that *-in* is an irregular plural morpheme (besides the morpheme order) is that it can never follow broken plurals unlike the regular *-ad*, thus **kileb-in*, which would be expected from a regular plural.

⁸Some speakers produce such forms in natural speech, but when asked about it, they do not like them.

– Note that this is attested in both regular, (20a), and double plural forms, (20b).

- (20) a. *xatan-lar* ‘son-in-law_{AA}-PL_{TK}’ cf. *xatan-ad*
 b. *kileb-ler* ‘dog_{AA}.PL_{AA}-PL_{TK}’ cf. *kileb-ad*

- This shows that the root (and maybe root + a higher functional head) might come from language A, whereas the top node, which is Num comes from language B.

Neo-Aramaic varieties illustrate other instances of CS in the nominal domain.

- Various Kurdish particles are incorporated into the morphology of the Jewish Aramaic dialects.
 - e.g., the Central Kurdish **definite article** suffix *-ake* (Khan and Mohammadirad 2024:104-105).

- (21) *kalba* ‘dog’ *kalb_{Aramaic}-ake_{Kurdish}* ‘the dog’

- The Kurdish **agent nominalizer** *-āna* is another affix used in the formation of nouns that are of Aramaic etymology. It is productively used to form active participles from the present stem of verbs:

- (22) Jewish Neo-Aramaic dialects (Khan and Mohammadirad 2024:100)
- | | | |
|------------------------------|---|-------------------------|
| <i>'axlāna</i> ‘(big) eater’ | < | <i>'x-l</i> ‘to eat’ |
| <i>qaryāna</i> ‘reader’ | < | <i>q-r-y</i> ‘to read’ |
| <i>yalpāna</i> ‘learner’ | < | <i>y-l-p</i> ‘to learn’ |

Potential restrictions:

Attempts to add outer morphemes are informative wrt potential restrictions that may be at work.

- Adding the possessive marker to the form in (19) is allowed only if the possessive morpheme is Arabic, as in (23).

- (23) a. *soba-d-i*
 stove_{TK}-PL_{AA}-1SG.POSS_{AA}
 ‘my stoves’
 b. cf. **soba_{TK}-d_{AA}-im_{TK}*

- The same holds for the examples in (20), involving the Turkish plural suffix as the outer-most morpheme. Any further morpheme can only be drawn from the Turkish inventory, and not Arabic, (24).⁹

- (24) a. *xatan-lar-ım*
 son-in-law_{AA}-PL_{TK}-1SG.POSS_{TK}
 ‘my sons-in-law’
 b. cf. **xatan_{AA}-lar_{TK}-i_{AA}*

A possible hypothesis (H1): Based on (23) and (24), one plausible explanation is that whichever language the least embedded head is from, any additional morpheme has to be from that language.

This hypothesis would also correctly capture patterns that involve ‘derivational’ suffixes.

⁹Another example involving multiple Turkish morphemes in an intra-word CS is below:

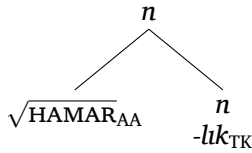
- (i) *forelesning-ler-e*
 lecture_{Norwegian}-PL_{TK}-DAT_{TK}
 ‘to the lectures’ (Jake et al. 2002:4)

- For example, 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.

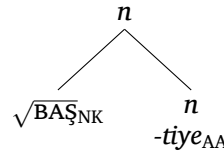
- (25) a. *salak* ‘stupid’ (Turkish root) + *-ti(y)ye* ‘state of’ (Arabic suffix) = *salaq-tiye* ‘stupidity’ (Sason Arabic)
 b. *sağ* (Turkish) / *baş* (Kurdish) ‘good, healthy’ + *-ti(y)ye* ‘state of’ (Arabic suffix) = *sağ-tiye* or *baş-tiye* ‘wellness’ (Sason Arabic)
 c. *dostane* ‘friendly’ (Turkish) + *-tiyye* ‘state of’ (Arabic) = *doştanətiyye* ‘friendship’ (Tillo Arabic, Lahdo 2009:217)
- (26) *hamar* ‘donkey’ (Arabic root) + *-lık* ‘state of’ (Turkish suffix) = *hamar-lık* ‘asininity, stupidity’

- The Root comes from one language and the categorizer *n* from another.

(27)

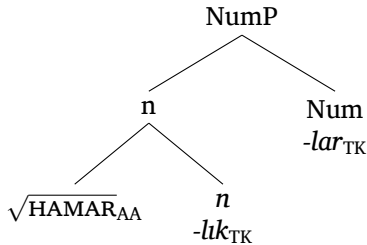


(28)



- This pattern also obeys the restriction noted above. The addition of a plural morpheme needs to be in the language expounded at the outer-most node.
 - in (25a), the plural needs to be the Arabic *-ad*.
 - in (26), the plural needs to be the Turkish *-lar*.

(29)



More combinations involving different morphemes, e.g., ablative *-DAn* are in (30c) and (31c).

- They show that another morpheme is also possible as long as it follows the restriction noted thus far.
- (31d) shows that replacing the Turkish ablative case with the Arabic preposition (as Arabic lacks overt case) is grammatical.

- (30) 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’

- (31) 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’

3.2 Patterns of trilinguals

Language-mixing patterns of trilingual speakers are crucial in demonstrating that the above hypothesis (H1) cannot be correct.

In all three examples in (32), the root *xatan* ‘son-in-law’ comes from Arabic. Yet, the examples differ wrt the ensuing morphemes.

- In (32a) and (32b), the plural is from Arabic just like the root, and the vocative case marker *-no* that follows the plural is from NK in (32a), whereas the possessive marker *-Im* in (32b) is from Turkish.
- It is even possible to have three morphemes each belonging to a different language. In (32c), the root is from AA, the plural is from Turkish, and the vocative is from NK.

- (32) a. *xatan-ad-no*
son-in-law_{AA}-PL_{AA}-VOC_{NK}
‘Sons-in-law!’
- b. *xatan-ad-im*
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!’

It is also possible to find acceptable forms with a longer sequence of morphemes, (33).

- (33) a. *xatan-lar-im-no*
son-in-law_{AA}-PL_{TK}-1SG.POSS_{TK}-VOC_{NK}
‘My sons-in-law!’
- b. *xatan-ad-im-no*
son-in-law_{AA}-PL_{AA}-1SG.POSS_{TK}-VOC_{NK}
‘My sons-in-law!’

Word-internal trilingual CS can be found in other languages as well, although very rarely.

- In Corfiot example in (34) the root comes from Hebrew, the verbalizer from Greek and the infinitival from Romance.

- (34) (Corfiot; Vardakis 2023:6)
dibur-efs-ár
talk_{Hebrew}-VRBZ_{Greek}-INF_{Romance}
‘talk’

Summary of Patterns: Overall, we can summarize the attested and unattested patterns as in (35).

- CS allows various patterns of morpheme insertion in (35a) through (35c), but not the pattern in (35d).
 - (35a) 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.
 - In (35b) the first two morphemes are from language A and they are followed by a morpheme from language B.
 - (35c) shows that each morpheme can belong to a different language,

- while in the ungrammatical (35d), 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.

- (35) **Morph1** **Morph2** **Morph3**
- Lang A < Lang B < Lang B
 - Lang A < Lang A < Lang B
 - Lang A < Lang B < Lang C
 - *Lang A < Lang B < Lang A

In fact, this generalization captures all the observed cases of intra-word CS in bilingual settings that involve many other languages, compiled in [Stefanich et al. 2019](#), as they form a small subset of (35).

3.3 Proposal: *No-Reversal Constraint*

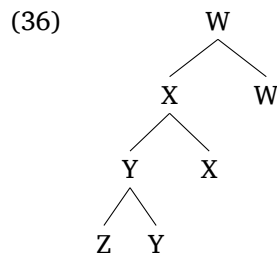
I propose a constraint called *No-Reversal Constraint* (NRC) which prohibits switching *back* to a language that has already contributed a morpheme earlier in the derivation (see [Akkuş 2024](#)).

- This constraint is at work both for the bilingual and trilingual speakers (and vacuously for monolingual speakers). This constraint is also compatible with examples that contain more than three morphemes in a single morphological word, as in (33). Those examples also follow from the No-Reversal Constraint.

↪ At what stage of the derivation does the NRC apply?

It applies to subwords in Morphology concomitant with (or prior to) linearization of MWds (morphosyntactic words), and before Vocabulary Insertion (following the definitions in [Embick and Noyer 2001](#)).¹⁰

Prediction 1: This imposes a restriction on MWds, but allows its obviation for larger structures.



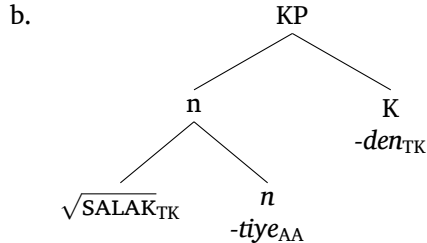
Suppose in (36), $X^0 = Z + Y + X$ forms an MWd, but W^0 itself is a separate MWd.

The No-Reversal Constraint applies to the subwords forming the complex word (i.e., any terminal that has undergone head movement in syntax to adjoin to another head). In the case at hand, it will place a restriction internal to $X^0 = Z + Y + X$, but not the set of terminal nodes $Z + Y + X + W$ since W forms a separate MWd itself.

¹⁰I have not been able to tease apart linearization from exponence. See also [Nevins 2012](#) for dissimilation effects that also apply within the M-Word.

A contrast:

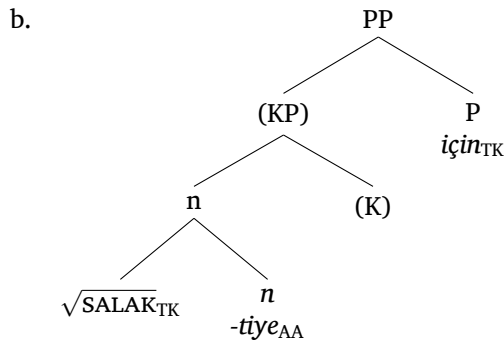
- (37) a. *salaq-tiye-den
 stupid_{TK}-DER_{AA}-ABL_{TK}
 ‘from the stupidity’



In (37b), the terminal nodes $\checkmark + n + K$ constitute an MWd, presumably due to head-movement.

On the other hand, in (38a), while the terminal nodes $\checkmark + n(+K)$ form an MWd, the preposition is itself part of a different MWd.

- (38) a. salaq-tiye için
 stupid_{TK}-DER_{AA} for_{TK}
 ‘for/because of the stupidity’



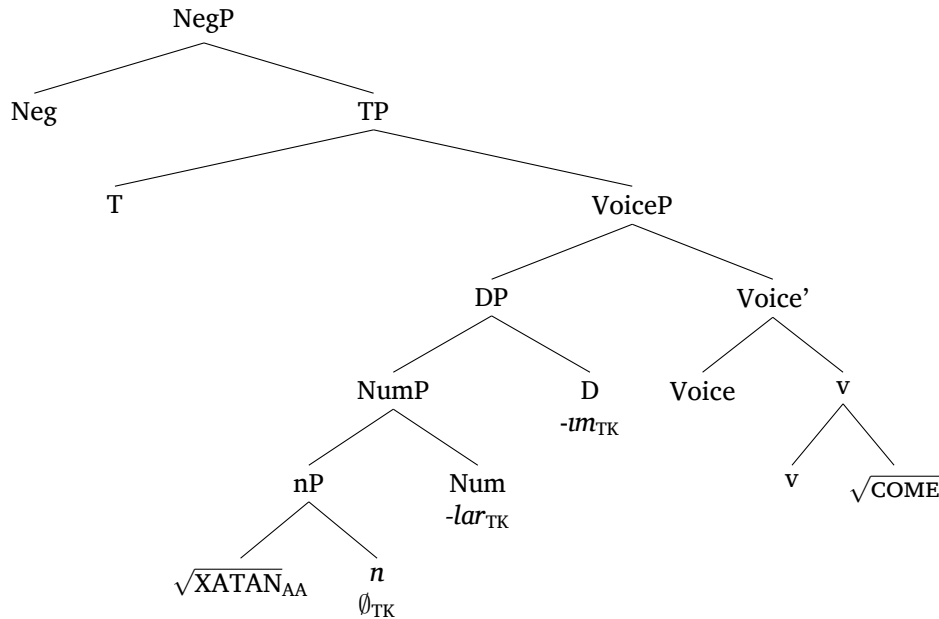
In fact, it is also possible to utter (39):

- (39) salaq-tiy-ad için
 stupid_{TK}-DER_{AA}-PL_{AA} for_{TK}
 ‘for/because of the stupidities’

This constraint could also explain the following possibilities. Note that (40c) involves combination of Spanish roots/lexical item plus Quechua inflection/grammatical properties (that are boldfaced).

- (40) a. ma-co xatan-lar-ım.
 NEG_{AA}-came_{AA} son.in.law_{AA}-PL_{TK}-1POSS_{TK}
 ‘My sons-in-law didn’t come.’
- b. bu model run-iyor.
 this_{TK} model_{TK} run_{English}-PROG_{TK}
 ‘This model is running.’ (Metehan Oğuz, March 24, 2024)
- c. (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.’

(41) Structure for (40a)



- In (40a), the terminal nodes $\sqrt{\text{XATAN}} + n + \text{Num} + \text{D}$ form a MWd, whereas the nodes $\sqrt{\text{COME}} + v + \text{Voice} + \text{T} + \text{Neg}$ forms another MWd.

Prediction 2: Nodes that do not form complex words via head-movement should not be subject to the No-Reversal Constraint. Indeed, the clitics $-(y)sa/ise$ ‘as for’, (42a), mi (interrogative), (42b), are apparent exceptions.

- (42) a. $soba-d = sa$
 $stove_{TK} - PL_{AA} = as.for_{TK}$
 ‘as for the stoves’
 b. $salaq-tiye = mi?$
 $stupid_{TK} - DER_{AA} = Q_{TK}$
 ‘is it the stupidity?’

3.4 Is the NRC string- or structure-based?

The NRC applies at the Morphology in the case studies at hand, but could it apply purely on the basis of hierarchical structures?

- The grammaticality of (42a) is suggestive that a purely string-based approach cannot work. Still it leaves some other options available.
- In particular, what is the nature of representation(s) referred to in the NRC? Suppose we have (43a) linearized as (43b):

- (43) a. Structure: $[[\text{Root } x] y]$
 b. Linearization: Root-x-y

- If the Root is Lg1, and x is Lg2, suppose we cannot switch back to Lg1 at y.
 - That is compatible with both a hierarchical and linear understanding of a NRC.

Two possible paths to tease apart linear vs. hierarchical approaches:

- (i) Linearly *; but hierarchically ✓
- (ii) Linearly ✓; hierarchically *

Making x or y a **prefix** would result in different predictions for linear vs. hierarchical approaches, and the languages in question allow us to test the option (ii), while I have not been able to find a configuration to test (i).

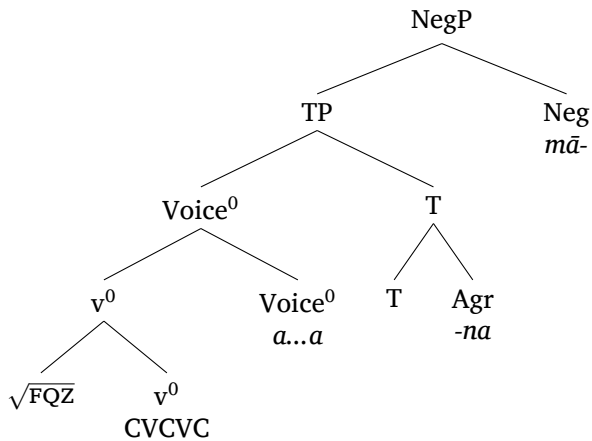
Assume y (which is structurally higher than x) is a prefix:

- On a linear approach, y_{Lg1} -Root $_{Lg1}$ - x_{Lg2} would involve no switching.
- But on a hierarchical approach, it would involve switching: it would go from Root $_{Lg1}$ to inner x_{Lg2} and then to outer y_{Lg1} .¹¹

Clausal spine in Arabic and Turkish Arabic and Turkish differ as to the position and linearization of the negative morpheme in the clausal spine. Despite its different height in the two languages, negation serves as a useful testing ground.

- (44) a. Clausal structure in Arabic: **Neg** > T > Asp > V¹²
- b. Surface order of negation: Negation is a prefix to the verbal complex. It has a long vowel and bears stress (which is shown via ' below).
- c. Sason Arabic
 'mā-faqaz-na.
 NEG-run.PST-1PL
 'We did not run.'

- (45) a. SA clausal hierarchy for (44c):
- b. SA linearization:
 Neg-TAV-Agr

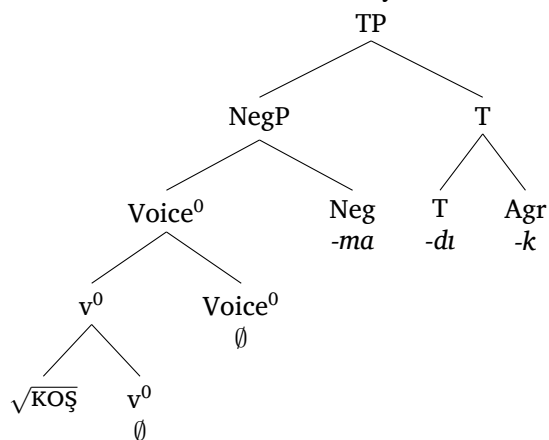


¹¹The same considerations would apply if x is linearized as a prefix, but again the languages in question don't seem to have a such a configuration.

¹²See Soltan 2007; Benmamoun et al. 2013 for Standard Arabic and several Arabic varieties, and Akkuş 2021b; Akkuş and Benmamoun 2024 for Sason Arabic verbal clauses.

- (46) a. Clausal structure in Turkish: T > Asp > Neg > V¹³
 b. Surface order of negation: Negation is a suffix to the verb. It has a short vowel obeying vowel harmony, and it is a pre-stressing affix, thus the preceding syllable bears stress.
 c. Turkish
 'koş-ma-dı-k.
 run-NEG-PST-1PL
 'We did not run.'

(47) a. Turkish clausal hierarchy for (46c):



b. Turkish linearization:
 V-Neg-Asp-T-Agr

It is possible to observe instances of intra-word CS that are informative as to the status of the NRC.

- Examples (48a-48b) show that an AA root/infinitival can be followed with multiple Turkish morphemes, such as negation, past tense and agreement.
- (48c) shows that an attempt to realize the Turkish negation as a prefix is ruled out.

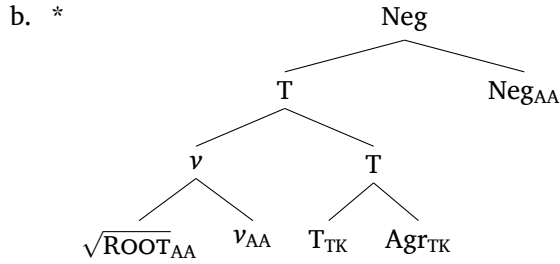
- (48) a. faqz-dı-k
 run.INF_{AA}-PST_{TK}-1PL_{TK}
 'We ran.'¹⁴
 b. 'faqz-ma-dı-k
 run.INF_{AA}-NEG_{TK}-PST_{TK}-1PL_{TK}
 'We didn't run.'
 c. *ma-'faqz-dı-k
 NEG_{TK}-run.INF_{AA}-PST_{TK}-1PL_{TK}
 'We didn't run.'

↪ Crucially, it is also not possible to use the Arabic negation, even in its prefixal position, (49).

- (49) a. *'mā-faqz-dı-k
 NEG_{AA}-run.INF_{AA}-PST_{TK}-1PL_{TK}
 'We didn't run.'

¹³See Keleşir 2001; Öztürk 2005; Akkuş 2021b; Paparounas and Akkuş 2024, a.m.o.

¹⁴Sason Arabic has developed infinitival forms (Akkuş 2021a, 2022) due to contact with Kurdish. It is also possible (but less preferred) to use the form *faqaz* 'run.PST', in which case it would result in a double marking of past tense information. This would make it similar to the double marking of conditional *he-be=se* 'if there is' in Turkish-Kurdish contact situations (Dorleijn 2006:84).



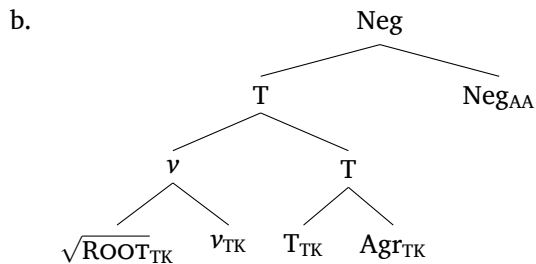
c. Linearization: Neg_{AA}-Root/v_{AA}-T_{TK}-Agr_{TK}

The ungrammaticality of (49a) is informative, this configuration is expected to be possible in terms of a linear approach, (49c), as the Arabic negation and the Arabic root/v are linearly adjacent.

However, for a hierarchical approach, the ungrammaticality is predicted since negation in Arabic is hierarchically separated from the Arabic root/v by intervening node(s), (49b).

- Again, intriguingly, (50) is licit, where the Root, Tense, Agr are from Turkish and the negation is from Arabic.

(50) a. 'mā-koş-tu-k.
NEG_{AA}-run_{TK}-PST_{TK}-1PL_{TK}
'We didn't run.'



c. Linearization: Neg_{AA}-Root/v_{TK}-T_{TK}-Agr_{TK}

4 Conclusions

- The long term bi- and multi-lingualism in southeastern Turkey has resulted in intriguing code-switching (CS) patterns.
- The trilingual CS obeys the the *No-Reversal Constraint (NRC)*, and prior approaches fail to capture the full range of the data:
 - This constraint applies to subwords in Morphology concomitant with (or prior to) linearization of morphosyntactic words (MWds).
 - Moreover, the NRC also exhibits a bottom-up effect (i.e., no [[A B] A] structures).

References

- Akkuş, Faruk. 2017. Peripheral Arabic Dialects. In *The Routledge Handbook of Arabic Linguistics*, ed. Elabbas Benmamoun and Reem Bassiouney, 454–471. Routledge.
- Akkuş, Faruk. 2021a. Evidence from Sason Arabic for \bar{A} -Movement Feeding Case-Licensing Relations. *Linguistic Inquiry* 53:589–607.

- Akkuş, Faruk. 2021b. (Implicit) Argument Introduction, Voice and Causatives. Doctoral Dissertation, University of Pennsylvania.
- Akkuş, Faruk. 2022. On Causee in Sason Arabic. *Syntax* 25:299–334.
- Akkuş, Faruk. 2024. No-Reversal Constraint and beyond: Word-internal language mixing in Anatolia. *Linguistic Variation* 1–43.
- Akkuş, Faruk, and Elabbas Benmamoun. 2024. Categorical Features and The Feature Structure of the Present Tense: Perspectives from Language Variation and Change. Ms., University of Massachusetts Amherst and Duke University URL https://www.dropbox.com/scl/fi/s7qq1hfp6n9m7wbqe9xu4/akkus-benmamoun-copula_Arabic-9-24.pdf?rlkey=io6domaa6viwiabu5mftg3ap7&st=9154qo6r&dl=0.
- Alexiadou, Artemis. 2017. Building verbs in language mixing varieties. *Zeitschrift für Sprachwissenschaft* 36:165–192.
- Alexiadou, Artemis, Terje Lohndal, Tor A Áfarli, and Maren Berg Grimstad. 2015. Language mixing: a distributed morphology approach. In *Proceedings of NELS 45*, ed. Thuy Bui and Deniz Özyıldız, volume 45, 25–38. Amherst, MA: GLSA.
- 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.
- Armstrong, Spyros, and Marilena Karyolemou. 2023. Contact-Induced Change in an Endangered Language: The Case of Cypriot Arabic. *Languages* 8:10.
- 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.
- Benmamoun, Elabbas, Mahmoud Abunasser, Rania Al-Sabbagh, Abdelaadim Bidaoui, and Dana Shalash. 2013. The location of sentential negation in Arabic varieties. *Brill's Journal of Afroasiatic Languages and Linguistics* 5:83–116.
- Citko, Barbara. 2014. *Phase theory: An introduction*. Cambridge University Press.
- 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.
- Dorleijn, Margreet. 2006. Turkish-Kurdish language contact. In *Turkic languages in contact*, ed. Hendrik Boeschoten and Lars Johanson, 74–94. Harrassowitz Wiesbaden.
- Embick, David, and Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32:555–595.
- González-Vilbazo, Kay, and Luis López. 2011. Some properties of light verbs in code-switching. *Lingua* 121:832–850.
- Harbour, Daniel. 2011. Valence and atomic number. *Linguistic Inquiry* 42:561–594.
- Jake, Janice L, Carol Myers-Scotton, and Steven Gross. 2002. Making a minimalist approach to codeswitching work: Adding the matrix language. *Bilingualism: language and cognition* 5:69–91.
- Jastrow, Otto. 1978. *Die mesopotamisch-arabischen qaltu-dialekte. band i: Phonologie und morphologie*. Wiesbaden.
- Kelepir, Meltem. 2001. Topics in Turkish syntax: Clausal structure and scope. Doctoral Dissertation, Massachusetts Institute of Technology.
- Khan, Geoffrey, and Masoud Mohammadirad. 2024. *Language Contact in Sanandaj: A Study of the Impact of Iranian on Neo-Aramaic*. De Gruyter.
- 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.
- Laks, Lior. 2014. The cost of change: Plural formation of loanwords in palestinian and jordanian arabic. *Zeitschrift für Arabische Linguistik* 5–34.
- López, Luis, Artemis Alexiadou, and Tonjes Veenstra. 2017. Code-switching by phase. *Languages* 2:9.
- 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.

- Muysken, Pieter. 1997. Media Lengua. In *Contact languages: A wider perspective*, ed. Sarah Grey Thomason, 365–426. John Benjamins.
- Nevins, Andrew. 2012. Haplological dissimilation at distinct stages of exponence. In *The morphology and phonology of exponence*, ed. Jochen Trommer, 84–116. Oxford University Press.
- Öztürk, Balkız. 2005. *Case, referentiality and phrase structure*. Amsterdam/Philadelphia: John Benjamins.
- Paparounas, Lefteris, and Faruk Akkuş. 2024. Anaphora and agreement in the Turkish DP: Delimiting binding-through-Agree. *Natural Language & Linguistic Theory* 42:633–700. URL <https://link.springer.com/article/10.1007/s11049-023-09583-4>.
- 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.
- Soltan, Usama. 2007. On formal feature licensing in minimalism: Aspects of standard arabic morphosyntax. Doctoral Dissertation, University of Maryland.
- Stefanich, Sara, Jennifer Cabrelli, Dustin Hilderman, and John Archibald. 2019. The morphophonology of intraword codeswitching: Representation and processing. *Frontiers in Communication* 4:54.
- Talay, Shabo. 2007. The influence of Turkish, Kurdish and other neighbouring languages on Anatolian Arabic. *Romano-Arabica* 179–189.
- 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>.
- Tucker, Matthew A., and Itamar Kastner. 2019. Non-concatenative Morphology. *Ms., Oakland University and The University of Edinburgh*.
- 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*.
- Wittrich, Michaela. 2001. *Der arabische Dialekt von Äzex*, volume 25. Otto Harrassowitz Verlag.
- Zabbal, Youri. 2002. The semantics of number in the Arabic number phrase. *Master's thesis, University of Alberta*.

5 Appendix: Comparison with alternative analyses

The NRC fares better compared to prominent alternative proposals, which fail to capture the full range of the attested patterns.

Poplack (1980, 1981): This approach argues that code-switching could only target two free morphemes, and it could not apply between morphemes within the same word.

This view effectively imposes a prohibition on code-switching in head movement and word-internal contexts.

Evaluation: I do not dwell on this proposal since studies have shown that this statement is too strong, and such patterns indeed are attested. In this talk alone, we have seen many examples to that effect, which shows that language-mixing is not restricted to free morphemes.

Belazi et al. (1994): This study 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.
 - 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.

Evaluation: This proposal is also too restrictive, in that it would incorrectly rule out examples in (25a) and (26), 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 2.1. Those examples involve the root coming from Turkish, whereas the categorizer *v* and Voice from Arabic.

More challenging are the examples in (32), repeated here as (51), which involve functional categories within the nominal phrase that belong to different languages.

- (51)
- a. xatan-ad-no
son-in-law_{AA}-PL_{AA}-VOC_{NK}
‘Sons-in-law!’
 - b. xatan-ad-im
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!’

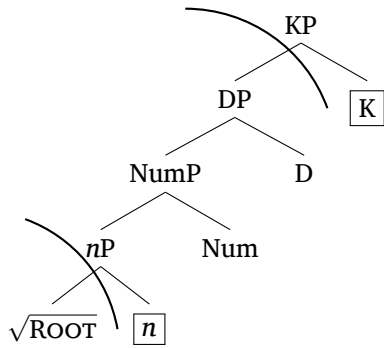
López et al. (2017): This analysis is built on the *phase*-theory and argues that code-switching is governed by phase-theoretic considerations.

López et al. (2017): bilinguals have multiple externalization systems or PFs, and 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.

Assuming the standard structure (52) for the nominal phrases (which also straightforwardly extends to Arabic/Semitic, Kurdish and Turkish, as evidenced also by the morpheme ordering) - López et al. (2017, (7)) suggest that *n* and *K* are phase heads (see Citko 2014 for the same phasehood properties in the nominal domain).

(52)



According to this formulation, the complement of a phase head and the phase head itself are transferred in different phases. In (52), 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*.

Evaluation: Despite capturing a large set of the examples (e.g., (19), (23), (24), (25a), (26), (30), (31)), it incorrectly rules out many acceptable forms too.

- e.g., those in (20), (32b), and (33b). For example, in (32b), 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 (20), 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.

Another possible direction: Configurations that involve combinations of prefixes and suffixes - which retain the same hierarchy across languages - would be informative too.