Tu+ 1:
Proceedings of the first workshop on
Turkish, Turkic and the languages of Turkey

November 21 - November 22, 2015
University of Massachusetts, Amherst

Edited by
Faruk Akkuş, İsa Kerem Bayırlı, Deniz Özyıldız

Thank you

The 1st Workshop on Turkish, Turkic and the languages of Turkey (Tu+1) was held at the University of Massachusetts, Amherst on 21-22 November 2015, in collaboration with Yale University. In addition to the two invited talks by Sabine Iatridou (MIT) and Jaklin Kornfilt (Syracuse University), the workshop hosted 23 oral and poster presentations. The presenters came from 17 different institutions, 6 of which were non-US institutions.

The articles in this proceedings volume are a selection of papers presented at Tu+1. Our hope is that this workshop will be held annually and that it will provide a platform for discussion of empirical and theoretical questions raised by Turkic languages and the languages spoken within Turkey.

We are grateful to the linguistics departments at Yale and at UMass for supporting the workshop, morally, technically and financially. Rosetta Berger, Rajesh Bhatt, Robert Frank, Vincent Homer, John Kingston, Tom Maxfield and Michelle McBride deserve special mention. We would also like to thank all those who gave a helping hand during the workshop. We are grateful to the Graduate Linguistic Student Association for publishing this volume, to Leah Chapman for agreeing to design the cover, and last but not least, to our presenters and audience for making this workshop an enjoyable learning experience.

The organizers and the editors

Table of Contents

On the complex connectives in Turkish

İsa Kerem Bayırlı

Massachusetts Institute of Technology

1.   Introduction

This paper is concerned with the absence of free choice-type inferences in the context of several expressions in Turkish. The complex disjunction *ya...ya...* and complex conjunction *hem...hem...* do not give rise to free choice-type effects (i.e. strengthening to wide scope conjunction) in some contexts in which their simple versions do. To capture these observations, it is claimed here that we need to revise the conditions on the distribution of the *exh* operator.

Spector (2014) has developed an analysis of French complex disjunction, in which the positive polarity nature of the complex disjunction is linked to its giving rise to "obligatory exclusivity inferences" (modelled as the obligatoriness of an *exh* operator c-commanding the complex disjunction). The analysis developed here retains this insight. It is shown, however, that the observation that *ya...ya...* is not acceptable in the scope of an existential modal is unexpected. The condition on the distribution of *exh* is revised so that more constructions (but the right ones) are ruled out. The analysis developed for *ya...ya...* is extended to the complex conjunction *hem...hem....* The absence of the free choice with *ya...ya...* and *hem...hem...* are shown to follow from similar assumptions.

Here is the outline. I briefly introduce (aspects of) the analysis of French complex disjunction by Spector (2014) and the derivation of free choice inferences in Fox (2007). I go on to explain why the unacceptability of the complex disjunction under an existential modal in Turkish does not follow from a combination of these two analyses and offer a new constraint on the distribution of the *exh* operator. Later in the paper, I tackle the issues concerning conjunction.

2.   Preliminary Remarks

In Turkish, there are (at least) two strategies for expressing disjunction. Two disjuncts can be separated by a single disjunction morpheme as in (1)

(1)   Ekin dondurma ya-da çikolatalı puding   yedi
       Ekin ice-cream or   chocolate pudding eat.PAST.3sg
       'Ekin ate ice cream or chocolate pudding'

Another strategy is to repeat the morpheme *ya* on both disjuncts. I will call this strategy "complex disjunction". "Complex" here is intended as a purely descriptive term in an attempt to contrast this construction to the "simple" one in (1).

(2)   Ekin ya dondurma ya-da çikolatalı puding   yedi
       Ekin or ice-cream or   chocolate pudding eat.PAST.3sg
       'Either Ekin ate ice cream or she ate chocolate pudding'

This paper is mainly about the properties of the complex disjunction in Turkish. In next section, I introduce some observations about *ya...ya...* with reference to the analysis of the complex disjunction in French developed by Spector (2014). Later, the analysis of free choice effects developed in Fox (2007) will be presented. These two studies form the basis for the analysis that is developed in this paper.

2.1   Spector (2014) on complex disjunction

In this section, I present the analysis of the French complex disjunction by Spector (2014). In doing so, I have two ambitions. First, this will give me a chance to report some properties of the Turkish complex disjunction that are relevant to the analysis that I deveop later in the paper. Second, Spector's analysis for the PPI-hood of the French complex disjunction will provide me with a "type of thinking" that will be at the core of my explanation for the unacceptability of *ya...ya...* under an existential modal. In introducing Spector's analysis, I will be using examples from Turkish since all the relevant judgments that are crucial to Spector's argumentation seem to be identical. That is to say, Spector's analysis of the PPI-hood of the French complex disjunction *soit...soit...* can be extended to *ya...ya...* without any immediate problem that I can see.

Let me start with simple disjunction in Turkish and contrast with *ya...ya...*. Simple disjunction gives rise to DeMorgan readings ($\sim(p \vee q) \equiv \sim p \ \& \sim q$) when it is in the scope of the negation:

(3)   Ekin dondurma ya-da çikolatalı puding   yemedi
       Ekin ice-cream or   chocolate pudding   eat.NEG.PAST.3sg
       'Ekin ate neither ice cream nor chocolate pudding'

Spector notes that the complex disjunction in French (*soit...soit...*) is judged to be unacceptable in the scope of negation. The same observation can be extended to *ya...ya...*.

(4)   # Ekin ya dondurma ya-da çikolatalı puding   yemedi
       Ekin or ice-cream or   chocolate pudding   eat.NEG.PAST.3sg
       'Ekin didn't eat ice cream or chocolate pudding'

Complex disjunction structures (in French and in Turkish) are licensed again when the negation is itself in the scope of a downward entailing operator (here negation).

(5) (Antakya'ya gitmişken) ya künefe ya baklava yememek kabuledilebilir
     (once in Antakya)   YA kunefe YA baklava eat.NEG   acceptable

     bişey   değil
     a.thing NEG
     'It is not an acceptable thing to not eat kunefe or baklava (once in Antakya)'

A description of these observations is to call *ya...ya...* a positive polarity item (PPI).[1]

The PPI-hood of *ya...ya...* seems to correlate with another property of this construction. Namely, it "triggers much stronger exclusivity inferences than" simple disjunction. The contrast in the following dialog exemplifies this.

(6) A: Ali Ankara'ya ya-da İzmir'e gidecek
         Ali Ankara-to or   İzmir-to go.will.3sg
         'Ali will go to Ankara or İzmir.'

     B: Kesinlikle! Hatta ikisine de gidecek.
         'Absolutely! In fact, he will go both'

(7) A: Ali ya Ankara'ya ya İzmir'e gidecek
         Ali or Ankara-to or İzmir-to go.will.3sg

     B: # Kesinlikle! Hatta ikisine de gidecek
         'Absolutely! In fact, he will go both'

The idea is that, by uttering *Kesinlikle!* "Absolutely", the speaker B commits herself to the truth of the first assertion. With simple disjunction, the exclusivity inferences can be cancelled and the speaker B's response is felicitous. With *ya...ya...*, on the other hand, these inferences are obligatory (or almost obligatory) and the speaker B's response is infelicitous as it does not observe this inference. As Spector notes, this is not reason to believe that the complex disjunction is the realization of the logical exclusive disjunction operator $\veebar$. As seen below, in the scope of a universal modal, the complex disjunction gives rise to readings that are incompatible with an exclusive analysis.

---

[1] See Szabolcsi (2004) on the distribution of PPIs.

For positive polarity items, anti-licensing is usually thought to be a local phenomenon (Szabolsci, 2004). That is, the positive polarity items are anti-licensed only in the scope of a clause-mate negation. With the complex disjunction, the situation is different. The complex disjunction in Turkish, just like the complex disjunction in French, is anti-licensed no matter how far negation is.

(i) # Ali'nin   ya Ankara'ya ya-da İzmir'e gideceğini sanmıyorum
       Ali-GEN or Ankara.to or   İzmir.to go.will.3sg think.NEG.IMPF.1sg
       Int. 'I don't think that Ali will go to Ankara or İzmir'

This leads Spector to conclude that the complex disjunction is a new type of a PPI: it is *a global PPI*.

(8)   A: Bu ödülü almak için, ya ilk ya ikinci soruyu cevaplaman lazım
             (to obtain this gift), or first or second question answer.2sg necessary
             'To obtain this gift, you must answer the first or the second question'

       B: Kesinlikle! Soruların ikisi birden cevaplamak da mümkün.
             'Absolutely! You can solve both!'

The salient reading of (8) is that it is enough to answer the first or the second question to obtain the gift. There is no extra requirement that you not answer the questions both, as would be predicted with an exclusive construal. To explain the observation that *soit…soit…* gives rise to exclusivity inferences without being an exclusive disjunction operator itself, Spector proposes the following constraint for French, which I adopt for Turkish since all the relevant judgments are identical.

(9)   *Constraint on the Turkish complex disjunction*
       ya…ya… must occur in the scope of an *exh* operator.

Spector adopts a relatively simple version of lexical entry for the *exh* operator. We will need a more involved entry in the next section.

(10) $[[exh(S)]] = \lambda w.[[S]](w)=1 \ \& \ (\forall \varphi \in ALT(S), \varphi(w)=1 \rightarrow ([[S]] \subseteq \varphi))$

That is, the prejacent is true and for any alternative of the prejacent, the alternative is true only if it is entailed by the prejacent. In other words, if an alternative is not entailed by the prejacent, then the alternative is false. Assuming that the only alternative to *must p or q* is *must p and q*,[2] it is not hard to see that *Exh(must p or q)* is compatible with the continuation in 8B (that is, *(exh(must p or q)) and (can p and q)* is not a contradiction).

It is time to present how Spector derives the PPI-hood of the complex disjunction from the assumption that it must occur in the scope of an exhaustivity operator. The central idea is that there is a condition on the *exh* operator that makes it necessary that it strengthen the meaning of the sentence to which it is attached. This is stated as the following constraint:

(11)   *Economy Constraint on the distribution of exh*
       An occurrence of *exh* in a given sentence S is not licensed if eliminating this occurrence leads to S' such that S' entails S.

With this, we can show that *exh* will not be licensed in DE contexts as its elimination does not weaken the meanings. We observe this situation below with negation as a special case. The first expression is the one in which *exh* comes between negation and the disjunction.

---

[2] There are some problems with this statement. Usually, *may* is taken to be an alternative to *must* and each disjunct is taken to be an alternative to the disjunction (Sauerland, 2004). As things currently stand, I have to restrict the alternatives to *must p and q* to get the readings I am interested in. Both problems can be solved when we adopt the *exh* operator in Fox (2007) to be discussed in the next section and the calculation of alternatives in Fox (2007, fn. 35).

(12)   Let S be ¬(exh (p ∨ q))
            S = ¬ (p ⊻ q)
            S' = ¬ (p ∨ q)
            S'⇒S

We do not get a strengthened meaning. Neither do we get a strengthening when *exh* is above the negation operator.

(13)   Let S be *exh* (¬ (p ∨ q))
            S = ¬ (p ∨ q)
            S' = ¬ (p ∨ q)
            S'⇒S

Therefore, *exh* is not licensed when *ya…ya…* is in the scope of a single neg operator. This conflicts with the requirement that *ya…ya…* occur in the scope of an *exh* operator. That is, complex disjunction cannot occur in the scope of a single negation. We describe this situation by calling it a Positive Polarity Item.

2.2   Fox (2007) on the free choice-type inferences

The free choice puzzle can be described as the observation that, in the scope of an existential modal, a disjunctive expression ends up having a meaning that is stronger than what is entailed by what seems to be its logical form.

(14)   Ekin dondurma ya-da çikolatalı puding yiyebilir
            Ekin ice-cream or   chocolate pudding eat.may.3sg
            'Ekin is permitted to eat ice-cream *and* she is permitted to eat chocolate pudding'

The total meaning of (14), given in (15)b, does not follow from the logical form of the sentence given in (15)a.

(15)   a.   ◊(p∨q)
            b.   ◊p & ◊q
                  where ◊(p∨q) ⊭ ◊p&◊q

It has been observed that this conjunctive inferences become unavailable when the disjunction is in the scope of a downward entailing operator (Alonso-Ovalle 2006, see also Fox 2007).

(16)   Ekin dondurma ya-da çikolatalı puding   yiyemez
            Ekin ice-cream or   chocolate pudding eat.may.NEG.PAST.3sg
            'Ekin can eat neither ice cream nor chocolate pudding'

It is an observed property of implicatures that they are cancelled in downward entailing environments (Gazdar 1979 a.o.) or, at least, strongly dispreferred (Chierchia et al. 2012).

If the conjunctive readings are implicatures, their being cancelled in a downward entailing environment is expected. This leads to the view that the conjunctive inferences must be derived as an implicature/inference in some way. That is, we do not take the conjunctive meanings to be the literal meaning of the sentence in (14).

We have seen that Spector (2014) has derived the obligatory exclusive inferences associated with the complex disjunction by appending an *exh* operator to a c-commanding position. Fox (2007) notices that, with a proper lexical entry for *exh*, doing this recursively, i.e. adding another *exh* to a sentence that already has an *exh* in it, will yield readings that entail the free choice inference associated with these sentences. The entry Fox proposes is the following:

(17) $[[\text{Exh}]](A_{<st,t>})(p_{st})(w) = [p(w) \ \& \ [\forall q \in \text{I-E}(p,A) \rightarrow \sim q(w)]]$

As before, the proposition expressed by the prejacent is true and some propositions that are alternatives to the prejacent are false. Which of the alternative propositions are false? These are now called the *innocently excludable* alternatives of the prejacent. A proposition is an innocently excludable alternative with respect to a prejacent only if its negation is member of every set that contains the prejacent and as many negated alternatives as possible. Here is a formal definition of this set given in Fox (2007).

(18)   I-E(p,A)= $\bigcap\{A'\subseteq A$: A' is the maximal set in A such that $A'^{\sim} \cup \{p\}$ is consistent$\}$
   where $A'^{\sim} = \{\sim p: p \in A\}$

What we do is look at those maximal (i.e. as many membered as possible) subsets of A whose exclusion from A is consistent with the prejacent. We, then, find the members that are in the intersection of all such maximally consistent sets. These are the alternatives that are negated. Let us see what happens when we append an *exh* operator to a sentence with a disjunction under existential modal. Following Sauerland (2004), we take each disjunct to be in the set of the alternatives.

(19)   Exh(C)(you may eat the cake or pizza)
        ALT$_{\text{AorB}}$ = {A or B, A, B, A and B}

By convention, the propositions are put together so that they get stronger as we move to the right. Propositions that are related by their strength are indicated with lines. With independent propositions, there are no lines.

(20)
$$C = \Diamond(p \lor q) \ \begin{array}{c} \diagup \Diamond p \diagdown \\ \diagdown_{\Diamond q} \diagup \end{array} \ \Diamond(p\&q)$$

There are two two-membered sets whose exclusion does not contradict the prejacent. The prejacent and the negation of the members of these sets are maximally consistent sets. (Notice that there are no three membered such set.)

(21)
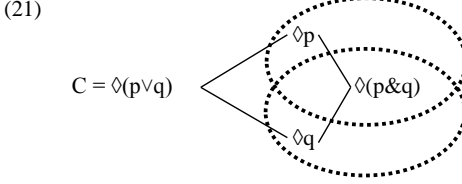
We notice that there is only one member in the intersection of these two sets. This will be the proposition that will be negated. We obtain the following results.

(22)   I-E($\Diamond$(p$\lor$q), ALT($\Diamond$(p$\lor$q))) = {$\Diamond$p, $\Diamond$(p&q)} $\cap$ {$\Diamond$q, $\Diamond$(p&q)}
        {$\Diamond$(p&q)} = I-E($\Diamond$(p$\lor$q), ALT($\Diamond$(p$\lor$q)))
        Exh(C)($\Diamond$(p$\lor$q)) = $\Diamond$(p$\lor$q) & ~$\Diamond$(p&q)

This does not entail the free choice reading (crucially, it is compatible with it). It would be a stipulation to assume that *exh* can be applied only once. If we do not make this stipulation, we see that the sentence has yet another parse. (Here I am simply repeating the calculations from Fox 2007)

(23)   Exh(C')[Exh(C) you may eat cake or pizza]
        where C'= {Exh(C)(p): p $\in$ C}

It turns out that when we compute the meaning that comes out of this parse, we get a reading that entails the free choice.[3]

(24)   C' = {1. Exh(C)($\Diamond$(p$\lor$q)), 2. Exh(C)($\Diamond$p), 3. Exh(C)($\Diamond$q), 4. Exh(C)($\Diamond$(p&q))}

We calculate each of these new "higher order" alternatives.

(25)   1. Exh(C)($\Diamond$(p$\lor$q)) = $\Diamond$(p$\lor$q) & ~$\Diamond$(p&q)
        2. Exh(C)($\Diamond$p) = $\Diamond$p & ~$\Diamond$q
        3. Exh(C)($\Diamond$q) = $\Diamond$q & ~$\Diamond$p
        4. Exh(C)($\Diamond$(p&q)) = $\Diamond$(p&q) (ignored as its negation is entailed by the prejacent)

C'= $\Diamond$(p$\lor$q) & ~$\Diamond$(p&q)

$\Diamond$p & ~$\Diamond$q

$\Diamond$q & ~$\Diamond$p

---
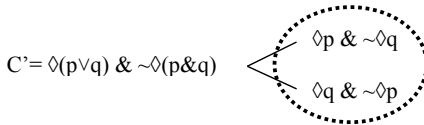
[3] The universal alternative of the existential modal is ignored here for illustrative purposes. The results are not affected by this omission.

The negation of both alternatives is consistent with the prejacent.

(26)   1. ◊(p∨q) & ~◊(p&q)   prejacent
       2. ~(◊p & ~◊q)   neg alt
       3. ~(◊q & ~◊p)   neg alt
       4. ~◊p ∨ ◊q   from 2
       5. ~◊q ∨ ◊p   from 3
       6. ◊p → ◊q   from 4
       7. ◊q → ◊p   from 5
       8. ◊p ←→ ◊q   from 6,7
       9. ◊(p∨q)   from 1
       10. ◊p∨ ◊q   from 9
       11. ◊p&◊q   from 8,10
       12. ~◊(p&q)   from 1
       13. (◊p&◊q)& ~◊(p&q)   from 11,12

We derive a proposition that entails the free choice reading.

3.   The puzzle of the Turkish complex disjunction

We have already seen that the Turkish simple disjunction gives rise to free choice-type inferences in the scope of an existential modal.

(27)   Ekin dondurma ya-da çikolatalı   puding   yiyebilir
       Ekin ice-cream or   chocolate pudding eat.may.3sg
       'Ekin is permitted to eat ice-cream *and* she is permitted to eat chocolate pudding'

We observe, however, that the same sentence becomes unacceptable when we use *ya…ya…* instead of the simple disjunction.

(28) # Ekin ya dondurma ya-da çikolatalı puding   yiyebilir
       Ekin or ice-cream or   chocolate pudding   eat.may.3sg
       Int. 'Ekin is permitted to eat ice-cream *and* permitted to eat chocolate pudding'

Notice the claim is something more than the unavailability of the free choice readings. The observation being made here is that *Turkish complex disjunction is not acceptable in the scope of the existential modal.*[4] We might perhaps hope that we can rule *ya…ya…* in the

---

[4] There is one prosodic shape of the sentence in which the free choice readings are again available. This is when there is strong accent on both disjuncts.

(ii)   Ekin ya DONDURMA   ya-da ÇİKOLATALI PUDİNG   yiyebilir
       Ekin or ice-cream   or   chocolate   pudding   eat.may.3sg
       'Ekin is permitted to eat ice-cream *and* she is permitted to eat chocolate pudding'

This problem will not be discussed in this paper due to space limitations.

context of the existential modal the way Spector has ruled out *soit…soit…* in the scope of negation. We might try to show that the *exh* operator is not licensed here due to the economy condition on the distribution of *exh* (repeated below):

(29)   *Economy Constraint on the distribution of exh:*
       An occurrence of *exh* in a given sentence S is not licensed if eliminating this occurrence leads to a sentence S' such that S' entails S.

It turns out that this strategy, as things stand right now, does not work. The *exh* strengthens the meaning when it is immediately above the disjunction.

(30)   Let S be ◊exh (p ∨q)
       $S = ◊((p ∨q) \& \sim(p\&q))$
       $S = ◊(p ⊻q)$
       $S' = ◊ (p ∨q)$
       S' ⇏ S

Consider the model under which all worlds are p-worlds and q-worlds. Under this model, S is false while S' is true. That is, S' does not entail S. This should get the *exh* operator licensed. Note that the state of affairs will not change if the *exh* operator is higher in the structure.

(31)   Let S be exh ◊(p ∨q)
       $S = ◊(p ∨q) \& \sim◊ (p\&q)$
       $S' = ◊ (p ∨q)$
       S' ⇏ S

Consider now the model under which there is a p and q-world. Under this model, S is false while S' is true. That is, S' does not entail S. Again, the *exh* operator should get licensed - contrary to the observations. It seems that if we want to get the unacceptability of the complex disjunction from the distribution of the *exh* operator, we need to revise the condition. How should we do that? Here is the intuition. We have noticed that both ◊*exh* *(p ∨q)* and *exh ◊(p ∨q)* are stronger than their *exh*-less alternatives. There could, however, be another alternative that is stronger than both these sentences. What if *exh(exh◊ (p ∨q))* was one of the alternatives against which both ◊*exh (p ∨q)* and *exh ◊(p ∨q)* are evaluated?[5]

(32)   exh' exh◊ (p ∨q) ↔   ◊p &◊q&~◊(p&q)
       exh◊ (p ∨q)   ↔   ◊(p ∨q) &~◊ (p&q)
       ◊ exh (p ∨q)   ↔   ◊(p ⊻q)
       ◊p &◊q&~◊(p&q) ⊨ₐ ◊(p ∨q) &~◊ (p&q)
       ◊p &◊q&~◊(p&q) ⊨ₐ ◊(p ⊻q)

We observe that *exh(exh◊ (p ∨q))* entails (in fact, stronger than) both ◊*exh (p ∨q)* and

---

[5] ⊨ₐ is asymmetric logical entailment.

exh $\lozenge(p \lor q)$. All we need now is to get *exh(exh$\lozenge$ (p $\lor$q))* to be a competitor. Here is one way to do it.

(33)   *The Strong Meaning Condition on exh*
         For some sentence S with an occurrence of *exh* in it
         Let S_0 be derived from S by deleting this *exh*
         Let $S_1...S_n$ for some n $\geq$0 be derived from S_0 by adding *exh* one by one to a
         position c-commanding the eliminated *exh*
             The occurrence of *exh* in S is not licensed if
                     $\exists$n such that S_n asymmetrically entails S

There are two novelties in the new constraint that I have formulated: the c-command condition and the asymmetric entailment condition. Here is the justification for the first one. If we do not add the c-command condition, we seem to rule out *(exh($\square$(p $\lor$q)))*, which we have already seen to be a possible representation. Its meaning is re-calculated below.

(34) $(exh(\square(p \lor q))) \leftrightarrow \square(p \lor q) \& \sim \square(p \& q)$
                        $\leftrightarrow \square(p \lor q) \& (\lozenge \sim p \lor \lozenge \sim q)$

Without c-command condition, we could eliminate *exh* and tuck in another *exh* between universal modal and disjunction.

(35) $(\square(exh(p \lor q))) \leftrightarrow \square(p \veebar q)$

$\square(p \veebar q)$ is strictly stronger than $\square(p \lor q) \& (\lozenge \sim p \lor \lozenge \sim q)$ as the readers can verify for themselves. Why asymmetric entailment? Consider the matrix disjunction in (36)

(36)   Let S= exh (p $\lor$q)
            S'= exh(exh (p $\lor$q))
            S $\Rightarrow$ S' & S' $\Rightarrow$ S

Fox (2007) has shown that a matrix disjunction is not strengthened beyond the first application of the *exh* operator. Then, if the relation was just entailment, since a proposition entails itself, we would have ruled out the matrix disjunction. Asymmetric entailment eliminates this possibility.
    There is a problem, though. How does this new constraint work with respect to negation? No issue arises when the *exh* operator comes between disjunction and negation. It is anti-licensed by the condition on the distribution of the *exh* operator. We eliminate *exh* and add zero *exh* to a commanding position.

(37)   Let S be ¬(exh (p ∨ q))
            S = ¬ (p $\veebar$ q)
            S' = ¬ (p ∨ q)
            S $\nRightarrow$ S' & S' $\Rightarrow$ S

However, when the *exh* operator is above the negation, we have a redundant *exh* operator. Redundancy of this kind is not ruled by our condition.

(38)   Let S be exh ($\neg$ (p $\vee$ q))
        S = $\neg$ (p $\vee$ q)
        S' = $\neg$ (p $\vee$ q)
        S'$\Rightarrow$S & S' $\Rightarrow$ S

I would like to introduce a new condition that will target specifically redundancy of this type.

(39)   *The Strong Meaning Condition on exh*
        For some sentence S with an occurrence of *exh* in it
        Let S_0 be derived from S by deleting this *exh*
        Let $S_1 \ldots S_n$ for some n ≥ 0 be derived from S_0 by adding *exh,* one by one, to a
        position c-commanding the eliminated *exh*
        The occurrence of exh in S is not licensed if
        i. $\exists$n such that S_n a-entails S or
        ii. $S_0 = S$

We still have a problem. We have seen that there is some way with which we can eliminate the expressions $\lozenge exh(p \vee q)$ and $exh\lozenge(p \vee q)$, namely by reference to the competitor $exh(exh\lozenge(p \vee q))$. However, we have not said anything about what will rule out *exh* $exh\lozenge(p \vee q)$ as the base against which the licensing condition is computed. Notice that there is no possible strengthening and no redundancy. What I would like to suggest is that *exh* $exh\lozenge(p \vee q)$ cannot be generated in the first place given the following constraint.

(40)   *Condition on Rescuers*
        Let A contain a lexical item whose occurrence context is not met,
        Let $A_1 \ldots A_j$ be the minimal licit continuation of A^6, then,
        If a continuation of A_j is grammatical, then A_j is grammatical

Note that the occurrence context of the complex disjunction is met when there is an *exh* operator c-commanding it. To see how the version of the condition works, consider how it rules out a sentence of the form *exh exh$\lozenge(p \vee q)$* with complex disjunction. Let A $= p \vee q$. Then one minimal licit continuation of A is $A_1 A_2$ where $A_1 = \lozenge(p \vee q)$ and $A_2 = exh\lozenge(p \vee q)$. We are asking whether or not $A_3 = exh\ exh\lozenge(p \vee q)$ can be obtained. If it can be obtained then A_2 is grammatical. Yet, we know already that A_2 is not grammatical given that the strong meaning condition anti-licenses it. Therefore, A_3 cannot be obtained. [7] In short, we

---

[6] $A_1 \ldots A_j$ is the minimal licit continuation of A if the requirement on A is not meet in any A_{j-n} for any n such that j>n>0 and the requirement on A is met in A_j

[7] This might look circular. Previously, I have used *exh(exh($\lozenge(p \vee q)$))* to eliminate *exh($\lozenge(p \vee q)$)*. Now, I am using *exh($\lozenge(p \vee q)$)* to eliminate *exh(exh($\lozenge(p \vee q)$))*. I would like to claim that it is not circular. The mechanism with which competitors for the strong meaning are constructed and evaluated is not the same mechanism with which expressions of a natural language are defined/generated. The competitor *exh(exh($\lozenge(p \vee q)$))* eliminates the natural language object *exh($\lozenge(p \vee q)$)* and the ungrammaticality of the natural language object

eliminate *exh exh◊(p∨q)* by the following requirement: the first point in the derivation at which the complex disjunction is licensed by an *exh* must be a grammatical step.

We can now talk about a prediction of the system. It has been noted that the singular indefinites do not give rise to free choice effects. This is exemplified below with Turkish simple disjunction.

(41)　Bir öğrenci Ankara'ya　ya-da İstanbul'a　gitti
　　　　a　　student Ankara-to　or　　İstanbul-to go.PAST
　　　　'A student went to Ankara or İstanbul'
　　　　cannot mean: 'A student went to Ankara and a student went to İstanbul'

In Fox (2007), the absence of the free choice inference is derived from the assumption that the Horn-set for the singular indefinite contains the numeral *two* as an alternative.[8]

(42)　Alt(some$_{sing}$)= {some, two}

With this set, after the first layer of exhaustification, we get the inference that a student went to Ankara or Istanbul and either no student went to Ankara or no student went to Istanbul.[9] This inference is incompatible with the free-choice interpretation. A further layer of will not change the situation. This leads to a prediction in the context of complex disjunction in the scope of a singular indefinite. Since the meaning is not strengthened after the second layer of exhaustification, *exh(exh(∃x(Px ∨Qx)))* will not rule out *exh(∃x(Px ∨Qx))*. If so, what we predict is that *ya…ya…* will be acceptable in the context of singular indefinites (assuming no other factor plays a role). This is a good prediction.

(43)　Bir öğrenci ya Ankara'ya　ya İstanbul'a　gitti
　　　　a　　student or Ankara-to　or İstanbul-to go.PAST
　　　　'A student went to Ankara or Istanbul'

The analysis developed in this paper finds a way to trace the acceptability differences between the existential quantifier and the existential modal back to the availability of free-choice inferences associated with them.[10]

---

exh(◊(p∨q)) explains the ungrammaticality of the natural language object *exh(exh(◊(p∨q)))*. The natural language object *exh(exh(◊(p∨q)))* and the competitor *exh(exh(◊(p∨q)))* are two distinct entities.

[8] See Klinedinst (2007) for an alternative. The exact account is not crucial here.

[9] Calculations can be found in Fox (2007).

[10] There is another prediction here. Namely, *exh(∃x(Px ∨Qx))* will block *(∃x (exh (Px ∨Qx))*. Consider the following logical forms (the first one is calculated in Fox,2007)

(iii)　　exh(∃x(Px ∨Qx)) ⊨ (∃x(Px ∨Qx) & ~(∃xPx &∃xQx))

(iv)　　(∃x (exh (Px ∨Qx)) ↔ ∃x (Px ∨̲Qx))

The first proposition is stronger. If *(∃x (exh (Px ∨ Qx))* is ruled out, then that it should not be possible to continue the sentence in (43) one with a sentence that corresponds to the proposition *(∃x (Px & Qx))*.

(v)　　Bir öğrenci ya Ankara'ya ya İstanbul'a gitti. Bir öğrenci Ankara'ya ve İstanbul'a gitti.
　　　　'A student went to Ankara or Istanbul. A student went to both Ankara and Istanbul'

3. The complex conjunction in the context of negated universal modal

In this section, we test another prediction of the current proposal. This time, we will be concerned with the properties of simple conjunction *ve* 'and' and complex conjunction *hem…hem…* in Turkish. Simple and complex here, again, refers to the number of the morphemes involved in the expression of the conjunction.

(44) a. Ekin dondurma ve çikolatalı puding yemedi
     Ekin ice-cream and chocolate pudding eat.NEG.PAST.3sg
     'Ekin ate neither ice cream nor chocolate pudding'

   b. Ekin hem dondurma hem çikolatalı puding yemedi
     Ekin and ice-cream and chocolate pudding eat.NEG.PAST.3sg
     'It is not the case that Ekin ate both ice cream and chocolate pudding'

What we observe is that the sentence in (44)a is strengtened from *~(p&q)* to *~p&~q*. The same is not true with the second sentence. The total meaning associated with the second sentence is *~(p&q)&(p∨q)*. Fox (2007) has argued that the wide scope conjunctive readings of the negated universal modals is comparable formally to the free-choice effects associated with the existential modals (see also Szabolcsi and Haddican 2004 for a discussion in the context of *homogeneity*).

(45) Ekin'in odasını toplamasına ve bulaşıkları yıkamasına gerek yok
     Ekin room.her clean and dish.PLU wash necessary NEG
     'Ekin is not required to clean her room and she is not required to wash the dishes'
     Basic Meaning: $\sim\Box(p\&q) \leftrightarrow \Diamond\sim(p\&q) \leftrightarrow \Diamond(\sim p\vee\sim q)$
     Free Choice-type inference: $\Diamond\sim p \& \Diamond\sim q$

More specifically, Fox has shown that the free-choice reading can be derived with the second layer exhaustification.[11] We observe that the complex conjunction does not give this free-choice type reading.

(46) Ekin'in hem odasını toplamasına hem bulaşıkları yıkamasına
     Ekin and room.her clean and dish.PLU wash
     'Ekin is not required to clean her room and she is not required to wash the dishes'

     gerek yok
     necessary NEG
     Basic Meaning: $\sim\Box(p\&q) \leftrightarrow \Diamond\sim(p\&q) \leftrightarrow \Diamond(\sim p\vee\sim q)$
     Inference: $\Box(p\vee q)$

There is something weird about this continuation although calling it unacceptable does not seem right. I point to two complications. There is strong specificity effect (a specific student) associated with the existential quantifier in an argument position. I find it hard to avoid the effect and it is unclear to me what role it plays in the calculation of the entailments. Secondly, assume that Spector's analysis of PPI-hood is general. That is, for any PPI, there is the requirement that it occurs in the scope of an *exh* operator. If so, the logical form, *(∃x (exh (Px ∨Qx))*, might be ruled out on independent grounds given that *bir* in Turkish and *some* in English are PPIs.

[11] for calculations, see Fox, 2007. I avoid the repetition of Fox's paper here.

If we make the assumption that we have made for complex disjunction, the absence of the homogeneity-type effects with the complex conjunction follows immediately.

(47)   *Condition on Turkish complex conjunction*
        Hem...hem... must occur in the scope of an *exh* operator.

We have a problem, though. We expect the complex conjunction expressions to be unacceptable (in the same way as the complex disjunction expressions are). This is not true. The sentences with the complex conjunction are not judged to be unacceptable. However, they clearly have a *metalinguistic* feel to them. Indeed, an approximation to the meaning of sentence in (46) is that *one cannot/should not say that Ekin must both clean the room and wash the dishes*. To explain this anomaly, I stipulate a condition of the following kind:

(48)   *Violations of Constraints on the Distribution of exh*
        An ungrammatical sentence S, in which an economy contraint is not obeyed, is
        associated with one of the following judgments:
        i.   The speaker judges S to be unacceptable
        ii. The speaker judges S to be metalinguistic

Perhaps, we observe in English that PPIs gives rise to metalinguistic feel rather than unacceptability in the context of negation. Szabolsci (2004) notes that even though the determiner *some* appears to be a PPI, it is licensed under a clausemate negation if it is in a denial context.

(49)   He found something.
        Wrong! He DIDn't / DID NOT find something.   not > some

PPIs are assumed to be unacceptable in the scope of negation; however, it seems that it is possible to get a metalinguistic feel rather than unacceptability.

4. Conclusion

Chierchia et al. (2012) suggests that the dispreference for deriving implicatures in Downward Entailing environments can be traced back to some type of Strong Meaning Hypothesis, discussed by various scholars (Dalrymple et al. 1998, Beck 2001 a.o). They consider two possibilities for how such a statement should be formulated: a strong version and a weak version. The proposal made in this paper for the Strong Meaning Hypothesis has the property that it is stronger than the weak version but weaker than the strong version.

References

Alonso-Ovalle, Luis. 2006. Disjunction in Alternative Semantics. PhD Dissertation, UMass Amherst.

Beck, Sigrid. 2001. Reciprocals are definites. *Natural Language Semantics*. 9: 69 – 138.

Chierchia, Gennaro, Danny Fox and Benjamin Spector (2012). Scalar implicature as a grammatical phenomenon. In *Semantics: An International Handbook of Natural Language Meaning*, ed. by C. Maienborn, P. Portner and K. von Heusinger. 3: 2297 – 2331. De Gruyter.

Fox, Danny.2007. Free choice and the theory of scalar implicatures. In *Presupposition and implicature in compositional semantics*, ed. by U. Sauerland and P. Stateva, 71-120. New York: Palgrave Macmillan.

Gazdar, Gerald.1979.*Pragmatics: Implicature, Presupposition and Logical Form*. New York: Academic Press.

Klinedinst, Nathan 2007. Plurals, Possibilities and Conjunctive Disjunctions. *UCL Working Papers in Linguistics.*

Sauerland, Uli. 2004. Scalar implicatures in complex sentences. *Linguistics and Philosophy*. 27: 367 – 391.

Spector, Benjamin. 2014.Global Positive Polarity Items and Obligatory Exhaustivity. *Semantics and Pragmatics*.7:11. 1-61.

Szabolcsi, Anna. 2004. Positive Polarity – Negative Polarity. *Natural Language and Linguistic Theory*. 22, 409 – 452.

Szabolcsi, Anna and Bill Haddican. 2004. Conjunction meets negation: a study in cross linguistic variation. *Journal of Semantics*. 21: 219 – 249.

İsa Kerem Bayırlı
isakerem@gmail.com

Feature domains and lexically conditioned harmony in Turkish

Jennifer Bellik

UC Santa Cruz

1.   Introduction

Walker (2012) describes vowel harmony as follows: "In vowel harmony, the vowels in a domain, such as the word, systematically agree, or 'harmonize', in some phonological property." Intuition suggests that this domain-level harmony be represented with a domain-level specification of the shared phonological property. This intuition is formalized in feature-spreading representations of vowel harmony, such as auto-segmentalism (e.g., Clements and Sezer 1982), as well as analyses situated in Optimality Theory (Prince and Smolensky 1993) that employ constraints such as ALIGN (Kirchner 1993, among others) or SPREAD (Kimper 2011).

Domain-level feature representations are formalized even more directly in Optimal Domains Theory (Cole and Kisseberth 1994), Span Theory (McCarthy 2004, O'Keefe 2007), and Smolensky's (2006) headed feature domains. In these theories, Gen constructs feature domains which incorporate multiple segments. A feature domain takes its value from its head (some segment in the domain); all the segments in a domain must realize the domain's feature value.

Though feature-spreading and feature-domains both use featural structures that transcend the segment, these larger structures still depend on an individual segment: the head of the domain, or the anchor for the spreading feature. This segment-dependence predicts that the value of harmony can always be traced to the value of a particular segment or segments. This paper examines a class of Turkish words which violate this prediction. The words of interest require front suffixes, and yet they contain no segment that can be a plausible trigger for front harmony. Some examples appear in (1).

(1)   *Turkish front harmony without a front trigger segment*

	Nom.		Dative		Gloss
a.	dikkat	→	dikkat-e,	*dikkat-a	'attention'
b.	harf	→	harf-e,	*harf-a	'letter'
c.	rab	→	rabb-e,	*rabb-a	'God'

Unlike the rest of the Turkish harmony system, the class of words exemplified in (1) has received little attention. Previously unnoticed is the fact that the vowel failing to trigger harmony is always /a/. There is no word like *hurf-ler* to correspond to *harf-ler*. The specialness of /a/ for this group of words is surprising given that /a/ is a normal participant in harmony for the vast majority of the Turkish lexicon. That is, /a/ only behaves as transparent in a small group of roots. I refer to this behavior as **lexically conditioned transparency**.

This paper proposes an analysis of lexically conditioned transparency using FEATURE DOMAIN THEORY (FDT). FDT extends Smolensky's (2006) headed feature domains, but innovates by positing feature domains that are specified in the *input* to phonology and are *independent* of individual segments. This extension connects FDT to gesturally-grounded models of phonology like Articulatory Phonology (Browman and Goldstein 1993), and enables FDT to provide a unified account of harmonic, disharmonic, and exceptional roots in Turkish. The paper is organized as follows: §0 describes the regular pattern of Turkish backness harmony, then presents the problem and the previous autosegmental account; §0 presents an analysis of lexically conditioned harmony using Feature Domain Theory; §4 provides an account of the specialness of /a/ in Turkish lexically conditioned transparency; and §5 concludes.

2.   Harmony and disharmony in Turkish

All eight vowels in the Turkish vowel inventory participate in backness harmony, alternating with the vowel that they match in for [high] and [round]. Suffixes harmonize in backness with the nearest root vowel, whether the vowels in the root are front as in (2) a, back as in (2) b, or both as in (2) c.

(2)   *Normal backness harmony (plural suffix /lEr/ is realized as [ler] or [lar])[1]:*
    a.   gündüz + lEr → gündüz-ler   'daytimes'
    b.   ayı   + lEr → ayı-lar   'bears'
    c.   kitap   + lEr → kitap-lar   'books'

The topic of this paper is a group of roots (3) in which a final /a/ fails to trigger back harmony—instead, these roots requires front suffixes, and also epenthetic front vowels where epenthesis is required (e.g. (3) j,k). The /a/ thus behaves as transparent, even though elsewhere in the lexicon /a/ triggers back harmony.

(3)   *Roots that fail to trigger harmony*
       Nom.   Dative   Gloss
    a.   infilʲak   infilʲakʲ-e,   infilak-a   'explosion'
    b.   helʲak   helʲakʲ-e,   *helʲakʲ-a   'ruin'
    c.   emlʲak   emlʲakʲ-e,   *emlʲakʲ-a   'real estate'
    d.   istimlʲak   istimlʲakʲ-e,   *istimlʲakʲ-a   'confiscation'
    e.   dikkat   dikkat-e,   *dikkat-a   'attention'

---

[1] I employ Turkish orthography throughout this paper. Turkish orthographic symbols mostly coincide with IPA symbols, but there are a few differences: ö = /ø/, ü = /y/, ı = /ɯ/, c = /dʒ/, ç = /tʃ/, ş = /ʃ/.

f.	menfaat	menfaat-e	*menfaat-a	'interest'
g.	beraat	beraat-e	*beraat-a	'acquittal'
h.	lʲiya:kat	lʲiya:kat-e	*lʲiya:kat-a	'merit'
i.	saat	saat-e,	*saat-a	'hour'
j.	kabir	kabr-e,	*kabr-a	'tomb'
k.	rahim	rahm-e,	*rahm-a	'womb'
l.	harf	harf-e,	*harf-a	'letter'
m.	rab	rabb-e,	*rabb-a	'God'

Several dozen words, mainly ending in /t/, /k/ or a cluster containing /r/, exhibit the same behavior as the examples above. In every case, the final vowel in the word is /a/.

The existence of these back roots that require front suffixes is noted in Lewis (1967) and Underhill (1976) as an exception to vowel harmony. Both note that all such words are borrowings from Arabic, and ascribe the requirement for front suffixes to the properties of the Arabic consonants involved. This is plausible as a diachronic explanation but does not explain the synchronic representation of the exceptional words, since modern Turkish does not exhibit a phonetic distinction between front and back consonants (except for /k g l/). Lewis (1967) implies that these words are exceptional because they are unassimilated—in modern terms, they belong to a distinct cophonology. However, no independent property picks out the words in which /a/ is transparent, making a cophonologies account is unprincipled.

The previous analysis of these roots is due to C&S, who model harmony as the spreading of [+/- back] features from underlyingly specified ("opaque") segments to underlyingly unspecified segments, according to an association convention (left-to-right, one to one, …). Directly extending this representation to consonants, they say that any consonant can be underlyingly marked as [-back], and then trigger front harmony on suffixes. This analysis is problematic because it implies that any consonant in Turkish can occur in three varieties: [-back], [+back], or unspecified, even though for consonants other than /l k g/, these three variants always sound exactly the same. And for all consonants but /l/, all three varieties sound exactly the same *at the end of a word*, which is precisely the environment where the variety needs to make a difference. In addition, though disharmonically front suffixes can occur following a wide variety of consonants, they occur after only one vowel: /a/. Under the consonant-driven analysis, this remains a suspicious accident.

The fact that all the exceptional words have /a/ for their last vowel suggests an approach in which the /a/ itself has a [-back] that triggers front suffixes. But since /a/ in these words is not audibly more front than in other words, an /a/-based approach shares the problems of the consonant-based account, albeit with only one segment occurring in [+back] and [-back] variants that are phonetically indistinguishable.

This shortcoming is not specific to an autosegmental approach. As discussed in the introduction, current theories of vowel harmony always require a segment as a harmony trigger. But the exceptional cases discussed here lack an audibly front segment to trigger the required front harmony on suffixes. In some cases, a front segment appears earlier in the word, but the most local trigger is back (*dikkat*); in other cases, there is no front segment in the word at all (*harf*). Consequently, any account that tries to pin the frontness

of the suffixes to a particular segment in the word will lack phonetic grounding. To avoid this pitfall, a segment-independent model of harmony is necessary.

3.   Proposal: Feature Domain Theory

Nothing in the surface phonology uniquely picks out the words that are exceptions to harmony in Turkish. Instead, a theory that accounts for the exceptional behavior of these roots must make reference to their underlying representations. However, these posited underlying representations should not make claims about segmental feature make-up that are not reflected in surface phonology.

An account of these exceptions to vowel harmony should capture the intuition that something about these particular *words* requires them to take front suffixes – not a characteristic of any of their *segments* in isolation. This intuition accords with the observation that harmony is a characteristic of the word: the requirement of harmony is that all the vowels in a *word* should agree in backness. Evidence from psycholinguistic (Harrison and Kaun 2001, Kabak et al. 2010) and articulatory (Boyce 1990) studies converges with the evidence of epenthesis and the distribution of vowels in the Turkish lexicon (Kabak and Weber 2013) to support the idea that harmony in Turkish acts as a word-level phenomenon. The theoretical representation of harmony should capture this by marking harmony at the word level, while at the same time remaining flexible enough to represent disharmonic words.

Domain-based theories of harmony couched in Optimality Theory meet these criteria. An example is Smolensky (2006), who represents harmony using headed featural domains. The basic idea is that featural domains are constructed by Gen, producing representations like those in (4). A feature domain's value (e.g., [+back] or [-back]) is determined by its head, a segment in it (underlined in (4)). Segments in a domain are required to realize that domain's feature value. Harmony is largely driven by the markedness constraint *HEAD. Since each domain must have a head, every domain incurs a *HEAD violation, and the number of domains per word is minimized in the optimal candidate.

(4)   *Turkish harmony with headed feature domains*
      a.   harmonic root: violates *Head minimally
            ayı +lEr   → (a̱yı-lar)
      b.   disharmonic root: violates *Head multiple times
            kitap+lEr   → (ki̱)(ta̱plar)
      c.   embedded domain: violates *Embed
            dikkat +lEr   → (di̱kk(a̱)t-ler)

For Smolensky (2006), transparency results when segments occur in embedded feature domains, as in (4) c. A segment in an embedded domain realizes the feature value of the innermost domain. Embedding is driven by markedness considerations, since embedding violates *EMBED but can enable a larger domain to satisfy ALIGN without any segments in it violating featural co-occurrence constraints. However, the constraint ranking which is needed to obtain the transparent structure predicts that /a/ will always be transparent

and will not have a front counterpart. In reality, though, /a/ is normally opaque (like all Turkish vowels), and alternates systematically with /e/ in suffixes.

The divergent behavior of *kitap* (which has the vowel sequence *i-a* and requires back suffixes) and *dikkat* (which has the same vowel sequence but requires front suffixes) indicates the need for either a different input structure or a different constraint ranking for these two types of words. A different constraint ranking is only available if *kitap* and *dikkat* belong to different co-phonologies, but this proposal seems suspect, as both words are loans from Arabic, there are no differences in stress assignement or clear segmental cues, etc. Therefore, a difference in input structure must account for the difference in output behavior.

In existing domain-based theories, however, feature domains are not present in the input. I therefore extend Smolensky's theory to use *headless* feature domains—independent of segments and present in the *input* to phonology. I term this approach FEATURE DOMAIN THEORY (FDT)[2].

(5)   ***Feature Domain Theory***
       For a feature F,
       a.   every value of F is represented in phonology as an F-domain, in both the input and the output.
       b.   F-harmony is the requirement that F-domains coincide with prosodic categories larger than the syllable—typically the word.
       c.   Segments inherit the F-value of the F-domain that contains them. When one domain is embedded in another, the deepest F-domain's F-value is realized in the output.

The central idea of FDT is that EVERY FEATURE IS A FEATURAL DOMAIN. Membership in an F-domain *replaces* segmental featural specification in a traditional theory of features. In a language without harmony, every feature is still a feature domain, but it happens to be a domain that coincides with a segment. So in a language like English, FDT's feature representation is largely isomorphic to traditional featural specifications. Meanwhile, in a language exhibiting harmony along some feature, the domains for that feature will regularly span entire words.

As will be shown in §3.2, encoding feature domains in the input to phonology enables faithfulness to drive transparency in *harf-ler* and *dikkat-ler,* so that the full range of harmonic behavior in Turkish is accounted for with a single constraint ranking.

3.1   The regular pattern of Turkish backness harmony in FDT

For Turkish backness harmony, the relevant F-domains are backness domains. In a harmonic root, such as *ayı* 'bear', there is only one underlying backness domain. An alternating suffix like the plural marker *-lEr* has no backness domain of its own. But all segments must be fully specified in the output: SPECIFY (6) is undominated. When the input contains underspecified material like *lEr*, then, either an existing feature domain can be expanded, or a new feature domain can be inserted. Expansion of an existing

---

[2] Thanks to Junko Ito for suggesting this name.

domain results in harmony and incurs a violation *EXPAND (7); insertion of a new domain could produce disharmony, and violates DEP-FD. Since suffixes harmonize, DEP-FD must outrank *EXPAND.

Also, since harmony applies within roots, we need a constraint that penalizes the presence of multiple backness domains, *FEATURE DOMAIN(BACK) (9). Every surface form violates *FD, though, so *FD must also be dominated by SPECIFY, for the ranking in (10).

(6)   SPECIFY: Every segment must be specified in the output of phonology.

(7)   *EXPAND: Don't expand a backness domain.

(8)   DEP-FD(back): Assign a violation for every backness domain in the output that has no correspondent in the input.

(9)   *F-DOMAIN(back) = *FD: Assign a violation for every backness domain.

(10)   SPECIFY, DEP-FD >> *FD, *EXPAND

(11)   *Incorporating a suffix into the existing F-domain avoids violating DEP.*

	/(EyI)$_B$ +lEr/	SPECIFY	DEP(FD)	*EXPAND	*FD
☞ a.	(ayı-lar)$_B$			***	*
b.	(ayı)$_B$ (-lar)$_B$		*!		**
c.	(ayı)$_B$ -lEr	*!			*

In a disharmonic root like *kitap* 'book', there are two contrasting backness domains in the input: (kI)$_F$ (tEp)$_B$. Since disharmonic roots surface faithfully, faithfulness to underlying featural domains must outweigh the pressure to harmonize within roots – *FD must be dominated by MAX-FD: "Assign a violation for every FD in the input that has no correspondent in the output." If *FD were to outrank MAX-FD, all words would be harmonic, because extra backness domains would be deleted.

When multiple domains are present, underspecified suffixes are incorporated into the later harmonic domain, since this minimizes violations of *EXPAND.

(12)   *Suffixes are incorporated into the nearest harmonic domain*

	/(kI)$_F$ (tEp)$_B$ +lEr/	SPECIFY	DEP(FD)	MAX(FD)	*FD	*EXPAND
☞ a.	(ki)$_F$ (taplar)$_B$				**	***
b.	(ki(tap)$_B$ ler)$_F$				**	****!**
c.	(kitepler)$_F$			*!		****!**

3.2 Apparent harmony failure as lexically-conditioned transparency

On the surface, *dikkat* looks like it should have the same structure as *kitap*: a front domain containing the front vowel, followed by a back domain containing /a/. However, this structure predicts that *dikkat* will behave like *kitap* and take back suffixes. Rather, *dikkat* requires an input containing an embedded domain. Such an input structure will allow FAITH to drive transparency.

The failure to phonetically realize a feature due to conflicting demands of multiple governing featural domains is anticipated and penalized by *EMBED (13) (Smolensky 2006). *EMBED is outranked by faithfulness constraints MAX(FD), DEP(FD), and *CONTRACT (14).

(13)   *EMBED: Assign a violation for every segment contained in multiple domains.

(14)   *CONTRACT: Don't remove segments from a featural domain.

*EMBED is violated by the /a/ in *harf* or *dikkat,* revealing that it must be ranked below the faithfulness constraints that preserve embedded featural domains.

(15)   *FAITH(FD) >> *EMBED*

	/(dIk(kEt)_B)_F /	SPECIFY	DEP(FD)	MAX(FD)	*CONTRACT	*EMBED	*EXPAND	*FD
☞ a.	(dik(kat)_B)_F					***		**
b.	(dik)_F (kat)_B				*!**			**
c.	(dikket)_F			*!				*

Suffixes must be incorporated into the nearest featural domain, not an embedded featural domain, both because of *EXPAND and because of *EMBED. Pulling a suffix into an embedded domain, as in (16)b, results in twice the number of *EXPAND violations as pulling it into the outermost domain as in (16)a.

(16)   *Expand prevents suffixes from being pulled into embedded domains.*

	/(dIk(kEt)_B)_F +E/	SPECIFY	FAITH(FD)	*EMBED	*EXPAND	*FD
☞ a.	(dik(kat)_B e)_F			***	*	**
b.	(dik(kata)_B)_F			****!	**	**

The losing candidate (dik(kata)_B)_F is harmonically bounded by the winning candidate, showing that embedding will always result in disharmony between the embedded vowel and any suffixes.

Introducing underlying embedded domains, then, allows the transparency of /a/ in certain Turkish words to fall out naturally from the previously defined constraint set. FDT can account for *harf-ler* and *dikkat-ler* where other domain-based theories could not, because in FDT, domains themselves are targets for faithfulness since they are present in the input, not just constructed by Gen.

This same constraint set and input representation also accounts for the apparently disharmonic epenthesis in words like *kabir ~ kabr-e* 'tomb (+dat)', but I do not present the tableaux here for reasons of space. It remains to be explained, however, is why only /a/ is able to be transparent in Turkish—the task of the next section.

4   Transparent /a/ and its cross-linguistic implications

Formally, the generalization that lexically-conditioned harmony always involves an embedded /a/ can be accounted for simply with the constraint *EMBED[-LOW] in (17).

(17)   *EMBED[-LOW]: Assign a violation for every segment specified [-low] that contained in two backness domains.

(18)   *EMBED[-LOW] >> *EMBED

The ranking in (18) captures the fatal markedness of embedding a high or mid vowel in backness domains in Turkish. Since *EMBED[-low] is unviolated, we can conclude that it is undominated. This ranking yields the desired result that only the low vowel /a/ is able to occur in embedded backness domains, as illustrated in (19), because in Turkish, the only transparent vowel is /a/, and the only low vowel is also /a/.

(19)   *Tableau for hypothetical /(h(U)ʙrf)ꜰ/*

/(h(U)_B rf)_F /	*EMBED[-low]	MAX[round]?	FAITH(BD)	*EMBED	*EXPAND	*FD	IDENT[low]	MAX[round]?
☞ a. (h(a)_B rf)_F						**	*	*
☞ b. (hürf)_F			*!			*		
c. (h(u)_B rf)_B	*!			*		**		

If *EMBED[-low] and MAX(round) outrank MAX(Backness Domain), an input embedded domain containing a high vowel will be deleted in the output, as shown in (19)b. This incurs a Max(FD) violation but produces a harmonic, entirely front word. On the other hand, if *EMBED[-low] does not dominate FAITH(Backness Domain) but does dominate IDENT constraints governing height and rounding, then the embedded vowel would instead be changed to become a low vowel to avoid violating *EMBED[-low]. This is shown by candidate (a). Thus, adding *EMBED[-LOW] as an undominated constraint in the Turkish constraint system enables FDT to capture the generalization that only /a/ is allowed to be transparent in Turkish.

There are both percepual and articulatory reasons for *EMBED(BACK)[-LOW]. Although transparent vowels are non-participants in harmony phonologically and acoustically, they still participate articulatorily to some degree (Gafos and Dye 2011). A transparent vowel must still be recognizable acoustically even when its articulation has been distorted by the surrounding context, and so only perceptually stable vowels (Stevens 1989) may be transparent. Since /a/ is perceptually stable, it is a good candidate for transparency—a perceptual motivation for a constraint like *EMBED(BACK)[-LOW]. In addition, since low vowels are less contrastive for backness than high vowels, it is less taxing articulatorily to produce *i-a-i* than *i-u-i*. Therefore a disharmonic *i-a-i* sequence will be perceptually and articulatorily more similar to a harmonic *i-æ-i* sequence than a disharmonic *i-u-i* sequence will be to a harmonic *i-ü-i* sequence. Lastly, /a/ occurs in a relatively sparse part of the Turkish vowel space, meaning it is less confusable with the other phonemes. /a/ would have to raise as well as front or round in order to be perceived as one of the other phonemes. (Even Turkish [æ] is higher than /a/ and is described by some as a low mid vowel.) In contrast, /i/ would be a poor candidate for transparency in Turkish since there are three vowels that it differs from only by one feature (/ɪ/ in backness, /ü/ in rounding, and /e/ in height).

The constraint *EMBED(BACK)[-LOW] may seem surprising, given that cross-linguistically, /a/ is uncommon as a systematically transparent vowel in backness harmony, while /i/ and /e/ are very common (Kramer 2003). But systematic transparency is driven by markedness and corresponds to gaps in the vowel inventory (Kiparsky and Pajusalu 2003), while lexically-conditioned transparency is driven by faithfulness, so differences in their patterning are not so unexpected.

5.   Conclusion

Motivated by a pattern of lexically conditioned transparency in some Turkish roots, this paper proposes Feature Domain Theory. FDT extends the theory of headed feature domains (Smolensky 2006). Headed feature domains are only targets for markedness constraints, not faithfulness, since they are entirely built by Gen. FDT innovates by proposing underlying, unheaded featural domains. Since FDT's feature domains are present in the input, they are targets for faithfulness as well as markedness. Thus, exceptional transparency as in *harf-ler* can be captured by using a marked structure (embedding) in the input. Normal harmonic processes remain in force. But since FAITH preserves embedding, the suffix does not harmonize with the linearly-adjacent but embedded /a/. Instead, it harmonizes with the outermost domain, which is [-back]. This is lexically-conditioned transparency. This structural analysis accounts for both regular harmony and apparent exceptions using the same constraint ranking and representational mechanisms.

References

Browman, C. P. and L. Goldstein. 1993. Dynamics and Articulatory Phonology. *Haskins Laboratories Status Reports on Speech Research.* SR-113 (51-62).

Boyce, S. E. 1990. Coarticulatory organization for lip rounding in Turkish and English. *Journal of the Acoustic Society of America*: 88 (6).

Clements, G. and E. Sezer. 1982. Vowel and Consonant Disharmony in Turkish. In *The Structure of Phonological Representations.,* ed. by H. van der Hulst and N. Smith. Dodrecht: Foris.

Cole, J. and C. Kisseberth. 1994. An optimal domains theory of harmony. *Studies in the Linguistic Sciences*: 24(2).

Gafos, A. and A. Dye. 2011. Vowel harmony: transparent and opaque vowels. In *Blackwell Companion to Phonology*, ed. by M. van Oostendorp and K. Rice (2164 – 2189), Wiley-Blackwell.

Harrison, K. D. & A. Kaun. 2001. Patterns, pervasive patterns, and feature specification. *Distinctive feature theory* 2, 211.

Inkelas, S., A. Küntay, R. Sprouse & O. Orghun. 2000. Turkish Electronic Living Lexicon (TELL). *Turkic Languages* 4, 253-275.

Kabak, B., K. Maniwa, & N. Kazanina. 2010. Listeners use vowel harmony and word-final stress to spot nonsense words: A study of Turkish and French. *Laboratory Phonology*: 207-224.

Kabak, B. & S. Weber. 2013. Markedness, context and directionality in vowel harmony: A corpus study on vowel co-occurence patterns. *Dilbilim Araştırmaları [Journal of Linguistic Research]*, 53-85.

Kimper, W. 2011. Competing triggers: Transparency and opacity in vowel harmony. University of Massachussetts, Amherst. Dissertation.

Kiparsky, P. and K. Pajusalu. 2003. Towards a typology of disharmony. *The Linguistic Review* 20:217-241.

Kirchner, R. 1993. Turkish vowel harmony and disharmony: an Optimality Theoretic account. *Rutgers Optimality Workshop I*.

Kramer, M. 2003. *Vowel harmony and correspondence theory*. Berlin: Mouton de Gruyter.

McCarthy, J. 2004. Headed spans and autosegmental spreading. *Linguistics Department Faculty Publication Serie*s. Paper 42.

O'Keefe, M. 2007. Transparency in Span Theory. In *University of Massachusetts Occasional Papers in Linguistics 33: Papers in Optimality Theory 3* ed. by L. Bateman, A. Werle, M. O'Keefe, and Ehren Reilly. Amherst, MA: GLSA.

Prince, A. and P. Smolensky. 1993/2002. Optimality Theory: Constraint Interaction in Generative Grammar. *Rutgers Optimality Archive*: ROA 537.

Smolensky, P. & Legendre, G. 2006. *The Harmonic Mind: From Neural Computation To Optimality-Theoretic Grammar. Vol. 2: Linguistic and Philosophical Implications*. MIT Press.

Stevens, K. N. 1989 On the quantal nature of speech. *Journal of Phonetics* 17. 3-46.

Walker, R. 2012. Vowel Harmony in OT. *Language and Linguistics Compass* 6/9: 575– 592.

Jennifer Bellik
jbellik@ucsc.edu

Subject marking and scrambling effects in Balkar nominalizations[*]

Tatiana Bondarenko

Lomonosov Moscow State University

1.   Introduction

In this paper I analyze the nominative/genitive alternation of subject marking in Balkar nominalized clauses (NCs):

(1)   ustaz   sabij-i /   sabij-i-**ni**   alma-sy-n   aSa-Ran-y-n   ešit-ti[1]
        teacher child-3.**NOM/** child-3-**GEN**   apple-3-ACC   eat-NMN-3-ACC   hear-PST
        'The teacher heard that (someone's) child had eaten (someone's) apple.'

Despite the fact that there are no noticeable interpretative differences between the nominalizations with nominative and nominalizations with genitive subjects, there is clear evidence that these nominalizations involve different syntactic structures. One piece of evidence comes from scrambling effects: it is possible to scramble direct objects of nominalizations with nominative subjects over the subject of nominalization (2) or over the matrix subject (3), but this kind of movement is impossible out of nominalizations with genitive subjects (4)-(5).

---

[1] I use subject nouns in the possessive declension throughout the paper, because in the non-possessive declension the case markers of nominative and genitive coincide with accusative: genitive looks the same as marked accusative, and nominative looks the same as unmarked accusative.

(2)   ustaz   alma-sy-n   sabij-i   aSa-Ran-y-n   ešit-ti
      teacher   apple-3-ACC   child-3-**NOM** eat-NMN-3-ACC   hear-PST
      'The teacher heard that it was (someone's) child that had eaten (someone's)
      apple.'

(3)   alma-sy-n   ustaz   sabij-i   aSa-Ran-y-n   ešit-ti
      apple-3-ACC   teacher   child-3-**NOM** eat-NMN-3-ACC   hear-PST
      'The teacher heard that it was (someone's) child that had eaten (someone's)
      apple.'

(4)   * ustaz   alma-sy-n   sabij-i-ni   aSa-Ran-y-n   ešit-ti
      teacher   apple-3-ACC   child-3-**GEN** eat-NMN-3-ACC   hear-PST
      Int. interpretation: 'The teacher heard that it was (someone's) child that had eaten
      (someone's) apple.'

(5)   * alma-sy-n   ustaz   sabij-i-ni   aSa-Ran-y-n   ešit-ti
      apple-3-ACC   teacher   child-3-**GEN** eat-NMN-3-ACC   hear-PST
      Int. interpretation: 'The teacher heard that it was (someone's) child that had eaten
      (someone's) apple.'

In contrast to direct object scrambling, scrambling of subjects is allowed out of both types
of nominalizations[2]:

(6)   sabij-i-**ni**   fatima ata-sy-ny   maSyna-sy-n   buz-Ran-y-n   bil-di
      child-3-**GEN** Fatima father-3-GEN   car-3-ACC   crash-NMN-3-ACC   know-PST
      'Fatima found out that (someone's) child smashed the father's car.'

(7)   dušman-ɨ   ustaz   šaxar-ɨ-n   cac-ʁan-ɨ-n   bil-di
      enemy-3.**NOM** teacher   city-3-ACC   destroy-NMN-3-ACC   know-PST.3SG
      'A teacher found out that (someone's) enemy destroyed (someone's) city.'

The observed correlation between the subject marking in nominalizations and the
extractability of direct objects suggests that there are some syntactic differences in the
make-up of nominalizations with nominative and genitive subjects which underlie the
alternation. This brings up the question of whether it is possible to account for the
scrambling phenomena in Balkar nominalizations under the previously proposed theories
of subject case licensing: subject case licensing dependent on tense (Chomsky 1981.
Lyutikova & Grashchenkov 2008) and subject case licensing dependent on phi-feature
agreement (Chomsky 1995, Kornfilt 1984, Kornfilt 2003, Miyagawa 2011)[3]. In section 2
of this paper I will address the question of whether subject case licensing in Balkar
nominalizations is dependent on the functional categories of the tense-aspect domain, while

---

[2] Both types of scrambling are driven by information structure. Interpretational effects associated with
scrambling are irrelevant for the line of reasoning I develop below and will be ignored throughout the paper.
[3] Whether the subject case is licensed by TAM features or by phi-features seems to be parameterized across
languages (Iatridou 1993, Ura 2000, Lavine & Freidin 2002).

in section 3, I will consider the question of whether the nominative/genitive alternation depends on the properties of agreement between nominalizations and their subjects.

Yet another issue is how other properties of these nominalizations are related to subject case licensing. I will present some properties that the two nominalizations have in common in section 2 while in section 4, I will address several properties with respect to which they differ. In section 4, I will argue that the differences between nominalizations with nominative subjects and nominalizations with genitive subjects in Balkar have a structural source: only the former nominalizations have a CP layer in their functional structure. I argue that this single structural difference explains all the different properties the two nominalizations exhibit: subject case, scrambling possibilities, number agreement, wh-words licensing and binding effects. In my analysis of Balkar nominalizations, I suggest that C (when it is present) licenses the nominative case on the subject, while D (in the absence of C) licenses the genitive case. In other words, while some languages have been shown to employ D-licensing (for example, Dagur see Hale 2002) or C-licensing (for example, Turkish see Kornfilt 1984, Kornfilt 2003) in the embedded contexts, Balkar can employ both.[4]

2. Does licensing of NOM/GEN depend on the functional categories of the tense-aspect domain?

There have been different views with respect to whether the functional categories of the tense-aspect domain play a role in licensing subject case (Chomsky 1981, Kornfilt 2003, Gülsat Aygen 2004). It is evident that in Balkar nominalizations tense is not related to licensing subject case, because nominalized clauses with nominative subjects and with genitive subjects don't differ in their (impoverished) tense specification. Both nominalizations have the same two allomorphs: 'ryR' for future interpretation, 'Ran' for non-future interpretation.

(8) ol   bala-sy/bala-sy-ny   **busaRatda**   alma-sy-n   aSa-**Ran**-y-n
     That child-3.NOM/child-GEN **now**   apple-3-ACC   eat-**NMN**-3-ACC
     ajt-a-dy
     say-IPFV-3SG
     'He is saying that (someone's) child is eating (someone's) apple now.'

(9) ol   bala-sy/bala-sy-ny   **tUnene**   alma-sy-n   aSa-**Ran**-y-n
     that child-3.NOM/child-GEN **yesterday**   apple-3-ACC   eat-**NMN**-3-ACC
     ajt-a-dy
     say-IPFV-3SG
     'He is saying that (someone's) child has eaten (someone's) apple yesterday.'

(10) * ol   bala-sy/bala-sy-ny   **tambla**   alma-sy-n   aSa-**Ran**-y-n
     that child-3.NOM/child-3-GEN **tomorrow**   apple-3-ACC   eat-**NMN**-3-ACC
     ajt-xan-dy
     say-PFCT-3SG

---

[4] Miyagawa (2011) proposes a similar analysis of nominative/genitive alternation in Japanese relative clauses ("ga/no Conversion"): C is responsible for nominative case licensing, and D – for genitive.

(11)   ol   bala-sy/bala-sy-ny   **tambla**   alma-sy-n   aSa-**ryR**-y-n
        that   child-3.NOM/child-3-GEN   **tomorrow** apple-3-ACC eat-**NMN**-3-ACC
        ajt-xan-dy
        say-PFCT-3SG
        'He has said that (someone's) child will eat (someone's) apple tomorrow.'

In (8)-(9) it is shown that nominalizations with the 'Ran' suffix are compatible with interpretations where the event described by the nominalization co-occurs or precedes the event described by the matrix verb. The ungrammaticality of (10) indicates that these nominalizations cannot describe an event that follows the event described by the matrix verb, while (11) shows that this interpretation obtains in nominalizations with the 'ryR' suffix. As we see from (8)-(11), there are no differences between nominalizations with nominative subjects and nominalizations with genitive subjects with respect to availability of temporal interpretations; therefore, subject case licensing is not dependent on tense specification of the nominalized clause.

The two nominalizations have a number of other properties in common. They exhibit the full argument structure of the verb and do not display the transitivity restriction (which is present in structures with genitive subjects in Japanese see Harada 1971, Watanabe 1996, Miyagawa 2011): overtly marked accusative object can intervene between the subject and the nominalized verb (1). Both nominalizations can include derivational morphemes; for example, passivized verbs can be nominalized:

(12)   fatima   Uj   iSle-**n**-gen-i-n   bil-di
        Fatima   house.**NOM** build-**PASS**-NMN-3-ACC   know-PST
        'Fatima found out that a house has been built.'

(13)   fatima   alim-ni   Uj-U-**nU**   iSle-**n**-gen-i-n   bil-di
        Fatima   Alim-GEN   house-POSS-**GEN**   build-**PASS**-NMN-3-ACC   know-PST
        'Fatima found out that Alim's house has been built.'

Moreover, both nominalizations can include negation in their functional make-up:

(14)   fatima sabij-i   /sabij-i-**ni**   alma-sy-n   aSa-**ma**-Ran-y-n
        Fatima child-3.**NOM** /child-3-**GEN**   apple-3-ACC eat-**NEG**-NMN-3-ACC
        eSit-ti
        hear-PST
        'Fatima heard that (someone's) child has not eaten (someone's) apple.'

The ability to host negation markers and to co-occur with temporal adverbs ((8)-(9), (11)) suggests that in both nominalizations the amount of structure that is nominalized is not smaller than TP. The only verbal morpheme that cannot occur in nominalizations is the question marker:

(15)   * fatima sabij-ler-i   /sabij-ler-i-**ni**   Saxar-Ra   bar-Ran-**my**-lar-y-n
        Fatima child-PL-3.**NOM** /child-PL-3-**GEN**   city-DAT   go-NMN-**Q**-PL-3-ACC
        sor-du
        ask-PST
        Int. reading: 'Fatima asked whether (someone's) children have gone to the city.'

In this section I have shown that nominalizations with nominative subjects and nominalizations with genitive subjects share a number of properties and are at least TP-level nominalizations. Most importantly, though, subject case marking of these nominalizations does not hinge upon their tense specification.

3. Does licensing of NOM/GEN depend on agreement?

A natural question then arises whether subject case licensing could be related to the differences in phi-feature agreement between subjects and corresponding nominalizations. At first sight, there does not seem to be any differences between the agreement with nominative and with genitive subjects: in both cases nominalizations display the same possessive markers that nouns in a possessive construction take[5] (cf. (16) and (18), (17) and (19)).

(16)   taryx-**ym**
        story-**1SG**
        'my story'

(17)   bala-**Ryz**-Ra
        child-**2PL**-DAT
        'to your (pl) child'

---

[5] Whether "nominal" agreement markers (i) are completely distinct from verbal ones is an unsettled question. There are two verbal agreement paradigms in Balkar (ii); the 1st set of markers is used in the present and future tenses, perfect and habitual, while the 2nd one is used in the past tense, in conditionals and in imperative forms. Verbal agreement with 3 person subjects is optional both in terms of person agreement and number agreement.

    i.   *Nominal agreement markers*

	Sg	Pl
1	-(y)m	-(y)byz
2	-(y)N	-(y)Ryz / -(i)giz
3	-(s)y	

    ii.   *Verbal agreement markers*

1st set of verbal agreement markers			2nd set of verbal agreement markers		
	Sg	Pl		Sg	Pl
1	-ma	-byz	1	-(y)m	-(y)q
2	-sa	-syz	2	-(y)N	-(y)Ryz / -(i)giz
3	-(dy)	-(dy)-(la)	3	-(dy)	-(dy)-(la)

As can be observed from the tables ((i)-(ii)), 1Sg and 2 person markers of the nominal paradigm are identical to the markers of the 2nd set of verbal agreement markers, while 1Pl marker of the nominal paradigm is identical to the corresponding marker of the 1st set of verbal markers. The only marker in the nominal paradigm that does not have a verbal counterpart is 3rd person marker ('(s)y'). Thus, it might be the case that agreement patterns of Balkar are best analyzed not in terms of verbal – vs – nominal split, but in some other manner.

(18)   annja   men   /men-**ni**   basta-ny
        mother   1SG.**NOM**   /1SG-**GEN**   porridge-ACC
        aSa-Ran-**ym**-my   bil-di
        eat-NMN-**1SG**-ACC   know-PST
        'The mother found out that I have eaten porridge.'

(19)   appa   siz   /siz-**ni**   basta-ny
        grandfather   2PL.**NOM**   /2PL-**GEN**   porridge-ACC
        aSa-Ran-**yRyz**-Ra   sejirsin-di
        eat-NMN-**2PL**-DAT   be.surprised-PST
        'The grandfather was surprised that you (pl) have eaten porridge.'

But while person agreement[6] is the same in both nominalizations and is identical to the agreement in the nominal paradigm, there are peculiar differences in number agreement which are reflected in the use of third person plural marker 'lar'. This marker occurs both on nouns and on verbs with the following difference: it conveys an interpretable number feature in the former case ((20)-(23)) and an uninterpretable number feature in the latter ((25)-(28)). I assume that nouns always bear an interpretable number feature; while the plural value of this feature is realized overtly by 'lar', the singular value has a null realization. I also assume that agreement between possessors and possessees is obligatory in Balkar: whenever a possessor is present, the possessee agrees with it. However, there is only one available morphological slot for realization of number features: the structure in (24), where one instance of lar denotes the plurality of a possessee while the other one is an agreement morphology triggered by the plurality of the possessor, is unavailable in Balkar, and, as far as I know, in other Turkic languages. Therefore, the two compete for the same morphological slot. When the plural marker occurs on possessees as in ((21),(23)), it denotes plurality of possessees and is never interpreted as denoting plurality of possessors. I suggest that this is the case because of the competition of the two number features, the interpretable number of the possessee and the uninterpretable number of the possessor for a single slot. The interpretable one wins and gets realized, as in (23).

(20)   sabij-i-ni   kitab   *SG possessor, SG possessee*
        child-3-GEN   book-3
        'child's book'

(21)   sabij-i-ni   kitab-**lar**-y   *SG possessor, PL possessee*
        child-3-GEN   book-**PL**-3
        'child's books'

(22)   sabij-**ler**-i-ni   kitab-y   *PL possessor, SG possessee*
        child-**PL**-3-GEN   book-3
        'children's book'

---

[6] Here by "person agreement" I mean agreement markers which are specified for person features (including cumulative 1PL and 2PL markers).

(23) sabij-**ler**-i-ni   kitab-**lar**-y   *PL possessor, PL possessee*
      child-**PL**-3-GEN   book-**PL**-3
      'children's books' /*'children's book'

(24) * sabij-**ler**-i-ni   kitab-**lar-lar**-y   *PL possessor, PL possessee*
      child-**PL**-3-GEN   book-**PL-PL**-3
      'children's books'

The plural marker ('lar') on verbs is a realization of an uninterpretable number feature (it does not denote plurality of events, (26)) which is valued through the Agree operation with the subject ((25)-(28)). Crucially, this type of agreement is optional: it can be absent on verb even when the subject displays plural marking (27).

(25) sabij-i   tau-Ra   Orle-gen-di   *SG subject, SG verb*
      child-3.NOM mountain-DAT   climb-NMN-3SG
      '(Someone's) child climbed the mountain.'

(26) * sabij-i   tau-Ra   Orle-gen-di-**le**   *SG subject, PL verb*
      child-3.NOM mountain-DAT   climb-NMN-3SG-**PL**
      Int. reading: '(Someone's) child climbed the mountain several times.'

(27) sabij-**ler**-i   tau-Ra   Orle-gen-di   *PL subject, SG verb*
      child-**PL**-3.NOM mountain-DAT   climb-NMN-3SG
      '(Someone's) children climbed the mountain.'

(28) sabij-**ler**-i   tau-Ra   Orle-gen-di-**le**   *PL subject, PL verb*
      child-**PL**-3.NOM mountain-DAT   climb-NMN-3SG-**PL**
      '(Someone's) children climbed the mountain.'

Clausal nominalizations are not expected to have interpretable number features (Alexiadou et al. 2010)[7], and this is borne out by both Balkar nominalizations: they cannot denote plurality of events when they receive plural marking ((30), (34)). But the two nominalizations differ with respect to the optionality of number agreement. The number agreement ('lar') in nominalizations with nominative subjects follows the verbal pattern in ((25)-(28)): the uninterpretable number feature (number of the nominalization's subject) is optionally realized on the nominalization ((29)-(32)):

(29) men   a-ny   sabij-i   tau-Ra   Orle-gen-i-n
      1SG   that-GEN   child-3.NOM mountain-DAT   climb-NMN-3-ACC
      bil-di-m
      know-PST-1SG
      'I found out that his child climbed the mountain.'

---

[7] Alexiadou et al. (2010) observe that nominalizations which have verbal projections above AspP cannot have interpretable number: it is not possible for them to denote plurality of events.

(30) * men   a-ny   sabij-i   tau-Ra   Orle-gen-**ler**-i-n
        1SG   that-GEN   child-3.NOM   mountain-DAT climb-NMN-**PL**-3-ACC
        bil-di-m
        know-PST-1SG
        Int. reading: 'I found out that his child climbed the mountain several times.'

(31)   men   a-ny   sabij-**ler**-i   tau-Ra   Orle-gen-i-n
        1SG   that-GEN   child-**PL**-3.NOM   mountain-DAT climb-NMN-3-ACC
        bil-di-m
        know-PST-1SG
        'I found out that his children climbed the mountain.'

(32)   men   a-ny   sabij-**ler**-i   tau-Ra   Orle-gen-**ler**-i-n
        1SG   that-GEN   child-**PL**-3.NOM   mountain-DAT climb-NMN-**PL**-3-ACC
        bil-di-m
        know-PST-1SG
        'I found out that his children climbed the mountain.'

The number agreement in nominalizations with genitive subjects ((33)-(36)) shows a different pattern: the number feature on these nominalizations is still uninterpretable (34), but agreement is obligatory and the agreement marker cannot be omitted (35).

(33)   ustaz sabij-i-ni   bir zat da   oqu-ma-Ran-y-n   bil-di
        teacher child-3-GEN   one.thing.NOT read-NEG-NMN-3-ACC   know-PST
        'The teacher found out that (someone's) child hasn't read anything.'

(34) * ustaz sabij-i-ni   bir zat da   oqu-ma-Ran-**lar**-y-n   bil-di
        teacher child-3-GEN   one.thing.NOT read-NEG-NMN-**PL**-3-ACC know-PST

(35) * ustaz sabij-**ler**-i-ni   bir zat da   oqu-ma-Ran-y-n   bil-di
        teacher child-**PL**-3-GEN   one.thing.NOT   read-NEG-NMN-3-ACC know-PST

(36)   ustaz sabij-**ler**-i-ni   bir zat da   oqu-ma-Ran-**lar**-y-n   bil-di
        teacher child-**PL**-3-GEN one.thing.NOT read-NEG-NMN-**PL**-3-ACC know-PST
        'The teacher found out that (someone's) children haven't read anything.'

What might be the reason for the difference in number agreement between the two nominalizations? I propose that nominalizations with nominative subjects display verbal number agreement, while nominalizations with genitive subjects exhibit nominal number agreement which is the same as in (20)-(23). As suggested above, in nominal possessor constructions agreement is obligatory, but the morphological component chooses to realize interpretable number features of possessees instead of the uninterpretable features of possessors. Since clausal nominalizations do not have interpretable number features, we expect uninterpretable number features to realize if the agreement is nominal. And that is exactly what we see in nominalizations with genitive subjects ((33)-(36)). On the other hand, number agreement between subjects and verbs is optional (27), and we expect to see

the same optionality of subject-nominalization agreement if it is verbal. As I have demonstrated (31), this expectation is borne out for nominalizations with nominative subjects.

In this section I have shown that the two nominalizations differ with respect to agreement with their subjects: while nominalizations with nominative subjects display verbal number agreement, nominalizations with genitive subjects exhibit nominal number agreement with a minor reservation that they do not have their own interpretable number. However, this difference on its own is not sufficient to explain the scrambling possibilities we observed in the first section: different agreement will not derive the scrambling effects if the left periphery of two nominalized clauses is identical. In the next section I will relate the difference in number agreement to scrambling effects and other properties which distinguish the two nominalizations from each other.

4.   C-licensing – vs – D-licensing, agreement and scrambling effects

The idea that phi-feature agreement occurs at a phase level (CP, vP, DP[8]), (Chomsky 2005, Boeckx 2003, Miyagawa 2005), naturally explains why D and C become the locus of subject case licensing (Miyagawa 2011): the phi-features are only introduced into the derivation by these phase. Both D-licensing (Saito 1983, Hale 2002, Miyagawa 1993, Miyagawa 2008, Miyagawa 2011) and C-licensing (Watanabe 1996, Hiraiwa 2001, Kornfilt 2003, Kornfilt 2008) have been proposed as the mechanisms of licensing subjects in embedded contexts; and C-licensing has been proposed for licensing both nominative and genitive subjects (Watanabe 1996, Hiraiwa 2001). In this section I will argue that Balkar employs C-licensing in nominalizations with nominative subjects and D-licensing in nominalizations with genitive subjects, and that this difference can explain the other properties in which the two nominalizations differ.

I propose that nominalizations with nominative subjects are nominalized CPs, while nominalizations with genitive subjects are DPs without the C projection in their functional make-up (D immediately takes TP as its complement). This is supported by several similarities between CPs and nominalizations with nominative subjects on one hand, and between DP[9]s and nominalizations with genitive subjects on the other hand. First, scrambling of direct objects out of CPs ((36)-(37)) is possible just like in nominalizations with nominative subjects ((2)-(3)), and unlike nominalizations with genitive subjects ((4)-(5)):

(37)   ustaz   **a-ny**   **surat-y-n**   sabij-i   iSle-gen-di   dep
        teacher **that-GEN picture-3-ACC** child-3.NOM   draw-NMN-3S   COMP
        ojla-dy
        think-PST
        'The teacher thought that it was (someone's) CHILD that drew his picture.'

---

[8] I assume that DP is also a phase. See (Chomsky 2001, Svenonius 2004, Dikken 2007, Ott 2009) for the relevant discussion.
[9] I assume the Universal-DP Hypothesis that implies that DPs are present in all languages, but that not all noun phrases are DPs within the languages (Pereltsvaig 2007, Lyutikova & Pereltsvaig 2015).

(38) **a-ny   surat-y-n**   ustaz   sabij-i   iSle-gen-di   dep
      that-GEN picture-3-ACC teacher child-3.NOM   draw-NMN-3SG   COMP
      ojla-dy
      think-PST
      'The teacher thought that it was (someone's) CHILD that drew his picture.'

Nominalizations with genitive subjects behave like DPs with respect to scrambling. No
material can be extracted from DPs apart from the genitive possessor ((39)-(40)); the
genitive possessor can be scrambled out and, for example, become separated from the
possessee noun by clausal material, e.g., by the temporal adverb of the clause (41).

(39) * sejir   sabij-i-ni   kitab-y
       interesting   child-3-GEN   book-3.NOM
       Intended reading: 'child's interesting book'

(40)   sabij-i-ni   sejir   kitab-y
       child-3-GEN   interesting   book-3.NOM
       'child's interesting book'

(41)   alim sabij-ler-i-ni   **tUnene**   kitab-y-n   tas   et-gen-di
       Alim child-PL-3-GEN   **yesterday** book-3-ACC   lost   do-PFCT-3SG
       'Alim lost the children's book yesterday.'

The pattern we observe with DPs is identical to that of nominalizations with genitive
subjects ((4)-(6)): only genitive subjects can be extracted out of nominalizations (6); no
other material can scramble out ((4)-(5)).
     Second, both nominalizations with nominative subjects and CPs can host wh-words in
their structure which are licensed by the question marker on the matrix predicate[10]; attempts
to license wh-words inside nominalizations with genitive subjects lead to ungrammaticality
(44).

(42)   ustaz   sabij-i   **ne-ni**   buz-du   dep   bil-di-mi
       teacher child-3.NOM   **what-ACC**   break-PST   COMP   know-PST-Q
       'About what thing did the teacher find out that (someone's) child broke it?'

(43)   ustaz   sabij-i   **ne-ni**   buz-Ran-y-n   bil-di-mi
       teacher child-3.NOM   **what-ACC**   break-NMN-3-ACC   know-PST-Q
       'About what thing did the teacher find out that (someone's) child broke it?'

(44) * ustaz   sabij-i-ni   **ne-ni**   buz-Ran-y-n   bil-di-mi
       teacher child-3.GEN   **what-ACC**   break-NMN-3-ACC   know-PST-Q

---

[10] This fact itself supports my proposal that nominalizations with nominative subjects have a CP layer in their
structure under the assumptions that wh-words undergo movement at LF (Huang 1982), that noun phrases
are islands for extraction (see (39)) and that wh-words move through the specifiers of C projections: if
nominalizations with nominative subjects did not have a CP layer, wh-words in (43) wouldn't be able to
move out of the nominalization and get interpreted in the matrix clause. Thus, the ungrammaticality of (44)
is predicted under my analysis which suggests that nominalizations with genitive subjects do not involve CP.

Third, both CPs and nominalizations with nominative subjects constitute the same binding domain with respect to binding of reflexives, which is different from the binding domain of nominalizations with genitive subjects. There is a reflexive pronoun 'kesi' which can participate in long-distance biding relations and be bound by the subject of a matrix clause. While CPs and nominalizations with nominative subjects are transparent for the binding of the simple reflexive 'kesi' (it can be bound by the subject of the matrix clause), ((45)-(46)), nominalizations with genitive subjects do not allow the material inside them to participate in long-distance binding relations (the simple reflexive cannot be bound by the matrix subject) ((47)):

(45) ustaz soxta-sy   kesi-n   iSle-gen-di   dep   ojla-dy
      teacher student-3.**NOM** self-ACC draw-NMN-3SG COMP think-PST
      'The teacher$_i$ thought that his student$_k$ drew himself$_k$.' – OK
      'The teacher$_i$ thought that his student$_k$ drew himself$_i$.' – **OK**

(46) ustaz soxta-sy   kesi-n   iSle-gen-i-n   ojla-dy
      teacher student-3.**NOM** self-ACC draw-NMN-3-ACC   think-PST
      'The teacher$_i$ thought that his student$_k$ drew himself$_k$.' – OK
      'The teacher$_i$ thought that his student$_k$ drew himself$_i$.' – **OK**

(47) ustaz soxta-sy-ny   kesi-n   iSle-gen-i-n   ojla-dy
      teacher student-3-**GEN** self-ACC draw-NMN-3-ACC   think-PST
      'The teacher$_i$ thought that his student$_k$ drew himself$_k$.' – OK
      'The teacher$_i$ thought that his student$_k$ drew himself$_i$.' – *****

In other words, nominalizations with nominative subjects exhibit CP-like behavior with respect to reflexives' binding, while nominalizations with genitive subjects are different in this respect.

There is further evidence which suggests that nominalizations with genitive subjects are DP[11]s: they can contain demonstratives:

(48) ustaz **a-ny**   soxta-sy-ny   maSina-sy-n buz-Ran-y-n
      teacher **that-GEN** student-3-GEN   car-3-ACC   break-NMN-3-ACC
      bil-di
      know-PST
      'The teacher found out about that breaking of the car by the student.'

In (48) the demonstrative 'any' can be interpreted as a modifier of the whole nominalization[12], which I take as evidence for the presence of D projection in the structure of the nominalization.

So far we have seen that nominalizations with nominative subjects are CP-like with respect to such properties as direct object scrambling possibilities, ability to host wh-

---

[11] The question of whether nominalizations with nominative subjects are DPs or NPs requires further investigation.
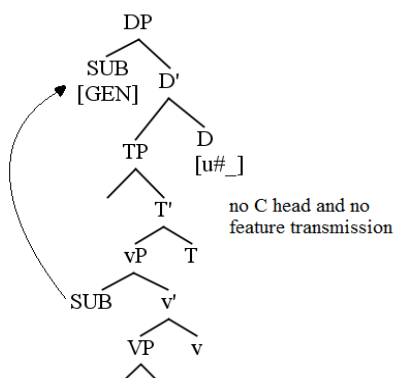[12] It can also be interpreted as a modifier of the noun phrase "student" ('soxtasyny').

phrases and reflexives that can enter into long-distance binding relations, while
nominalizations with genitive subjects behave in a different manner and are DP-like with
respect to scrambling and the ability to host demonstratives. These differences between the
two nominalizations can be naturally connected with the difference in number agreement
under the proposed analysis, which suggests that only nominalizations with nominative
subjects have a CP layer in their structure. If it is C that introduces verbal phi-features into
the derivation (by being a phase head), then it is expected that verbal agreement is possible
only when C is present in the structure, and is impossible otherwise. However, the C under
consideration is a special case, because it is nominalized. I propose that in configurations
with a nominalized C some of its phi-features can be overridden by the phi-features of the
nominalizing head. In case of the Balkar nominalization, the person feature is the one being
overridden: as we have seen in section 3, in nominalizations with nominative subjects
person agreement seems to be nominal. But crucially, the presence of the verbal number
agreement in nominalizations with nominative subjects reveals that the C projection is
present in the structure and that its number feature is retained and transmitted to T
(Chomsky 2008), which can then act as a probe and attract the subject of nominalization
to its specifier for the purposes of agreement, valuing its case feature nominative. The
proposed structure for nominalizations with nominative subjects is presented in (49).[13]

(49)   *Nominalizations with nominative subjects*

---
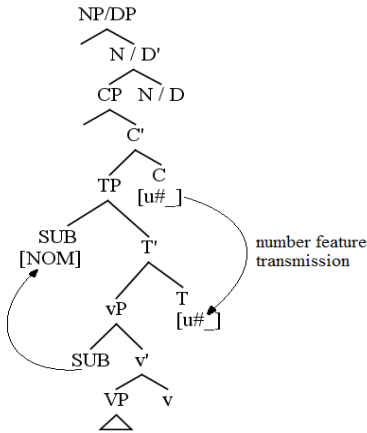
[13] I will leave unresolved the issue of whether nominalizations with nominative subjects are NPs or DPs,
because I have no evidence that would argue in favor of one possibility over another. These nominalizations
could be also thought of as KPs (see Bittner & Hale).

(50)   *Nominalizations with genitive subjects*

```
                    DP
                   /  \
                SUB    D'
               [GEN]  /  \
                    TP    D
                   /  \  [u#_]
                  T'      no C head and no
                 /  \     feature transmission
               vP    T
              /  \
           SUB    v'
                 /  \
               VP    v
               /\
```

I argue that in nominalizations with genitive subjects there is no C projection. This immediately explains the absence of any verbal agreement between the nominalization and its subject: since C introduces phi-features into the derivation, there is no source of verbal phi-features in the absence of C. Consequently, T does not have phi-features, cannot act as a probe, enter into the Agree relation with the subject and value its case feature nominative. The subject with an unvalued case feature remains available for further interactions and enters into the Agree relation with D. This results in nominalization displaying the nominal agreement and in valuing subject's case feature genitive. In this way, the existence/ absence of CP in the functional make-up of nominalizations determines whether verbal agreement can or cannot be present[14].

Scrambling properties of the two nominalization directly follow from the proposed structures. In nominalizations with a CP layer (49) nominative subjects remain in the Spec, TP position. Thus, these nominalizations contain an escape hatch (Spec, CP – A'-position) through which direct objects can be scrambled into the matrix clause (3). Direct objects can also stop directly in the Spec, CP position of the nominalized clause (2). There is no such escape hatch in the functional make-up of nominalizations with genitive subjects: DP is a phase, and Spec, DP is an A-position to which subjects move in order to get their case valued. Hence, under the proposed analysis it is predicted that scrambling of direct objects to the left periphery of nominalizations with genitive subjects and into the matrix clause should be impossible: there is no position through which such movements would take place. Note that the possibility of genitive subjects' scrambling out of such nominalizations is also predicted: being in the specifier of a phase (DP), these subjects can move out to the left periphery of the matrix clause (6) just like genitive possessors can be extracted outside noun phrases (41).

---

[14] I will not address the issue of how person agreement works in Balkar nominalizations (see footnote 4). With the stipulation that it does not require the movement of nominative subjects to positions higher than Spec, TP, it does not affect my analysis.

5.   Conclusion

In this paper I proposed an analysis of nominative/genitive alternation in Balkar nominalized clauses. I have argued that nominalizations with nominative subjects and nominalizations with genitive subjects project distinct syntactic structures: only the former have a CP layer in their functional make-up. I have shown that this single difference in the functional structure of the two nominalizations can derive various discrepancies between them: scrambling possibilities and nature of number agreement. Besides, under the CP-analysis the parallelism between nominalizations with nominative subjects and finite CPs as to their ability to host wh-words and to allow of long-distance binding of reflexives comes out with no effort at all. There have been proposed two types of subject licensing operative in different languages: C-licensing and D-licensing. In this paper I have argued that subjects of Balkar nominalized clauses employ both mechanisms: C licenses nominative subjects, while D licenses genitive ones.

<div align="center">References</div>

Alexiadou, A., Iordachioaia, G., Soare, E. 2010. Number/Aspect Interactions in the Syntax of Nominalizations: A Distributed Morphology Approach. *Journal of Linguistics* 46: 537-574.

Bittner, M., Hale, K. 1996. The structural determination of case and agreement. *Linguistic Inquiry* 27.1:1-68.

Boeckx, C. 2003. *Islands and Chains: Resumption as Stranding*. John Benjamins Publishing Company

Chomsky, N. 1981. *Lectures on Government and Binding*. Foris Publications, Dordrecht.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.

Chomsky, N. 2001. Derivation by phase. In *Ken Hale: A Life in Language*. ed by Kenstowicz, M., MIT Press, Cambridge, Mass., pp. 1–52.

Chomsky, N., 2005.Three factors in language design. *Linguistic Inquiry* 36, 1-23.

Chomsky, N. 2008. On Phases. In *Foundational issues in linguistic theory: essays in honor of Jean-Roger Vergnaud.* ed. by R. Freidin et al.., MIT Press, Cambridge, Mass.

den Dikken, Marcel. 2007. Phase extension: Contours of a theory of the role of head movement in phrasal extraction. *Theoretical Linguistics* 33: 133–163.

Gülsat, A. 2004. Finitness, Case and Clausal Architecture. *MIT Occasional Papers in Linguistics* 23.

Hale, K. 2002. On the Dagur object relative: some comparative notes. *Journal of East Asian Linguistics* 11, 109-122.

Harada, S.-I. 1971. Ga-no conversion and ideolectal variations in Japanese. *Gengo Kenkyu* 60, 25-38.

Hiraiwa, K. 2001. On nominative-genitive conversion. In *MIT Working Papers in Linguistics 39:A Few from Building E39* ed. by Guerzoni, E., and O. Matushansky. MITWPL: Cambridge

Huang, C.-T. J. 1982. Logical relations in Chinese and the theory of grammar. Ph.D thesis, MIT.

Iatridou, S. 1993. On Nominative Case Assignment and a few related things. *MIT Working Papers in Linguistics* 19:175-198.

Lyutikova, E. and Grashchenkov, P. 2008. Nominalizacija i semantiko-sintaksičeskij interfejs [Nominalization and the Semantic-Syntax Interface]. In *Issledovanija po glagol'noj derivacii [Studies of verb derivation],* ed. by V.A. Plungjan and S.G. Tatevosov. Moscow: Jazyki slavjanskix kul'tur.

Lyutikova, E. and Pereltsvaig, A. 2015. The Tatar DP. *The Canadian Journal of Linguistics/La revue canadienne de linguistique,* 60/3, 289-325.

Kornfilt, J. 1984. Case Marking, Agreement, and Empty Categories in Turkish. Ph.D.dissertation, Harvard University.

Kornfilt, J. 2003. Subject case in Turkish nominalized clauses. In *Syntactic Structures and Morphological Information*. ed. by Junghanns, U., Szucsich, L.Mouton de Gruyter, Berlin/New York, 129-215.

Kornfilt, J. 2008. Subject case and Agr in two types of Turkic RCs. In *Proceedings of Workshop on Altaic Formal Linguistics (WAFL4)*, ed. by Ulutas, S.,Boeckx, C. (Eds.), MITWPL56, Cambridge, Mass., 145-168.

Lavine, J and Freidin, R. 2002. The subject of defective Tense in Slavic. *Journal of Slavic linguistics* 18: 1, 101–130.

Miyagawa, S. 1993. Case-checking and minimal link condition. In *MIT Working Papers in Linguistics 19: Papers on Case and Agreement II*. ed. by Phillips, C., MITWPL, Cambridge, Mass, 213-254.

Miyagawa, S. 2005. On the EPP. In *Proceedings of the EPP/Phase Workshop*. ed. by McGinnis, M., Richards, N. MITWPL, Cambridge, Mass. 201–236.

Miyagawa, S. 2008. Genitive Subjects in Altaic. In *Proceedings of the Workshop on Altaic Formal Linguistics (WAFL4),* ed. by Ulutas, S., Boeckx, C. MITWPL 56, Cambridge, Mass. 198.

Miyagawa, S. 2011. Genitive subjects in Altaic and Specification of Phase. *Lingua* special volume ed. by Jaklin Kornfilt and John Whitman. Vol 121, 1265-1282.

Ott, Dennis. 2009. The conceptual necessity of phases: Some remarks on the minimalist enterprise. In *Explorations of phase theory: interpretation at the interfaces*, ed. by Kleanthes K. Grohmann. Mouton de Gruyter, Berlin, 253-277.

Pereltsvaig, A. 2007. On the universality of DP: A view from Russian. *Studia Linguistica* 61(1): 59-94.

Saito, M. 1983. Case and government in Japanese. *Proceedings of the Second West Coast Conference on Formal Linguistics*. CSLI, Stanford, pp. 247-259.

Svenonius, P. 2004. On the edge. In *Syntactic Edges and Their Effects* ed. by Adger, D., de Cat C., Tsoulas, G. Kluwer, Dordrecht, pp. 261–287.

Ura, H. 2000. *Checking theory and grammatical functions in universal grammar*. Oxford University Press.

Watanabe, A. 1996. Nominative-genitive conversion and agreement in Japanese: A crosslinguistic perspective. *Journal of East Asian Linguistics* 5, 373-410.

Tatiana Bondarenko
t.i.bond@yandex.ru

Auxiliaries in North Azeri and some related issues[*]

Colin P. Davis

Massachusetts Institute of Technology

1.   Introduction

In this paper, I analyze the distribution and form of auxiliary *be* in North Azeri (Turkic), based on which I go onto discuss a few related issues. Centrally, I demonstrate that auxiliary patterns in North Azeri provide evidence for a theory like that expressed in Bjorkman (2011), in which auxiliary *be* is inserted to realize features that are structurally stranded from a verbal element. For North Azeri, I show that when certain projections intervene between V^0 and T^0, auxiliary *be* is inserted to support T^0. However, due to substantial allomorphy of copulas to the features of T^0, the morpho-syntactic regularity of auxiliary insertion in North Azeri is not straightforwardly surface-apparent.

(1-3) below give some preliminary data. In all these examples we see aspect marking on the verb, and, I argue, some form of *be* supporting tense morphology:

(1) *Past tense auxiliary*
män gat͡ʃ-ad͡ʒaɣ [i] -di-m
1SG run-PROS [BE] -PST-1SG
'I was going to run.'

(2) *Present tense auxiliary*
sän jat-ad͡ʒaɣ- [Ø] -Ø-san
2SG sleep-PROS- [BE] -PRES-2SG
'You (sg.) are going to sleep.'

(3) *Nominalized clause auxiliary*
[o-nun   je-jir   [ol] -duɣ-u   ] halva
[3SG-GEN eat-IMPV [BE] -NFUT-3SG] halva
'the halva that he/she is/was eating'

We see three allomorphs of the copula in the above examples (*i, ol, -∅*) and these copular forms have further complications that render the underlying morpho-syntactic consistency of the language unclear on the surface. I argue, however, that when such factors are controlled for, the underlying principles at work are regular. Namely, whenever intervening heads like Asp(ect)^0 interrupt V^0 to T^0 movement, T^0 is supported by auxiliary *be*.

I go on to argue that the facts about copular allomorphy in North Azeri provide evidence for a theory in which allomorphic conditioning is structurally constrained. (Bobaljik 2012). That is, the North Azeri copula only suppletes when structurally local to a T^0 bearing relevant features. Allomorphic conditioning cannot be triggered over unbounded distance.

Finally, I focus on one particular case of auxiliary *be*. In the first part of the paper, I argue that when the null present tense T^0 requires auxiliary support, it takes an auxiliary *be* that is itself silent, as in (2) above. What does it mean for an auxiliary, intuitively a support element, to be phonologically null, and support an element that is itself null? I suggest that this configuration is evidence that auxiliary *be* is not reducible to a requirement of the PF interface. The auxiliary is arguably not motivated by syntax proper or LF either. Not being reducible to pure syntax or the interfaces, auxiliary insertion acts like an operation endemic to a separate morphological component of the grammar (Harley & Noyer 1991), providing independent evidence for such a component.

1.1   Background on North Azeri

First, some background. Being Turkic, North Azeri is head-final, has *pro*-drop, and has a fair degree of agglutinative morphology. The North Azeri verbal complex hosts aspect, modality, tense, and agreement morphology:

(4)   män gat͡ʃ-ir-∅-di-m
       1SG run-IMPV-BE-PAST-1SG
       'I was running.'

I assume that the order of the units of verbal inflection reflects a hierarchy of corresponding functional projections. (Baker 1985) I claim the clausal structure in (5) for North Azeri. While I do not literally take aspect and modality to compete for the same syntactic position, treating these as occupying the same position is sufficient for the analysis at hand.

(5)   [[[V^0] Asp(ect)/Mod(al)^0] T^0]

Though I assume at least v/voiceP dominating VP, I abstract away from this in what follows, using V^0 to refer to a complex head of at least V^0+v^0. While all clauses contain tense morphology, aspect/modality morphology is optional. For simplicity I assume that Asp/ModP is absent when no corresponding morphology is present. The table in (6) below shows the basic tense/aspect/modality morphology of North Azeri:[1]

---

[1]Distributional facts actually suggest a dedicated projection for the perfect, between Asp/Mod^0 and T^0. Also, I gloss over the fact that perfects are not aspects nor modals, but these details are not relevant to the

(6) *North Azeri TAM*[2]

Aspect/Modality	Tense
-(j)Ar = Epistemic modal (*may/probably*), habitual	*-∅* = Present
-mAlI = Epistemic/deontic modal (*must/should*)	*-DI* = Past
-(j)Ir = Imperfective	*-DIG* = Non-future
-mIʃ/-(j)Ib = Perfect	
-(j)Adʒ͡AG = Prospective (*going to, intend to*)	

Main clauses and non-nominalized embedded clauses use the present or past tense, while nominalized embedded clauses (NCs) use only the non-future tense *-DIG*. I hypothesize that *-DIG* represents a defective T^0 (Chomsky 2001, Miyagawa 2012) which is featurally and thus semantically impoverished. Like those of Turkish, North Azeri NCs have genitive subjects and use agreement morphology from the nominal paradigm, otherwise used in possessive constructions where a possessum agrees with the genitive possessor. The details of NCs are not important here, though a few will appear throughout the paper.[3]

With the background out of the way, in the next section I get into the details of the auxiliary data, which also entails a discussion of copulas in North Azeri generally.

2.   Copulas, auxiliary and otherwise

In this section, I survey a number of configurations in North Azeri which I argue represent a very regular morpho-syntax, which morpho-phonological idiosyncrasies obscure.

First, I show that whenever an Asp/ModP sits between VP and TP, an auxiliary copula is inserted to host tense morphology. A straightforward auxiliary context is NCs, where when Asp/Mod⁰ is present, a copular form *ol* arrives to support tense (*-duɣ*), as in (7-8):

(7) [o-nun   je-**miʃ** *(**ol**)-**duɣ**-u   ] halva
     [3SG-GEN eat-**PRF**   **BE-NFUT**-3SG] halva
     'the halva that he/she has/had eaten'

(8) [o-nun   je-**jädʒ͡äj** *(**ol**)-**duɣ**-u   ] halva
     [3SG-GEN eat-**PROS**   **BE-NFUT**-3SG] halva
     'the halva that he/she is/was going to eat'

While non-nominalized clauses do not require an auxiliary copula, they optionally use the copular stem *i* as an auxiliary in the past tense when there is also Asp/Mod⁰, as in (9-10).

---

points of the paper. The *-(j)Ib* allomorph of the perfect is triggered by non-1st person subjects. (Authier 2010) The morpheme *-mIʃ/-(j)Ib* which I've labeled the perfect also has an inferential evidential reading, a function typical of the Turkish equivalent *-mIş*. (Göksel & Kerslake 2005) In North Azeri, the evidential reading is quite marked, the perfect reading having become the default, as Schönig (2006) and Öztopçu (2003) observe.

[2]Following the terminology of the Turkish equivalent, I refer to *-(j)Ar* as the 'aorist' (AOR) though this label does not really match the semantics involved.

[3]See Davis (2015) for more on North Azeri NCs.

(9)   män halva je-**jir**   **(i)**-**di**-m
       1SG halva eat-**IMPV** **BE**-**PST**-1SG
       'I was eating halva.'

(10)   män uʃaɣ ol-anda halva je-**jär**   **(i)**-**di**-m
       1SG child BE-when halva eat-**AOR** **BE**-**PST**-1SG
       'When I was a child I used to eat halva.'

The auxiliaries *ol* and *i* have the same distribution in their respective contexts. Both are ungrammatical if there is only tense, and no Asp/Mod^0 morphology, as in (11-12):

(11)   [män-im je (*****ol**)-duɣ-um] halva
       [1SG-GEN eat **BE**-NFUT-1SG] halva
       'the halva that I eat/ate'[4]

(12)   män je (*****i**)-di-m
       1SG eat **BE**-PST-1SG
       'I ate.'

   Why is auxiliary *i* optional while auxiliary *ol* is obligatory? Looking at standard past tense copular constructions, we see two copular allomorphs in free variation, *i* and *-j*. The *-j* variant deletes when following a consonant, which we can see when comparing copular constructions with vowel final and consonant final predicates, as in (13) and (14). In the former, there is no reason to delete the copula, which surfaces as *i* or *-j*. In the latter, the consonant-final predicate creates conditions that will result in *-j* deleting, but *i* which is not subject to those deletion conditions can persist on the surface.[5]

(13)   män xästä i/-j/*-∅-di-m   (14)   män d͡ʒözäl i/*-j/-∅-di-m
       1SG sick BE-PST-1SG   1SG pretty BE-PST-1SG
       'I was sick.'   'I was pretty.'

   With one exception, Asp/Mod^0 morphemes in North Azeri end in a consonant. As an auxiliary supporting tense morphology occurs just to the right of Asp/Mod^0 marking, that auxiliary will be adjacent to any final consonant of the Asp/Mod^0 morpheme. Therefore in such contexts if there is an auxiliary *-j*, it will delete. The exception is the vowel-final modal *-mAlI*, where as expected an auxiliary *-j* does not delete, as in (15):

(15)   sän d͡ʒet-mäli-[**j**]-di-n
       2SG go-MOD-**BE**-PAST-2SG
       'You (sg.) needed to go.'

---

[4]This example in its grammatical form would in fact be pronounced *je-dij-im* due to vowel harmony.
[5]Kelepir (2001) makes a similar claim of the Turkish copula *-y-*, which I have drawn from in analyzing North Azeri.

In summary, the auxiliary *i* in past tense clauses merely appears optional because it has a free variant *-j* which, for phonological reasons, usually deletes in auxiliary contexts. So far, I've demonstrated the presence of an auxiliary in past tense non-nominalized clauses as well as NCs. In the next section, I argue for the presence of auxiliaries in present tense clauses as well, the other sort of non-nominalized clause.

2.1   The silent present tense auxiliary

The evidence for this silent auxiliary comes from stress placement. As a point of comparison, Turkish typically has word-final stress, but Kornfilt (1996) argues that when a word contains a copular morpheme, stress shifts to precede the copula. Similarly, stress in North Azeri is also typically word-final, as (16) shows with a simple past tense verb form with no Asp/Mod^0 morphology.[6]

(16)   jat-ˈ**dı-n**
         sleep-PST-2SG
         'You (sg.) slept.'

However, in single-word copular verbal complexes, as in (17-18), stress precedes the copula, shifting onto the predicate. In (17) we see that the copula is zero in the present tense, as is cross-linguistically common. (Kornfilt 1996, Payne 1997) In (18) we see an overt copula *-j* which we've seen before in past tense contexts.

(17)   sän tälä'**bä**-∅-∅-sän   (18)   o   xäs'**tä**-j-di-∅
         2SG student-BE-PRES-2SG   3SG sick-BE-PST-3SG
         'You (sg.) are a student.'   'He/she was sick.'

A parallel fact is that in present tense verb forms with Asp/Mod^0 like (19-20), stress shifts onto the Asp/Mod^0 morphology, unlike the word-final stress in verbal complexes without such morphology like (16). I infer that the present tense zero copula which we saw in (17) appears as an auxiliary in such verb forms, supporting the silent present tense, and shifting stress onto Asp/Mod^0, as that morpheme directly precedes the copula:

(19)   ja'**t-ır**-∅-∅-am   (20)   sän jat-a'**dʒaɣ**-∅-∅-san
         sleep-IMPV-BE-PRES-1SG   2SG sleep-PROS-BE-PRES-2SG
         'I sleep/am sleeping.'   'You (sg.) are going to sleep.'

Following this analysis, the sequence of morphemes in (19-20) is exactly the same as that of auxiliary constructions seen already, in past tense clauses and NCs. In all if we have a verb with Asp/Mod^0 marking, there is an auxiliary *be* carrying tense morphology. It appears that the underlying structure is precisely the same across all the constructions reviewed, then, but details of allomorphy and phonology obscure this underlying regularity. Similar auxiliary phenomena have been discussed for Turkish matrix clauses. (Kornfilt

---

[6]Stress is represented with the IPA convention of the symbol ⟨ ˈ ⟩ preceding a stressed syllable.

1996, Kelepir 2001) However, analogous phenomena in NCs have been analyzed as occurring for different reasons than auxiliaries in non-nominalized clauses. (Göksel 2001, 2003) I claim that in North Azeri, all the auxiliaries seen here are fundamentally the same, only superficially differing in morpho-phonology.

    Further, we've seen a regularity in the environments that trigger certain copular forms. Both auxiliary copulas and copular main verbs are subject to the same rules of phonological exponence, in that they are tense-sensitive. To formalize this regularity, I use the framework of Distributed Morphology (Halle & Marantz 1993, Harley & Noyer 1991) in which syntactic terminals are assigned phonological information post-syntactically, based on rules of contextual Vocabulary Insertion (VI). The following set of VI rules, where the biconditional arrow relates an underlying phonological form to a syntactic context, captures the facts about present and past tense copular forms in North Azeri:

(21)   /i/ $\vee$ /-j/ $\leftrightarrow$ [_, +PST]   (22)   /-Ø/ $\leftrightarrow$ [_, +PRES]

We also saw that in NCs the auxiliary has the form *ol*. In what follows, I argue that *ol* is the elsewhere form of the copula that occurs when neither of the above rules is able to apply.

2.2   An account of the auxiliary

In this section I give an account of auxiliary copulas which will contextualize the discussion of allomorphic locality, and inform the analysis of the present tense auxiliary in the latter part of the paper. In analyzing similar facts about Turkish, Kelepir (2001) claims that when an intervening AspP blocks the typical V^0 to T^0 movement, a copula is inserted at T^0 to check an assumed [Verbal] feature, yielding configurations like (23):

(23)
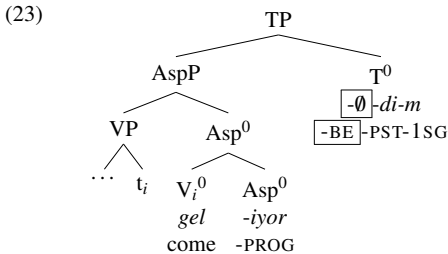
Kelepir's evidence for this interruption of movement in Turkish comes from clausal coordinations with what is termed *suspended affixation*. (Kornfilt 1996, Lewis 1975) Such coordinations have two verbs with Asp/Mod^0 morphology, and one instance of tense scoping over the verbal coordination. North Azeri has these too, as in (24-26):

(24)   [män-im   [[pʃir-**är**   ] vä   [je-**jär**   ]] ol-**duɣ**-um   ] halva
          [1SG-GEN [[cook-**AOR**] and [eat-**AOR**]] BE-**NFUT**-1SG] halva
          'The halva that I probably will/used to cook and eat.'

(25)   män [[pul-u   d͡ʒöhtyr-**är**] vä [o-na   ver-**är**   ]]-∅-**di**-m
          1SG [[money-ACC take-**AOR**   ] and [3SG-DAT give-**AOR**]]-BE-**PST**-1SG
          'I used to take the money and give it to him/her.'

(26)   män [[pul-u   d͡ʒöhtyr-**yr**] vä [o-na   ver-**ir**   ]]-∅-∅-äm
          1SG [[money-ACC take-**IMPV**] and [3SG-DAT give-**IMPV**]]-BE-**PRES**-1SG
          'I am taking the money and giving it to him/her.'

Applying the analysis of Kelepir for Turkish, such constructions involve an Asp/ModP coordination, out of which no V^0 moves high as T^0, because such movement would result in a Coordinate Structure Constraint violation. (CSC, Ross 1967) Rather, no V^0 moves higher than its local Asp/Mod^0. Further head movement of that head-complex to T^0 would need to be Across-The-Board (ATB) in order to not violate the CSC, but such movement is not happening here. ATB movement involves two identical elements in each conjunct moving to a higher position and coalescing as a single element, and in (24-26) we clearly see two separate syntactically low instances of the coordinated verb.

    In contrast, when there is no Asp/Mod^0, suspended affixation is impossible, as is expected if V^0 must move out of VP. (27) demonstrates. If there is no Asp/Mod^0, and the CSC prevents movement of V^0 to T^0, V^0 is stuck in VP if ATB movement is not available. I suggest that this is the cause of the ungrammaticality of (27). Notice in (27) that auxiliary insertion does not save the construction. This shows that ungrammaticality here is not reducible to V^0 failing to move to T^0 for T^0's own needs, but that V^0 itself must move.

(27)   *män [pul-u   **d͡ʒöhtyr**] vä [o-na   **ver**] (i)-di-m
          1SG [money-ACC **take**]   and [3SG-DAT **give**] COP-PST-1SG
          'I took the money and gave it to him/her.'

    I infer that as a general principle in North Azeri V^0 moves to T^0, but Asp/Mod^0 blocks that movement when present. Thus examples like (24-26) do not violate the CSC, as there is no head movement out the coordination. If Asp/Mod^0 blocks V^0 to T^0 movement, and we see auxiliaries at T^0 only when there is Asp/Mod^0, we can conclude that when V^0 does not make it to T^0, auxiliaries appear, as in the schema in (28):

(28)

Bjorkman (2011) argues that auxiliary *be*, which is cross-linguistically common, arises post-syntactically to ensure that a functional head's features are supported by a verbal el-

ement.[7] In other words, inflectional features require the presence of a V^0 element in order to be well-formed. The behavior of North Azeri appears to provide evidence for such an account. When V^0 and T^0 are not local, tense requires a V^0, motivating auxiliary insertion.

2.3   Copular allomorphy and locality

I have shown at the beginning of section 2 that the copular allomorph in the past tense for both copular constructions and auxiliary constructions is *i/-j*, though the latter variant usually deletes in auxiliary contexts. I have also argued in section 2.1 that the copula in both present tense auxiliary constructions and copular constructions is null. We've also seen section 2 that in NCs, the auxiliary has the form *ol*.

A further fact is that copular constructions with Asp/Mod^0 morphology always use the copular form *ol*, as in (29-30). The former is a present tense clause, while the latter is past tense, yet the copular main V^0 does not show allomorphy sensitive to that T^0. Rather we see *ol* in both of these cases, where something intervenes between V^0 and T^0.

(29)   män xästä **ol-muʃ**-∅-∅-am
      1SG sick   BE-**PRF**-BE-PRES-1SG
      'I have been/become sick.'

(30)   män häkim **ol-ad͡ʒaɣ**-∅-di-m
      1SG doctor **be/become**-PROS-BE-PST-1SG
      'I was going to be/become a doctor.'

I argue that this *ol* is the elsewhere form of the copula, which occurs when a copula is not in a context sufficient to trigger its allomorphy. With this hypothesis in place, all the North Azeri facts show that the copula's allomorphy is conditioned by the features of T^0, but under certain conditions that allomorphy fails to apply.

When does copular allomorphy fail? When the copula is not local enough to T^0, I argue. What is the locality condition? Bobaljik (2012) argues that an XP level between two syntactic elements blocks allomorphic conditioning between them, as in (31) where α and β are two syntactic objects in a potential allomorphy relationship:

(31) $\alpha \ldots]_{X^0/^*XP} \ldots \beta$

The distribution of copular allomorphy in North Azeri provides evidence for a proposal of this nature.[8] Based on what I argued above, we expect a copular main V^0 to move to T^0 when there is no Asp/Mod^0 in the structure to interfere with that movement, and therefore in such cases copular allomorphy sensitive to T^0 should occur, and there is no XP level

---

[7]Bjorkman (2011) uses a framework where heads transmit inflectional features downward through a reverse Agree operation. Auxiliary *be* hosts features which fail to agree with and spell out on V^0.

[8]This topic is taken up in greater detail in Davis (2017), along with the puzzle of why in (29-30) there is both a *be* and *become* interpretation.

between V^0 and T^0 within the resulting head-adjunction structure. That allomorphy goes through in such scenarios is confirmed by (13), (14), (17), and (18) above.

I have argued that the presence of Asp/Mod^0 blocks V^0 to T^0 movement, and when this interruption occurs, we predict based on the schema in (31) that a copular main V^0 cannot take allomorphy sensitive to T^0. This is because in such a structure Asp/ModP intervenes between V^0 and T^0, as in (32), where V^0 does not reach T^0:

(32) $[\, [\, [\, \ldots \, t_i \,]_{VP} \, V^0_{[BE]i}\text{-Asp/Mod}^0 \,]_{\underline{Asp/ModP}} \, T^0 \,]_{TP}$

This prediction is borne out in (29-30) above, where we have Asp/ModP intervening between the copular main V^0 and T^0, and V^0 takes the elsewhere form *ol*.

Finally, why does the auxiliary copula undergo allomorphy in present tense clauses like (2) or past tense clauses like (1), and yet the auxiliary has the form *ol* in NCs like (3)? I argue the NC T^0 *-DIG* is an instance of T^0 which is featurally impoverished, resulting in its underspecified interpretation. The lack of specific tense features like [+PAST] or [+PRESENT] also leaves *-DIG* unable to condition copular allomorphy. Therefore any copula in the local context of this T^0 cannot take anything but elsewhere allomorphy. While auxiliaries, as V^0 elements inserted directly at T^0, are inherently local enough to T^0 for allomorphy to occur, if the relevant features are absent, allomorphy fails anyway.

3.   Silent auxiliaries and their motivation

In this section, I discuss the nature of the present tense, and go on to consider the present tense silent auxiliary. I've argued that in North Azeri there is evidence for a phonologically null auxiliary copula supporting the also null present tense. The primary evidence for this came from a parallel with stress irregularities in typical copular constructions. For example, in (17) and (18) we saw present and past tense predicative copular constructions. In the latter, we see an overt copula *j*, and stress shifts to precede this copula. In the former, the copula is null due to the allomorphy facts about North Azeri, yet word stress shifts to a non-final position that is precisely preceding where we expect the copula to be, despite the fact that it is not phonologically overt.

There is also irregular word-medial stress in present tense verb forms with Asp/Mod^0 marking like (19-20). I have argued that the irregular stress in these present tense verb forms with Asp/Mod^0 morphology stems from the presence of a copula in the verbal complex supporting present tense morphology. Present tense morphology is null in this language, as is not cross-linguistically unusual, and furthermore the auxiliary copula supporting that tense morpheme is also null, as this is the copula's realization when local to a present tense T^0. I have argued that this analysis brings all the North Azeri data together under a unified morpho-syntactic analysis. As a general principle, the presence of Asp/ModP prevents V^0 to T^0 movement, resulting in auxiliary *be* insertion at T^0. The irregular stress in present tense examples like (19-20) falls out from this, as the presence of auxiliary copulas within the morphological word results in stress shifts that we independently know to occur in the language, as in copular constructions like (17-18).

The presence of the auxiliary supporting the present tense T^0 entails that there is really a present tense T^0 there in the syntax. Alternatively, we might suppose that there is in fact no present tense T^0, thereby accounting for its nullness in that it could not be pronounced if it is not syntactically present. Under this view, present tense is a default interpretation provided pragmatically when there is no T^0 node. Such an analysis yields a relatively irregular syntax, where root clauses sometimes contain TP and sometimes do not. The analysis I have argued for maintains that TP is always present. If we do not have a present tense T^0, there is no reason to expect an auxiliary copula to ever appear supporting this T^0, and without an auxiliary copula we lose the explanation for the irregular stress in (19-20).

Additionally, look back at the suspended affixation constructions in (24-26). Example (26) has, at least in the interpretation, present tense scoping over two aspect marked verbs. The two verbal conjuncts in (26) which I have argued are an AspP coordination which is the complement of T^0, are not units that could ever stand alone as root clauses in this context where we have a 1st singular subject. While 3rd singular agreement morphology is null in North Azeri, 1st person agreement is not. The verbal conjuncts in (26) are therefore not units that could be well-formed as stand-alone elements in (26), and as such they are presumably coordinated under more structure, containing a head or heads that bear the 1st singular agreement morphology that we see in (26), at the right edge of the construction. If there is no present tense T^0 in the syntax, we might ask what higher head the coordination is the complement of, and what head bears that agreement morphology. Alternatively, if we simply suppose that there really is a present tense T^0, there is nothing curious about (26).

3.1   Null auxiliaries

Having argued that there is a present tense T^0, now I go on to consider what is entailed by the silent auxiliary that sometimes supports this null T^0. Bjorkman (2011) argues that auxiliary *be* is inserted post-syntactically in what are fundamentally well-formed syntactic structures. Namely, while the syntactic structures may be satisfactory in of themselves, there is a morphological well-formedness requirement that inflectional features have a local V^0 element. When no V^0 is sufficiently local, auxiliary *be* is inserted at the stranded feature. Bjorkman argues for this post-syntactic view of auxiliaries, rather than a view where auxiliary verbs are selected in syntax, because of auxiliary patterns that have a character of overflow, rather than selection by any particular functional element.

Bjorkman demonstrates the overflow pattern in languages like Latin, Arabic, and Kinande. Put schematically, given two functional categories F and G, the presence of just F or G does not result in an auxiliary in the overflow pattern. Rather, it is only the combination of F and G that results in an auxiliary. The Latin examples in (33-35) demonstrate this scenario. In (33) the verb has perfect marking, and in (34) passive marking, with no auxiliary in either example. In (35) however, which combines the perfect and the passive, an auxiliary *be* appears:

(33)   Puellae   crustulum   consumpserunt.
        girl-PL.NOM small.pastry-ACC eat-PL.PFV
        'The girls ate the little pastry.'

(34) Crustulum   consumitur.
     small.pastry-NOM eat-PRES.PASS
     'The little pastry is (being) eaten.'

(35) Crustulum   consumptum   **est**.
     small.pastry-NOM eat-PASS.PTCP be.3SG.PRES
     'The little pastry was/has been eaten.'   (Bjorkman 2011, pp. 27)

The intuition here is that no specific element can be singled out that selects for the auxiliary. Rather, the auxiliary is required when the presence of too many functional elements results in some, typically structurally higher, elements being stranded from V^0 such that they require an auxiliary. That is, there is no single functional head that selects an AuxP, rather auxiliaries arise based on the configurations produced by syntax, without reference to what the specific elements of that configuration are.

If this view is correct, in that auxiliary *be* is a post-syntactic repair element not selected for by anything in the syntax proper, the conclusion is that auxiliaries are not relevant to the syntax itself. We might further ask whether there is a way to conceive of auxiliaries in terms of requirements of the interfaces. (Chomsky 2001, 2008, 2013 and many more.) Namely, we could appeal to the sensory-motor (SM) interface which is concerned with externalization, and the conceptual-intentional (CI) interface which is concerned with meaning. These interfaces are respectively fed by the syntactic representations phonological form (PF) and logical form (LF). Auxiliary *be* does not appear to introduce any additional semantics, but simply supports elements that themselves are semantically interpreted. This being the case, auxiliary *be* is irrelevant to any requirement of LF or the CI interface.

What about the SM interface and PF? Chomsky argues the SM interface is concerned with the externalization of language, that is to say, the encoding of the abstract hierarchical syntactic representation into something transmittable by the mouth, in the case of spoken language. Intuitively, then, processes occurring in response to PF requirements ought to relate to the optimization of the syntactic representation for externalization. Examples of this would be the linearization of hierarchical structure, and the insertion of phonological information so that the structure can be pronounced.[9] The insertion of auxiliary *be* could potentially be reduced to some PF requirement of a given language that inflectional morphology be realized or pronounced local to a verbal element.

Consider the silent auxiliary scenario argued for in this paper, however. In this configuration, the element being supported is the phonologically null present tense T^0, and due to rules of contextual allomorphy, the auxiliary *be* supporting this T^0 is itself null. If auxiliary *be* and what it supports can both be phonologically inert, such a configuration demonstrates that auxiliary *be* cannot be attributed to any requirement of PF, if PF is strictly concerned with feeding the SM interface a phonologically interpretable, utterable, object.

This is so for two reasons. First, if PF requirements care about pronounceability and phonological well-formedness for the sake of externalization, how would such requirements be violated by an un-supported functional morpheme that is not pronounced? Such

---

[9]Assuming a Distributed Morphology style Late Insertion framework.

a morpheme ought to be invisible to requirements that reference pronounced material. Second, if auxiliary *be* were motivated by such a PF requirement, how would the presence of a phonologically empty auxiliary satisfy that requirement? In short, then, if auxiliary *be* is PF-motivated, why is the auxiliary inserted to support elements that are phonologically absent, and how can the auxiliary satisfy any PF requirement if it is itself also silent? If auxiliaries are post-syntactically inserted following Bjorkman, contribute nothing to semantic well-formedness, and do not necessarily contribute anything to phonological well-formedness either, then auxiliary *be* is evidently not motivated by any straightforward requirement of syntax, LF, or PF. The auxiliary and what it supports can of course be pronounced, but the fact that they need not be indicates that the auxiliary requirement is not fundamentally related to a condition on PF. If this is accurate, what is the nature of the requirement to which we should attribute the motivation for auxiliary *be*?

I see two directions to go from here. One would be to hypothesize that the PF representation can have requirements that do not reduce to optimization for the purpose phonological well-formedness, or externalization. That is, a language could require its PF representations to have a V^0 element local to inflectional features, such that when there is no local V^0, an auxiliary is inserted. In other words, PF might require as its input a particular sort of syntactic structure, in which inflectional features are local to verbs.[10]

Consider however that since this statement of the motivation for auxiliaries has nothing to do with any need of the SM interface, it would be equivalently valid to posit that this auxiliary requirement is in fact motivated by a condition that LF imposes on a syntactic structure. This is because if the auxiliary requirement is fundamentally irrelevant to SM and CI, and therefore also to PF and LF, there is no obvious basis from which to claim with any strength that the requirement should be imposed by one interface rather than the other. The auxiliary is irrelevant both to externalization and interpretation, so neither interface is intuitively a more likely candidate for the locus of the motivation for auxiliary insertion. This line of reasoning where PF or LF can have requirements that do not reduce to well-formedness for the interface they relate to is, therefore, not very informative, and weakens the theory of what the interfaces are and do.

A second path to take would be to state that the auxiliary requirement is truly not motivated by syntax proper, nor LF, nor PF. Rather, the auxiliary could be motivated by some language-specific post-syntactic morphological process, occurring in a component such as what Harley & Noyer (1991) term Morphological Structure. Here apply post-syntactic operations that delete features (Impoverishment), combine the features of multiple terminals (Fusion), split individual terminals apart (Fission) and other operations like Lowering which modify the output of syntax in constrained ways. Bjorkman (2011) in fact formalizes auxiliary insertion as an instance of Fission, where a head that is not sister to a V^0, bearing an inflectional feature, is split into a V^0 node and a terminal bearing the inflectional feature:

(36) $[_{X^0} \, u\text{INFL}{:}\alpha \,] \rightarrow V^0 + [_{Infl} \, u\text{INFL}{:}\alpha \,]$
         where there is no V^0 already sister to X^0   (Bjorkman 2011, pp. 82)

---

[10]Thanks to Christopher Hammerly for this suggestion.

While nothing that I am claiming hinges directly on positing that auxiliary insertion is caused by a Fission operation, I suggest that auxiliary insertion indeed has the character of such a post-syntactic operation that is endemic to a separate morphological component. This is because the auxiliary requirement is not clearly reducible to a requirement of any other part of the grammar.

Following Bjorkman's arguments about overflow auxiliary patterns, auxiliary *be* should not be understood as arising in the syntax, but rather post-syntactically. If we do not weaken the theory of the interfaces so as to state that PF or LF can have requirements that do not reduce to any well-formedness condition relating to externalization or interpretation respectively, then based on the facts about North Azeri, we should not claim that the auxiliary is motivated by the PF (or, of course, LF) interface. This being said, the auxiliary must be motivated by some other part of the grammar.

In this sense, then, the facts about auxiliaries in North Azeri are evidence that we really do need a post-syntactic component of the grammar which has its own (constrained) set of operations, as has been proposed in the Distributed Morphology literature already. It would be ideal if we could minimize the addition of special representational levels to our theory of grammar, however I currently have developed no other way to incorporate the auxiliary facts argued for in this paper. Future work should examine to what extent the phenomena invoked as post-syntactic morphological operations, like Fusion, Fission, and based on what I've said here, auxiliary insertion, can be abolished or reduced to something more principled or not.

4.   Conclusion

In this paper I addressed several points, building off an argument about auxiliary verb phenomena in North Azeri. First, I argued that this language has a morpho-syntactically regular distribution of auxiliary *be*, which is obscured by morpho-phonological factors. Based on that discussion, I went on to argue that the distribution of allomorphy in North Azeri is evidence for a theory where allomorphic conditioning between two elements is structurally constrained. Finally I focused on the consequences of a particular sub-case among the auxiliaries, that of the present tense. After arguing that it is sensible to hypothesize that despite being null a present tense TP is truly present in the syntax, I considered the consequences of a null auxiliary element supporting another element which is itself phonologically null. I argued that as the motivation for auxiliaries is not reducible a purely syntactic requirement nor an interface requirement, and therefore we really may need a post-syntactic level of representation where such morphological operations apply.

References

Authier, Gilles. 2010. Azeri morphology in Kryz (East Caucasian). In Lars, Joanson. (ed.) *Turkic Languages 14*, 14-42.

Baker, Mark. 1985. The Mirror Principle and morphosyntactic explanation. *Linguistic Inquiry* 16.3: 373-415.

Bobaljik, J. 2012. *Universals in Comparative Morphology: Suppletion, Superlatives, and the Structure of Words*. MIT Press, Cambridge, MA.

Bjorkman, Bronwyn. 2011. *BE-ing default: The morphosyntax of auxiliaries.* Doctoral dissertation, Cambridge, Mass: MIT Press.

Chomsky, N. 2001. Derivation by phase. In *Ken Hale: A Life in language*, ed. M. Kenstowicz: 1 - 52. Cambridge, Mass.: MIT Press.

Chomsky, N. 2008. On Phases. In *Foundational Issues in Linguistic Theory*, ed. R. Freidin, C. P. Otero & M. L. Zubizarreta, 13366. Cambridge, MA: MIT Press.

Chomsky, N. 2013 The Problem of Projection. In *Lingua* 130: 3349.

Davis, C. 2015. Unifying Nominalized Clauses, Auxiliaries, and Copular Allomorphy in North Azeri. *Linguistic Portfolios*: Vol. 4, Article 6. St. Cloud State University.

Davis, C. 2017. Locality and copular allomorphy in North Azeri. In *Proceedings of the 12th Workshop on Altaic Formal Linguistics*. MIT Working Papers in Linguistics.

Göksel, Aslı. 2001. The auxiliary verb *ol* at the morphology-syntax interface. In Taylan E. Eser (ed.), *The Verb in Turkish*, 151-181. Amsterdam/Philadelphia: John Benjamins.

Göksel, Aslı. 2003. Syntactic head or morphological buffer? The dual status of *ol*. In A.S. Özsoy, E. Taylan, A. Koç, D. Akar and M. Nakipoğlu-Demiralp (eds.) *Studies in Turkish linguistics*, pp. 48-57. Boğaziçi University Press.

Göksel, A., & C. Kerslake. 2005. *Turkish: A Comprehensive Grammar.* Abingdon, UK: Routledge.

Halle, M & A, Marantz. 1993. Distributed morphology and the pieces of inflection. In K. Hale and S.J. Keyser (ed.): *The View From Building 20*, 1-52. Cambridge, MA: MIT Press.

Harley, H & R, Noyer. 1999. Distributed morphology. In *Glot International* 4:4. 3 - 9.

Kelepir, M. 2001. *Topics in Turkish syntax: Clausal Structure and Scope.* Doctoral dissertation, MIT.

Kornfilt, Jaklin. 1996. On some copular clitics in Turkish. In A. Alexandiou, N. Fuhrhop, P. Law and S. Loehken (eds.): *Papers in Linguistics.* Berlin: Zentrum für Allgemeine Sprachwissenschaft 6, p. 96-155.

Lewis, G. L. 1975. *Turkish grammar*. Oxford University Press: Oxford.

Miyagawa, Shigeru. 2012. *Case, argument structure, and word order.* New York, NY: Routledge.

Öztopçu, Kurtuluş. 2003. *Elementary Azerbaijani.* Santa Monica, California.

Payne, Thomas 1997. *Describing morphosyntax: A guide for field linguists.* Cambridge, UK: Cambridge University Press.

Ross, J.R. 1967. *Constraints on variables in Syntax.* Doctoral dissertation, MIT.

Schönig, Claus. 2006. Azerbaijanian. In: Lars Johanson & Éva Ágnes Csató (eds.) *The Turkic languages.* London & New York: Routledge. 248-260.

Colin P. Davis
colind@mit.edu

Relative clause strategies in languages of East Anatolia: Divergence and convergence

Ophélie Gandon

University Sorbonne Nouvelle Paris 3 & UMR7528 Mondes iranien et indien

1.   Introduction

Turkey, especially Eastern Anatolia, displays a non-trivial linguistic diversity with languages belonging to numerous distinct genetic families: beside Turkish, Iranian languages (Zazaki and Kurmanji Kurdish), European languages (Romeyka and Armenian), Semitic languages (Arabic and Neo-Aramaic dialects in South-Eastern Turkey), and Caucasian languages (especially South Caucasian with Laz) are spoken in the area. Though the proportion of speakers of minority languages has generally significantly decreased since the beginning of the 20th century, these languages have been in a long-standing contact in the past and possibly have influenced one another. The possibility of a linguistic area in East Anatolia where unrelated languages would converge regarding their linguistic features has been mentioned in Matras (2010, p. 75–76), and then refuted in a recent paper by Haig (to appear). Building on various phonological and morphosyntactic features, Haig shows that there are rather two "areal epicenters": the "southeastern Kurdish/NENA/Arabic 'Mesopotamian' region'", and the "northern Turkic/Kartvelian/Armenian 'Caspian/Caucasian' region" (*ibid*).

This paper investigates divergences and convergences regarding relativization strategies in languages spoken in East Anatolia, an issue that has not been addressed in Haig's paper. Our definition of relative clauses (henceforth RC) combines syntactic and semantic criteria: a RC is a subordinate clause that modifies a noun, and the modified noun (head noun) has a syntactic function in both the matrix clause and the relative one. The constructions singled out by this definition may vary according to numerous parameters, leading to different relativization strategies. RCs thus present an interesting situation because the diffusion of various features may be investigated. In this paper, I confine myself mainly to three variation parameters, *i.e.* the position of the RC with respect to the head noun, the finiteness of the verbal form in the RC, and the presence or not of a relativizer and accordingly its nature. Some more specific features such as the use of resumptive pronouns or the possibility of free RCs are considered as well, when information is available. When doing so, two clearly distinct areas appear, which interestingly overlap with the two areas identified in Haig's paper: an area in the south

where there is no real linguistic convergence between languages, and an area in the north where languages converge more or less.

The first part briefly presents the participial strategy of Turkish, the dominant language in Turkey nowadays, and its specific features. The second part describes the relativization strategies in minority languages spoken in the southern part of Eastern Anatolia and shows the absence of real convergence. The third part deals with minority languages spoken in the northern part of Eastern Anatolia: they all diverge from their genetic families and converge more or less regarding various features. Finally, the fourth part attempts to provide some explanations for these differences of behavior observed between languages in contact.

2.   Turkish: Prenominal participial RCs

Turkish is the official language of Turkey and the majority of the population speaks Turkish. Turkish resorts to prenominal participial RCs as its main relativization strategy (ex. (1)), which is the common strategy for Turkic languages (see ex. (2) in Old Turkic, and ex. (3) in Chuvash, genetically the most distant of Turkic languages and spoken in Russia)[1]:

(1)   *Standard Turkish*
        Erik-ler   al-dığ-ım   adam pazar-ın   baş-ın-da
        plum-PL-ACC take-PTC.POSS.1S man   market-GEN head-POSS.3S-LOC
        dur-uyor.
        stand-PROG
        'The man I bought the plums from stands at the beginning of the market.'

(2)   *Old Turkic (Erdal 2004, p. 441)*
        [bo   kiši-niŋ   ätöz-in   yu-mïš]   suv
         DEM   person-GEN body-POSS.3SG.ACC   wash-PTC   water
        'the water (with) which this person washed his body'   (DhāSū 31)

(3)   *Chuvash (Ersoy 2010, p. 452)*
        [epir   şit-ni]   şir
         PRO.1PL reach-PTC place
        'the place we reach'

This strategy can relativize all kind of syntactic functions, and generally does not resort to resumptive pronouns[2]. Note that the subject of the RC in Turkish is marked with the

---

[1] For all examples, the RC is between square brackets, the head noun is underlined, and the relativizer if any is in bold.
[2] The reflexive pronoun *kendi(si/ler)* may be used as a resumptive pronoun when relativizing syntactic functions very low on the hierarchy, but it is not very common (see Aslı Göksel and Kerslake 2005, p. 384) for some examples).

genitive case, which is a common feature to several other Turkic languages (see ex. (2) in Old Turkic)[3].

Unlike other Turkic languages however, the participle *-GAn* is mainly subject-oriented (or oriented toward a genitive possessor extracted from the subject position) in Turkish (and to a lesser extent in Azeri)[4]. Finally, free relative clauses (that is, without any head noun or a substitute) are possible:

(4)   *Standard Turkish*
        Ali [telefon-da y̲ü̲k̲s̲e̲k̲ ses-le   konuş-an]-lar-a   çok kız-ıyor.
        Ali phone-LOC high   sound-COM speak-PTC-PL-DAT very get_angry
        'Ali gets angry with people speaking loudly on the phone.'

The participle thus takes the nominal marks that would normally bear the head noun, plural and dative in the example above.

3.   The southern part of Eastern Anatolia

Besides Turkish, Iranian languages (Kurmanji and Zazaki) and Semitic languages (Arabic and Neo-Aramaic dialects) are spoken in South Eastern Anatolia. Though it is difficult to give a precise number of speakers for each of these languages, Iranian languages are by far the most widespread minority languages in Turkey, especially with Kurmanji Kurdish which displays several million of speakers (the estimation varies between 8 and 15 million (Öpengin and Haig, to appear)). Paul estimates the number of Zaza speakers in Southeastern Anatolia between 1,5 and 2 million (Paul 2009, p. 545).

Semitic languages are spoken in the South near the border with Irak and Syria, by much less speakers. 365 340 speakers were reported for Arabic in Turkey in the 1965 census (P. A. Andrews 1989, p. 148). The number of Neo-Aramaic speakers was perhaps up to 50 000 until fifty years ago according to the references given in P. A. Andrews (1989, p. 161); but most of them have emigrated the West since.

Iranian and Semitic languages both resort to a similar strategy, which diverges totally from the Turkish one, that is, finite postnominal RCs introduced by a complementizer or an invariable relativizer[5].

---

[3] Unlike other Turkic languages however, the possessive mark which refers to the subject of the RC is realized on the participle in Turkish, while it is usually realized on the head noun in other Turkic languages.
[4] It can also occasionally relativize other functions in case the subject of the relative clause has a low degree of individuation and has minimal control over the event described by the verb; cf. Haig for more details (1998, p. 174–85).
[5] Iranian and Neo-Aramaic languages resort to a complementizer to introduce RCs (that is, an element that can introduce other kinds of subordinate clauses as well), while Arabic dialects resort to a simple relativizer (specific to relative clauses, and which is usually very similar or formally identical to the definite article).

(5) *Kurmanji Kurdish (Iranian)*
Bajar-ê [**ku** ez Lê mezin bûm xweş e
town-LK **COMP** I PREP.PRO.3SG big be.PST.1SG nice COP.3SG
'the town I grew up in is nice.'

(6) *Zazaki (Iranian) (Paul 2009, p. 574–75)*
girwey-o [**ki** mā nē-zāne
work-LK **COMP** we NEG-know.PRS.PL
'a work that we don't know'

(7) *North-Eastern Neo-Aramaic, dialect of Jilu (Hakkâri) (Semitic) (Fox 1997, p. 81)*
en zuz-e [**t**-itwa šqile mənn-an]
these coin-PL COMP-COP.2MSG.PST taken.PL from-1PL
'the money that you have taken from us'

(8) *Arabic dialect of Mardin (Grigore 2007, p. 309–10)*
ayn tmān [**la** kān yəlzəm tə 'tī-ni]
where money **REL** was necessary for give-me
'Where is the money you had to give me?'

When relativizing syntactic functions low on the hierarchy, both Iranian and Semitic resort to resumptive pronouns (*e.g.* ex. (5) where *lê* is a contraction of the preposition *li* 'in' and the pronoun 3SG.F.OBL *wê*). Iranian and Semitic languages diverge only on a few minor points, such as the presence of a linker on the head noun for Iranian languages, or the circumstances for the omission of the relativizer/complementizer[6].

However, the resemblance between Iranian and Semitic cannot be accounted for by any convergence phenomenon, since exactly the same features are found in other Iranian and Semitic languages spoken outside Eastern Anatolia, as shown by the examples below. Persian is an Iranian language spoken in Iran:

(9) *Persian*
un(ân) doxtar-i [**ke** (man) mi-xâh-am] in nist
that girl-LK **COMP** I IPFV-want-1SG this is.not
'This is not the girl whom I want'

Hebrew is a sister language of Aramaic spoken mainly in Israel (Hebrew and Aramaic both belong to the Northwest branch of the Semitic family):

(10) *Hebrew (Vaillette 2001, p. 305)*
ha-yeled$_i$ [**še** raʔiti (ʔoto$_i$)]
the-boy$_i$ **COMP** saw.1SG him$_i$
'The boy that I saw.'

---

[6] The complementizer may be omitted when the head noun is indefinite and/or when the RC is non-restrictive in Neo-Aramaic and Arabic dialects, while in Kurmanji Kurdish it can be omitted when relativizing the syntactic function of object.

And finally, Levantine Arabic, the Arabic dialect spoken in Lebanon, resorts to finite postnominal RCs introduced by a relativizer similar to the definite article, as do Arabic dialects spoken in South of Eastern Anatolia:

(11)   *Levantine Arabic (McLoughlin 2003, p. 63)*
        hum   'ir-<u>rijaal</u> [**'illee** ruHt 'ila bayroot fee   sayyaarat-hum]
        PRO.3PL.M DEF-men **REL**   I.went to Beirut   with car-POSS.3PL.M
        'They are the men in whose car I went to Beirut.'

All these languages can resort to resumptive pronouns as well when relativizing syntactic functions low on the accessibility hierarchy (see *e.g.* ʔoto 'him' in ex. (10)).

   Thus, the shared features of Iranian and Semitic languages spoken in Southeastern Anatolia regarding relativization strategies cannot be considered as an instance of convergence, since they seem to be inherited. Moreover, the finite postnominal strategy is worldwide typologically very common (Creissels 2006, p. 223, 240 vol. 2): the similarity between Iranian and Semitic is thus most probably the result of likelihood. Besides, the prenominal participial strategy of Turkish diverges totally (prenominal position of the RC, non-finite verbal form, no relativizer, no resumptive pronouns). Therefore, overall[7] no convergence or linguistic area is identifiable in Southeastern Anatolia with respect to relativization strategies.

4.   The north part of Eastern Anatolia

The North part of Eastern Anatolia on the other hand shows a quite different picture. Various minority languages are spoken there beside Turkish: particularly Laz (a South Caucasian language), Homshetsi (a dialect of Eastern Armenian), and Romeyka (a Greek language). Interestingly these languages all diverge more or less from their respective families regarding relativization strategies. Besides, they all converge regarding at least one feature, the prenominal position of the RC.

4.1   Romeyka: Prenominal finite RCs

Romeyka designates the variety of Pontik Greek spoken by Muslim Greeks who remained in the province of Trabzon along the Black Sea after the population exchange with Greece in 1923. 4 535 speakers were reported in the 1965 census of Turkey (P. A. Andrews 1989, p. 145).

   Romeyka resorts to finite prenominal RCs, which is a typologically uncommon combination: prenominal RCs are most often non-finite worldwide (Keenan 1985, p. 160, A. D. Andrews 2007, p. 208, *inter alia*). The relative clause is marked with a relativizer, which precedes the verb:

---

[7] The only sign of an eventual convergence phenomenon is the placement of the complementizer in Neo-Aramaic languages, which tends nowadays to attach enclitically to the head noun instead of proclitically to the first constituent of the RC, thus possibly patterning the placement of the ezafe (linker) in Kurmanji Kurdish (Cohen 2015). Rather than a direct influence of Kurmanji Kurdish, Gutman (to appear) suggests an areal preference for head-marking.

(12)   *Romeyka (Neocleous, in preparation)*
       [pso'min   **'d**=enʤen]   o   pe'ðas,   t=e'mon
        bread.ACC **REL**=bring.PST.3SG the.NOM child.NOM the.NOM=my
       d=a'nepsin   'eton
       the.NOM=nephew.NOM be.IMPF.3SG
       'The child that brought bread was my nephew'

Other Greek languages spoken outside Anatolia commonly resort to finite postnominal RCs, introduced either by a complementizer or a relative pronoun[8]:

(13)   *Modern Greek (Holton, Mackridge, and Philippaki-Warburton 2012, p. 532)*
       O   kathiyitís [**pou**   mas ékane istoría] ítan polí kalós
       the teacher   COMP   us   taught history was very good
       'The teacher who taught us history was very good.'

(14)   *Modern Greek (Holton, Mackridge, and Philippaki-Warburton 2012, p. 533)*
       I   sizítisi   [**tin**   **opía**   kaname] prépi
       the conversation ART.DEF.F.SG.ACC RPRO.F.SG.ACC we.made must
       na   míni metaxí   mas
       to   stay between us
       'The conversation which we had must remain between us.'

Thus Romeyka diverges from its genetic family, particularly regarding the position of the RC. Moreover, the construction it resorts to is typologically uncommon. Given the fact that surrounding languages (namely Turkish, Laz and Homshetsi) all resort to prenominal RCs (see section 2. above and sections 4.2 and 4.3 below), it is conceivable that the prenominal position is the result of an areal influence. Note that interestingly, speakers of Pontik Greek who went back to Greece after the population exchange now resort to postnominal RCs (Drettas 1997, p. 347–57)[9].

4.2   Homshetsi: Prenominal participial RCs

Hemshinli are Armenians settled in Artvin and Rize provinces (along and near the Black Sea). Nowadays only Hemshinli of Artvin keep speaking Homshetsi. The number of speakers is estimated at 26 000 (Simonian 2007, p. xxi). Homshetsi is considered as a dialect of Western Armenian, tough there is no mutually intelligibility between both (Vaux 2007, p. 257).
     Homshetsi is probably the most convergent with Turkish regarding relativization strategies: it resorts to prenominal participial relative clauses.

---

[8] The postnominal relative pronoun strategy (which was the common one in Ancient Greek) is more specific to formal language and/or when relativizing syntactic functions low on the accessibility hierarchy for Modern Greek (Holton, Mackridge, and Philippaki-Warburton 2012, p. 532).
[9] A similar situation is reported for Cappadocian and Pharasiot Greek (two other dialects of Asia Minor Greek which were spoken more in the West in the center of Turkey), which used to resort to prenominal RCs and whose speakers nowadays settled in Greece resort to both, prenominal and postnominal RCs (Bağrıaçık 2015).

(15) *(Dumézil 1964, p. 19)*
    [šad   paa   unn-oɣ] <u>marte</u>
     many money have-PTC man
    'the man who has a lot of money'

(16) *(Dumézil 1964, p. 19)*
    [im   des-aj] <u>marte</u>-s
     PRO1SG.GEN see-PTC man-POSS.1SG
    'the man I saw/see'

(17) *(Dumézil 1964, p. 19)*
    [axčkone des-aj] <u>marte</u>-s
     girl.DAT see-PTC man-POSS.1SG
    'the man whose daughter I saw

(18) *(Dumézil 1964, p. 19)*
    [c'e   xabr-aj] <u>marte</u>-t
     PRO2PL.GEN talk-PTC man-POSS.2PL
    'the man whom you are talking about'

By contrast, Eastern Armenian (spoken in Armenia) resorts to finite postnominal RCs introduced by relative pronouns as its main relativization strategy, which was also the case for Classical Armenian:

(19) *Eastern Armenian (Dum-Tragut 2009, p. 479)*
    <u>ɫarabałc'i</u>-ner   ēl   k-an   [**or-onc'** hamar
    Karabakhian-PL.NOM also   exist-PTC.PRS **which-PL.DAT** POST
    Samvel-ě   heros   ē]
    Samvel.NOM-the hero.NOM he is
    'There are also Karabakhians for whom Samvel is a hero.'

Interestingly, no mention of such a postnominal relative pronoun strategy is made in Dumézil for Homshetsi. Eastern Armenian also resorts to participial RCs, but as a secondary strategy which is restricted to the relativization of the subject and the object (Hewitt 1978, p. 129), while it can relativize various syntactic functions in Homshetsi (as seen in the examples (15) to (18)), exactly as it is the case for Turkish (see section 2). Another interesting fact is the genitive marking of the subject of the relative clause, similar to Turkish again (see ex. (16) and (18))[10].

---

[10] This feature is also present in Eastern Armenian when the subject of the RC is a personal pronoun (Dum-Tragut 2009, p. 508), so Homshetsi does not diverge from Eastern Armenian with this respect. However it is worth noticing that Eastern Armenian itself has Turkic neighboring languages, Turkish on the West and Azeri on the East. The genitive marking of subjects of RCs is actually a feature that extends beyond Turkish borders and is found in various other non-related languages of the area (Gandon, in preparation).

4.3   Laz

Laz is a South Caucasian language spoken in the very North-East of Turkey along the Black Sea and near the border with Georgia. In the census of 1965, there were 26 007 persons declaring speaking Laz as a mother tongue and 59 101 as a second language in Turkey (P. A. Andrews 1989, p. 176). Laz may thus be considered as the most important minority language in Northeastern Anatolia nowadays. Laz mainly resorts to prenominal finite RCs, and to participial RCs as a secondary strategy. For both strategies, Laz diverges from the other South Caucasian languages with respect to various features. Some of these features are shared with other languages of the area (Turkish, Romeyka).

4.3.1   Prenominal finite RCs

The main relativization strategy of Laz is very similar to the one of Romeyka, that is, a prenominal finite RC (see ex. (20)), which is again typologically uncommon. Interestingly, Laz has been in a long standing contact with Pontik Greek (Drettas 2006).

(20)   *Laz of Pazar (Öztürk and Pöchtrager et al. 2011, p. 134)*
        [eni opşa **na**   m-a-şk'urin-e-n]   şuroni   layç'i o-n
         the most **COMP** O1-APPL-fear-TS-PRS.3SG animal.NOM dog   COP-PRS.3SG
        'The animal that I am the most afraid of is the dog.'

Unlike Romeyka, Laz resorts to a complementizer instead of a simple relativizer. The position uses to be very similar, that is, the preverbal one. Note however that for the dialect of Arhavi the complementizer may also attach enclitically to one of the constituents preceding the verb if any (see Lacroix 2009, p. 750–54 for more details).
This strategy may relativize a wide range of syntactic functions and does not make use of resumptive pronouns (gap strategy) (Lacroix 2009, p. 754–55). Unfortunately, I do not have information for Romeyka about these more specific features.
     Among their main relativization strategies, South Caucasian languages commonly resort to two distinct strategies: postnominal finite RCs introduced by relative pronouns (almost absent in Laz (Holisky 1991, p. 419, p. 457, Lacroix 2009, p. 768–69)), and pre-/(post)nominal RCs with a complementizer. The complementizer strategy is available in Mingrelian and in Georgian[11] (it is not reported for Svan, Tuite 2004):

(21)   *Georgian (Nash 2002)*
        [nino **rom** imas-tan   cxovrobs] me   im   kal-s   vicnob
         Nino **COMP** DEM.DAT-with he.lives   I   DEM woman-DAT I.know
        'I know the woman that Nino lives with.'

Though the appellation is similar (finite RC with a complementizer), the strategy of Laz actually differs from the ones found in Georgian and Mingrelian with respect to several

---

[11] For Georgian, the complementizer strategy is reported to be more specific to the spoken language, while the relative pronoun strategy is more common to the formal/written language (Aronson 1972, p. 139; Hewitt 1987, p. 187; Hewitt 1995, p. 606; Harris 1992, p. 394).

features. Due to lack of space, I will confine myself here to a comparison with Georgian only, which is the best documented South Caucasian language.

(i)   In Laz the RC is always prenominal while for Georgian the postnominal position of the RC is also possible:

(22)   *Georgian (Hewitt 1987, p. 194)*
        vxedav   im   saxl-s,   [zurab-i   **rom**   cxovrobs]
        I.see.it.PRS that   house-DAT Zurab-NOM   **COMP**   live.3SG.PRS
        'I see the house where Zurab lives.'

Given the fact that languages of the area all resort to prenominal RCs, and that the combination of a finite verbal form with the prenominal position in typologically uncommon for RCs, an areal influence is worth considering here.

 (ii)   Laz allow free relative clauses with this strategy (that is, with no head noun and no substitute), while this does not seem possible with the complementizer strategy of Georgian (Georgian does display free RCs, but with the relative pronoun strategy):

(23)   *Laz of Arhavi (Lacroix 2009, p. 759)*
        [si-**na**   çk'om-i]-pe-k   va   g-o-dzg-es-na…
         2SG-**COMP** manger-AOR-PL-ERG NEG srII2-VAL1-rassasier-AOR.srI.3PL-if
        'Si ceux que tu as mangés ne t'ont pas rassasié…'   (K'72.137)

Surprisingly, a finite verbal form thus takes nominal marks (plural and ergative in the example (23) given here). The fact that the verb takes the nominal marks of the head noun in case of free relatives is exactly the pattern noticed in Turkish: see ex. (4) of section 2. Thus once again an areal influence may be considered here.

(iii)   Georgian resorts to resumptive pronouns when relativizing syntactic functions low on the hierarchy (see the demonstrative *imas* in the example (21)); which is not the case for Laz, which resorts to the gap strategy.

Interestingly again, this is in line with the Turkish participial strategy. Note however this time that the use of resumptive pronouns with prenominal RCs is anyway typologically not common (Keenan 1985, p. 148–49, Dik 1997, p. 2:46, Creissels 2006, p. 239, p. 242 vol. 2, *inter alia*). Consequently the similarity between Laz and Turkish regarding this feature may just be the result of likelihood.

(iv)   When the RC is prenominal, the head-noun must be introduced by a demonstrative in Georgian (Nash 2002), which is not the case for Laz (see ex. (20) (nor for Turkish).

(v)   Finally, the form and the placement of the complementizer are different[12]. Note
        that some authors suggest a possible influence of the modal particle *na* of Pontik
        Greek in the origin of the complementizer *na* in Laz, see Drettas (2006) and
        Lacroix (2012) for two different scenarios.

4.3.2 Prenominal participial RCs

Beside the complementizer finite strategy, Laz also resorts to participial prenominal RCs
as a secondary strategy. South Caucasian languages as well resort to participial
prenominal RCs as a secondary strategy. However Laz again diverges from other South
Caucasian languages with respect to at least two points:

First, South Caucasian languages display oriented participles, while this does not
seem to be the case for at least Laz of Pazar and Laz of Arhavi. Only one main participle
-eri is mentioned for these two dialects, which can relativize either subjet or object
(Lacroix 2009, p. 657, Öztürk & Pöchtrager et al. 2011, p.130). Note that this situation of
Laz is interesting since Turkish as well displays oriented participles. Laz thus diverges
from both cognate and neighboring languages.

Secondly, Laz of Arhavi presents an even more interesting situation: besides the
subject and objects, this main participle can also relativize other syntactic functions,
while it is mainly restricted to the relativization of the subject and the object in other
South Caucasian languages. The example (24) below illustrates the relativization of a
dative argument:

(24)   *Laz of Arhavi (Lacroix 2009, p. 661)*
        a.   p'et'mezi-s   mtugi dol-a-şkid-u
              treacle-DAT   mouse PV-VAL5-drown-AOR.srI3SG
              'a mouse drowned in treacle'   (Ž.39)

        b.   [mtugi dolo-şkid-eri]   p'et'mezi
               mouse PV-be_drowned-PTC treacle
              'treacle in which a mouse drowned'   (Ž.39)

The subject of RCs is not marked with the genitive case unlike Turkish and Homshetsi;
note however that subjects of other kinds of non-finite subordinate clauses may in Laz of
Pazar (2011,p. 132), and in Laz of Arhavi with intransitive verbs (Lacroix 2009, p. 649–
50)[13].

5.   Observations and suggestions

To summarize, no switch in relativization strategies is observable in the south part of
Eastern Anatolia: each language keeps what seems to be its inherited strategy. Beside,
languages spoken in the north part of Eastern Anatolia diverge more or less from their

---

[12] For Georgian the complementizer seems to avoid the first position and is somewhere between the first
constituent and the verb (Hewitt 1987, p. 187).
[13] A similar situation is reported for Georgian (Harris 1981, p. 156–58; Hewitt 1987, p. 187).

respective genetic families and converge together at least according to one feature, the prenominal position.

Thus, languages behave differently in language-contact situations. What are the circumstances that may favor or not the diffusion of a linguistic feature and linguistic convergence? In this section, I suggest some factors that may be relevant for a language-contact induced change to occur or not.

5.1 Sociolinguistic factors

Sociolinguistic factors such as the number and the degree of bilingualism of speakers may be involved. Turkish in contact with Iranian and Semitic languages in the South part of Eastern Anatolia does not switch to a postnominal strategy, though other Turkic languages did in Iran: *e.g.* Khalaj, Sonqor Turkic and Kashkay (Kıral 2000, p. 183, Bulut 2005, p. 264, Dolatkhah 2012, p. 190). The reason here seems obvious: Iranian and Semitic languages are minority languages in Turkey, while Turkish is dominant as the unique official language. The major part of the population is monolingual in Turkish and thus is not in a language contact situation; only speakers of minority languages are bilingual.

5.2 Cognitive factors

Semitic and Iranian languages on the other hand do not switch either to the participial prenominal strategy of Turkish (within Turkey). Cognitively, one may expect that a postnominal RC will be easier to proceed than a prenominal one: indeed with a prenominal RC, the speaker has to keep in mind the RC until she accedes the noun it modifies. Note that postnominal RC strategy is by far the dominant strategy world-wide (see de Vries 2001, p. 235, Creissels 2006, p. 223 vol. 2, Dryer 2013, *inter alia*). This could explain why Iranian and Semitic languages resorting to a postnominal strategy do not switch to the Turkish participial prenominal one, despite the fact that an important part of their speakers are bilingual and thus in a language-contact situation.

5.3 Linguistic structural factors

However, the cognitive factor does not account for the situation of minority languages spoken in the northern part of Eastern Anatolia, which all resort to prenominal RCs while their cognates spoken outside Turkey resort to postnominal RCs. Linguistic structural factors may be suggested here: for all these languages, adjectives precede the nouns they modify (including their cognates spoken outside the area), while in most of Iranian languages and in Semitic ones including those spoken in Turkey, adjectives follow the noun they modify. Thus one may expect that it will be easier for a language to switch to prenominal RCs when adjectives are already preceding the nouns they modify.

Linguistic structural factors may also provide an explanation for the fact that Romeyka does not switch to the participial strategy of Turkish, as did Homshetsi. Indeed, Homshetsi already displayed participles inherited from its ancestor language, while only a few adjectives formed with old participles are available in Pontik Greek (Drettas 2006, p. 12).

5.4   Typological tendencies and universals

Romeyka and Homshetsi have both lost their relative pronouns (while Ancient Greek, Classical Armenian, Modern Greek and Eastern Armenian all display relative pronouns). Interestingly, the combination of prenominal RCs and relative pronouns seems impossible: it is not attested worldwide (Downing 1978, p. 382, p. 396; Keenan 1985, p. 149; Creissels 2006, p. 240 vol. 2; A. D. Andrews p. 2007, p. 218, *inter alia*). Thus, a change regarding one feature (*e.g.* the loss of relative pronouns) may be linked to another one (*e.g.* the position of the RC).

5.5   Contact time-depth

Finally, the contact time-depth certainly has an influence. One could expect that the longest in time two languages stay in contact, the more they have probability to influence one another and to converge. Turkish has become a dominant language in the area only recently, with the foundation of the Republic of Turkey in 1923 based on the policy of one unique language, Turkish. Before this, Laz has been in a long-standing contact with Pontic Greek (Drettas 2006). Thus this can be an explanation for why Laz does not switch to the participial prenominal strategy of Turkish as dominant strategy, though structurally it would be suitable for this.

6.   Conclusion

Considering various variation parameters of relativization strategies in languages spoken in Eastern Anatolia, two distinct areas arise. In the south part, languages (Iranian, Semitic and Turkish) do not diverge from their respective genetic families and no linguistic convergence can be identified. Iranian and Semitic languages resort to a similar strategy, but this is most probably the result of likelihood. In the north part on the other hand, languages (Laz, Romeyka, Homshetsi) all diverge from their respective genetic families and converge together more or less regarding various features. One feature is shared by all of them: the prenominal position of the RC.

Thus, languages show quite different behaviors regarding contact situations. Several factors may be relevant for a language-contact induced change to occur or not: a few have been suggested here but there are most certainly several other factors that may be involved (geographical, educational, etc.). The fact that Laz displays non-oriented participles while both cognate and neighboring languages display oriented participles for example is an interesting case that cannot be accounted for with the factors suggested here.

Finally, in a further work it would be interesting to investigate some more specific features that have not been considered here due to lack of place and/or lack of documentation, and some more languages (*e.g.* Circassian dialects spoken in Turkey: Circassian is a North-West Caucasian language, and North-West Caucasian languages indeed display a typologically very uncommon relativization strategy). The increasing documentation of minority languages during the last years will hopefully allow for leading such studies to a wider extent in a near future, possibly including some statistical tools.

Abbreviations

1/2/3: 1st/2nd/3rd pers, ACC: accusative, AOR: aortist, APPL: applicative, ART: article, COM: comitative, COMP: complementizer, COP: copula, DAT: dative, DEF: definite, DEM: demonstrative, ERG: ergative, F: feminine, GEN: genitive, IMPF: imperfect, IPFV: imperfective, LK: linker, LOC: locative, M: masculine, NEG: negative, NOM: nominative, O: object marker, OBL: oblique. PL: plural, POSS: possessive, POST: postposition, PREP: preposition, PRO: pronoun, PROG: progressive, PRS: present, PST: past, PTC: participle, PV: preverb, REL: relativizer, RPRO: relative pronoun, SG: singular, srI/II: series markers, TS: thematic suffix, VAL1/5: valence operator.

References

Andrews, Avery D. 2007. Relative Clauses. In *Language Typology and Syntactic Description*, ed. by Timothy Shopen, 206–36. Cambridge: CUP.

Andrews, Peter A., ed. 1989. *Ethnic Groups in the Republic of Turkey*. Wiesbaden: Reichert Verlag.

Aronson, H. I. 1972. Some Notes on Relative Clauses in Georgian. In *The Chicago Which Hunt: Papers from the Relative Clause Festival*, ed. by Paul M. Peranteau, Judith N. Levi, and Gloria C. Phares, 136–143. Chicago: Chicago Linguistic Society.

Bağrıaçık, Metin. 2015. Relative Clauses in Pharasiot Greek. presented at the Workshop Typologie aréale et stratégies de relativisation, Labex EFL/INALCO/MII/Sedyl: Paris.

Bulut, Christiane. 2005. Iranian Influence on Sonqor Turkic. In *Linguistic Convergence and Areal Diffusion: Case Studies from Iranian, Semitic and Turkic*, ed. by Éva Ágnes Csató, Isaksson Bo, and Jahani Carina, 241–270. Abington, New York: RoutledgeCurzon.

Cohen, Eran. 2015. Head-Marking in Neo-Aramaic Genitive Constructions and the Ezafe Construction in Kurdish. In *Semitic Languages in Contact*, ed. by Aaron M. Butts, 114–25. Leiden: Brill.

Creissels, Denis. 2006. *Syntaxe Générale : Une Introduction Typologique*. Paris: Lavoisier.

de Vries, Mark. 2001. Patterns of Relative Clauses. *Linguistics in the Netherlands* 18: 231–43.

Dik, Simon C. 1997. *The Theory of Functional Grammar. Part 2: Complex and Derived Constructions*. ed by Kees Hengeveld. Vol. 2. Berlin, New York: Mouton de Gruyter.

Dolatkhah, Sohrab. 2012. Elements for a Grammar of Kashkay (a Turkic Language of Iran). Paris: Ecole Pratique des Hautes Etudes.

Downing, Bruce T. 1978. Some Universals of Relative Clause Structure. In *Universals of Human Language*, ed. by Joseph H. Greenberg, Charles A. Ferguson, and Edith A. Moravcsik, 4:375–418. Stanford: Stanford University Press.

Drettas, Georges. 1997. *Aspects Pontiques*. Paris: ARP (Association de Recherches Pluridisciplinaires).

———. 2006. Les Joncteurs Subordonnants Na (Modal) et Oti (Complémentiseur) Du Grec Ont-Ils Été Empruntés En Laze et Pourquoi ? In *Atelier International de Morphosyntaxe Des Langues Du Caucase*. Paris, 11-14 décembre.

Dryer, Matthew, S. 2013. Order of Relative Clause and Noun. In *The World Atlas of Language Structures Online*, Max Planck Institute for Evolutionary Anthropology. Leipzig: Dryer, M. & M., Haspelmath. http://wals.info/chapter/90.

Dumézil, Georges. 1964. *Notes Sur Le Parler D'un Arménien Musulman de Hemşin*. Bruxelles: Palais des Académies.

Dum-Tragut, Jasmine. 2009. *Armenian: Modern Eastern Armenian*. Amsterdam: Benjamins.

Erdal, M. 2004. *A Grammar of Old Turkic*. Lieden : Brill.

Ersoy, Feyzi. 2010. *Çuvaş Türkçesi Grameri*. Gazi Kitabevi. Ankara.

Fox, Samuel. 1997. *The Neo-Aramaic Dialect of Jilu*. Wiesbaden: Harrassowitz Verlag.

Gandon, Ophélie. To appear. Convergence Areas in the Caucasus-Western Iran Area with Respect to Relativization Strategies. In *RJC 2015 - Actes Des 18èmes Rencontres Jeunes Chercheurs de l'ED 268 : Contact de Langues : Situations, Représentations, Réalisations.* Sorbonne Nouvelle, Paris.

———. in preparation. La Relativisation Dans Une Perspective Aréale: L'aire Caucase Iran-Anatolie. Université Sorbonne Nouvelle.

Göksel, Aslı, and Celia Kerslake. 2005. *Turkish: a comprehensive grammar*. London, New York: Routledge.

Grigore, Georges. 2007. *L'Arabe Parlé À Mardin. Monographie D'un Parler Arabe "périphérique."* Editura Universității din Bucureşti. Bucharest.

Gutman, Ariel. to appear. Can Pattern Replication Be Easily Established? The Case of the Neo-Aramaic Neo-Construct. In *RJC 2015 - Actes Des 18èmes Rencontres Jeunes Chercheurs de l'ED 268 : Contact de Langues : Situations, Représentations, Réalisations.* Sorbonne Nouvelle, Paris.

Haig, Geoffrey. to appear. East Anatolia: Language Contact in a Transition Zone. In *The Cambridge Handbook of Areal Linguistics*, ed. by Raymond Hickey. Cambridge: Cambridge University Press.

———. 1998. *Relative Constructions in Turkish*. Wiesbaden: Harrassowitz.

Harris, Alice C. 1981. *Georgian Syntax. A Study in Relational Grammar*. Cambridge: CUP.

———. 1992. Changes in Relativization Strategies: Georgian and Language Universals. In *Caucasologie et Mythologie Comparée, Actes Du Colloque International Du CNRS - IV Colloque de Caucasologie, Sèvre, 27-29 Juin 1988*, edited by Catherine Paris, 391–403. Paris: Peeters.

Hewitt, Brian G. 1978. The Armenian Relative Clause. In *International Review of Slavic Linguistics*, 3:99–138.

———. 1987. *The Typology of Subordination in Georgian and Abkhaz*. Berlin: Mouton de Gruyter.

———. 1995. *Georgian, a Structural Reference Grammar*. Amsterdam: Philadelphia : John Benjamins.

Holisky, D. A. 1991. Laz. In *The Kartvelian Languages*, ed. by Alice C. Harris, Greppin, J. A. C., 395–472. The Indigenous Languages of the Caucasus 1. Delmar, New York: Caravan Books.

Holton, David, Peter Mackridge, and Irene Philippaki-Warburton. 2012. *Greek: A comprehensive grammar*. 2nd ed. London, New York: Routledge.

Keenan, Edward L. 1985. Relative Clauses. In *Language Typology and Syntactic Description. Volume II: Complex Constructions*, ed. by Timothy Shopen, 141–170. Cambridge: Cambridge University Press.

Kıral, F. 2000. Copied Relative Constructions in Khalaj. In *Studies in Turkish and Turkic Languages: Proceedings of the 9th International Conference on Turkish Linguistics*, ed. by A. Göksel and C. Kerslake, 181–188. Wiesbaden: Harrassowitz.

Lacroix, René. 2009. *Description Du Dialecte Laze d'Arhavi (Caucasique Du Sud, Turquie). Grammaire et Textes. Thèse de Doctorat*. Université Lumière Lyon 2.

———. 2012. The Multi-Purpose Subordinator in Laz. In *Clause Linkage in Cross-Linguistic Perspective: Data-Driven Approaches to Cross-Clausal Syntax*, ed. by Volker Gast and Holger Diessel, 77–104. Berlin, Boston: Mouton de Gruyter.

Matras, Yaron. 2010. Contact, Convergence, and Typology. In *Handbook of Language Contact*, ed. by Raymond Hickey, 66–85. Oxford: Blackwell.

McLoughlin, Leslie J. 2003. *Colloquial Arabic (Levantine)*. 2nd edition. London, New York: Routledge.

Nash, Léa. 2002. Pre-Nominal Relative Clauses in Georgian. presented at the Topics in Georgian Syntax, September 13, MIT.

Neocleous, Nicolaos. in preparation. Headedness in Romeyka: Continuity, Contact and Change in the Hellenic Varieties of Pontus. United Kingdom: University of Cambridge.

Öpengin, Ergin, and Geoffrey Haig. to appear. Kurmanji Kurdish in Turkey: Structure, Varieties, and Status. In *Linguistic Minorities in Turkey and Turkic Speaking Minorities of the Peripheries*. Wiesbaden: Harrassowitz.

Öztürk, Balkız, and Markus A. Pöchtrager, eds. 2011. *Pazar Laz*. München, Allemagne: LINCOM Europa.

Paul, Ludwig. 2009. Zazaki. In *The Iranian Languages*, ed. by Gernot Windfuhr, 545–86. London, New York: Routledge.

Simonian, Hovann H. 2007. *The Hemshin. History, Society and Identity in the Highlands of Northeast Turkey*. London, New York: Routledge.

Tuite, Kevin. 2004. A Short Descriptive Grammar of the Svan Language. Université de Montréal. https://www.uni-jena.de/unijenamedia/Downloads/faculties/phil/kaukasiologie/Svan%5Bslightlyrevised%5D.pdf.

Vaillette, Nathan. 2001. Hebrew Relative Clauses in HPSG. In *Proceedings of the 7th International HPSG Conference, UC Berkeley (22-23 July, 2000)*, ed. by Dan Flickinger and Andreas Kathol. CSLI online publications. http://csli-publications.stanford.edu/.

Vaux, Bert. 2007. In *The Hemshin. History, Society and Identity in the Highlands of Northeast Turkey*, ed. by Hovann H. Simonian. London, New York: Routledge.

Ophélie Gandon
ophelie.gandon@gmail.com

<center>

Turkish scrambling within single clause wh-questions[*]

Tamarae Hildebrandt

University of Michigan

</center>

1.   Introduction

Turkish is an Altaic language that exhibits agglutinative properties. Verbs bear tense, aspect, mood, and/or agreement affixes, while nouns host Case and/or agreement affixes (1).

(1)   Ben-ø hiç   sen-i-n   köpeğ-i-n-i   yıka-ma-dı-m[1]
        I-Nom never   you-Gen-2SG dog-Gen-2SG-Acc   wash-Neg-Past-1SG
        'I never washed your dog.'

Additionally, (1) demonstrates the option of genitive and accusative Case stacking on *köpeğ* plus agreement features. Turkish has relatively free word order with the default order being Subject, Object, Verb (SOV). The placement of the verb is more restricted and typically appears sentence-finally, as Turkish is a head-final language (Kornfilt 1997).

1.1   Turkish Questions

According to Cheng (1997), languages either use wh-movement or are wh-in-situ. Wh-in-situ languages mark questions through use of a question particle. Turkish is wh-in-situ (2).

(2)   *Turkish Questions*
        a. Ben kitap-ı-m-ı   arkadaş-ı-m-a   ver-di-m-mi?
           1SG book-Gen.1SG-Acc friend-gen-1SG-Dat   give-Past-1SG+Q
           'Did I give my friend a book?'

---

    b. Deniz-i   ne-ø   vur-du-ø?
        Deniz-Acc what-Nom hit-Past-3SG
        'What hit Deniz?'

    c.* Deniz-i   ne-ø   vur-du-ø-mu?
        Deniz-Acc what-Nom hit-Past-3SG+Q
        'What hit Deniz?'

In both yes/no questions and wh-questions in in-situ languages (2), question particles are not required to appear overtly. Languages such as Chinese and Japanese have a question particle in both types of questions, but question-particles only appear in yes/no questions in Turkish (2a). Wh-questions with the question particle (2c) are ungrammatical, while the omission of the question-particle (2b) yields a grammatical question.

1.2   Scrambling

Turkish exhibits a movement operation called scrambling, which moves phrases to an argument (A-scrambling) specifier position within TP or to a non-argument (A'-scrambling) specifier position within the CP domain. Ross (1967) first defined scrambling as a stylistic operation that optionally applies in the grammar to account for word order differences without any semantic effects. In contrast, Öztürk (2005) argues for two types of scrambling that differ in binding theory, weak crossover, and reconstruction effects. There may be some semantic effects on one or both argument and non-argument types of scrambling.

    I argue that wh-movement and scrambling are distinct movement operations. In English, movement of at least one wh-phrase is obligatory to check the question-feature in the syntax[2]. In Turkish, movement of a wh-phrase is due to scrambling, not question-feature checking.

    I will focus on reconstruction to differentiate A- and A'-scrambling in Turkish. A-scrambling is not reconstructable, only one scopal reading is available (3b). A'-scrambling allows for reconstruction indicating that multiple scopal readings are possible (4).

(3)   *Turkish Scope SOV and OSV*   (Öztürk 2005, ex. 137)
    a. [Ali]$_k$ t_k [bütün test-ler-e]   gir-me-di.
        [Ali]$_k$ t_k [all   test-Pl-Dat]   take-Neg-Past
        'Ali did not take all the tests.'   not > all, *all > not

    b. [Bütün test-ler-e]$_i$ [Ali] t_i gir-me-di.
        [all   test-Pl-Dat]$_i$ [Ali] t_i take-Neg-Past
        'Ali did not take all the tests.'   all > not, *not > all

(3a) shows the default reading of negation taking scope over *bütün*; however (3b) indicates that the dative object moved into [Spec, TP]. The scope of (3b) shows that *bütün* takes

---

    [2]Wh-phrases can stay in-situ in an English question like '*Who saw what?*' *Who* moves cyclically to [Spec, CP]. *What* stays in-situ until LF indicating that only one wh-phrase must move overtly to [Spec, CP].

scope over negation and reconstruction is unavailable. Based on the position of the dative object and the one and only new scopal reading, (3b) is an example of A-scrambling.

(4)   *Evidence for Reconstruction in Turkish*   (İşsever 2008, ex. 25)
   a. [Üç   kişi]$_k$ t_k herkesi   suçladı.
       [three people-Nom]$_k$ t_k everyone-Acc accuse-Past-3SG
       'Three people accused everyone.'   3 > all

   b. [Herkesi]$_i$ [üç kişi]$_k$ t_i suçladı.
       [everyone-Acc]$_i$ [three people-Nom]$_k$ t_i accuse-Past-3SG
       'Three people accused everyone.'   all > 3, 3 > all

(4) demonstrates that in Turkish scopal relationships can be read off of the surface structure. (4a) shows the default reading where *üç kişi* takes scope over *bütün*. A new reading is available in (4b), in which *bütün* scopes over *üç kişi*. This initially looks like A-scrambling, we observe that the reconstructed meaning is also available in (4b). The default meaning of (4a) can be reconstructed for (4b) in LF by interpreting *üç kişi* in its first merged position. Therefore, by hypothesis, A'-scrambling has occurred in (4b) and A-scrambling has occurred in (3b), based on the evidence from reconstruction.

1.3   The Problem

Most of the current proposals (e.g. Miyagawa 1997, 2001, 2003, 2004) regarding scrambling are sufficient to accurately account for scrambling in Turkish wh-questions. I will attempt to develop a syntactic proposal that accounts for the data (5-7). I will explain why constituents scramble and why additional projections both above and below TP are necessary.

(5)   *Accusative Object*
   a. Deniz-i   ne-ø   vur-du-ø?
       Deniz-Acc what-N hit-Past-3SG
       'What hit Deniz?'

   b.* Ne-ø   Deniz'-i   vur-du-ø?
       What-Nom Deniz-Acc hit-Past-3SG

(5a) shows that the accusative object has scrambled to a position above the nominative subject. (5b) shows that scrambling of the accusative object is mandatory with nominative wh-phrases. My fieldwork data show that a dative wh-phrase can scramble (6a), but it needn't (6b).

(6)   *Dative wh-phrase*
   a. Kim-e to top-ø   atıl-dı-ø?
       Who-Dat ball-Nom   throw-past-3SG
       'To whom was the ball thrown?'

   b. Top-ø   kim-e   atıl-dı-ø?
       Ball-Nom who-Dat throw-Past-3SG
       'To whom was the ball thrown?'

The data in (7) further complicate our understanding of scrambling, as the first three constituents may appear in any order.

(7)   *Possible Scrambling of First Three Constituents*
       Top-ø   ne-den³   Deniz'-i   vur-du-ø?
       Ball-Nom   What-Abl   Deniz-Acc   hit-Past-3SG
       'Why did the ball hit Deniz?'

Both the ablative and accusative constituents can scramble to a position above the nominative subject. The only constant in (7) is that the verb must appear sentence-finally, indicating it is not eligible for scrambling.

2.   Miyagawa on Scrambling

Miyagawa (2004) notes similarities between Japanese and Turkish concerning overt morphological Case, scope, scrambling, V-to-T movement, and wh-in-situ. He conjectures that his proposal could extend to Turkish without providing specific examples. First, I will present a discussion of Miyagawa's (1997, 2001, and 2003) analysis of Japanese scrambling. Then I will argue against the extension to Turkish because the proposal does not accurately represent one of the key differences between A- versus A'-scrambling.

2.1   Japanese Scrambling

Miyagawa (2001, 2003) argues that the Extended Projection Principle (EPP) triggers A-scrambling, while A'-scrambling is motivated by focus. EPP features are associated with different core functional categories, which include C, T, and *v* (Chomsky 2000). EPP is a required and universally strong feature (e.g. Alexiadou and Anagnostopoulou 1998) that must be checked in every language. The parameter for checking can vary with respect to verbal agreement morphology. In languages with 'rich' verbal agreement morphology, EPP must be checked through head movement. Languages lacking rich verbal agreement morphology checks EPP via phrasal movement to the specifier of the functional category carrying the EPP feature.

   Japanese lacks a rich agreement system, so phrasal movement or merger will satisfy EPP. Miyagawa's proposal concentrates on the EPP feature located on T. Following Chomsky 1995, he assumes that either DP (9) or wh-PP (10) can move to satisfy EPP. Both phrases are equidistant from T because V has raised to T. A summary of Miyagawa's proposal of EPP and scrambling in Japanese is given in (8).

(8)   *EPP and Scrambling in Japanese*
       a. A-scrambling is triggered by the EPP feature located on T.
       b. EPP can be satisfied by a DP or a wh-PP because of the equidistant
           relationship created by V-to-T raising.

---

³In Turkish, *neden* means why and it is comprised of what (*ne*) and the ablative case marker (*den*).

Miyagawa (2001) demonstrates through scopal evidence (9) that the nominative DP or the accusative DP moves to [Spec, TP] to satisfy EPP. An additional scopal reading is available in (9b) when the accusative DP satisfies the EPP feature on T, as opposed to (9a).

(9)   *Japanese Scopal Evidence*
        a.   [Zen'in-ga]$_k$ t_k sono tesuto-o   uke-nakat-ta
             [all-Nom]$_k$ t_k that test-Acc   take-Neg-Past
             'All did not take that test.'   all > not, *not > all

        b.   [Sono tesuto-o]$_i$ [zen'in-ga]$_k$ $(t_k)^4$ t_i uke-nakat-ta
             [that test-Acc]$_i$ [all-Nom]$_k$ (t_k) t_i take-Neg-Past
             'That test, all didn't take.'   not > all, (all > not)

The default interpretation of (9a) is that 'No one person took the test,' meaning that *zen'in* takes scope over negation. The nominative DP moves to [Spec, TP] thereby satisfying EPP. (9b) contains a primary and secondary reading, which suggests that Japanese can reconstruct in LF. To get the primary reading, A-scrambling has occurred—meaning *zen'in* stays in-situ, while the accusative DP *sono tesuto* to [Spec, TP] satisfying EPP. A second derivation is possible through A'-scrambling. The accusative DP moves to a position higher than [Spec, TP], as the nominative DP occupies this position. (9b) is an example of A-scrambling because of the new scopal reading, but the reconstructed reading (9b) suggests that A'-scrambling occurred.

   Like in (9b), there are two derivations possible in (10). For the primary reading, in which negation takes scope over *zen'in*, the wh-PP moves to [Spec, TP] to satisfy EPP. Additionally this movement to [Spec, TP] also satisfies the WH-feature on T, which is a part of a feature movement model Miyagawa (2004) adopts for questions.

(10)   *wh-PP Satisfying EPP*   (Miyagawa 2001, ex. 53)
        [Dare-to]$_i$ [zen'in-ga]$_k$ (t_k) t_i asoba-nakat-ta no?
        [who-with]$_i$ [all-Nom]$_k$ (t_k) t_i play-Neg-Past Q
        'With whom, all didn't play?'   not > all, (all > not)

In this feature movement model, an unvalued Question-feature is located on C and an unvalued WH-feature is located on T. The feature movement model creates a problem for A'-scrambling and the ability to account for the reconstructed reading. The nominative DP moves to [Spec, TP], while wh-PP moves to a non-argument specifier for A'-scrambling. Does the WH-feature require the wh-phrase to be in the [Spec, TP] or can probe/goal agree satisfy the WH-feature?

(11)   *The Morphological Differences between Wh-movement and Wh-in-situ*
        a.   Movement: impossible morphological separation of the wh-feature and phrase
        b.   In-situ: wh-feature morphologically separates from the wh-phrase

---

[4]The trace in parenthesis indicates the A'-scrambling structure, in which the nominative DP moves to [Spec, TP] and the other constituent moves to a specifier above TP.

All languages use wh-movement in the narrow syntax, but differences arise in the morphology (11). In wh-movement languages, the wh-phrase and the wh-feature cannot be morphologically separated from one another. The entire phrase pied-pipes to [Spec, CP] to check the Q-feature on C. In wh-in-situ languages, the wh-feature can morphologically separate from the wh-phrase. Only the wh-feature moves in the overt syntax to [Spec, CP] to check the Q-feature on C.

2.2   Problems with the Proposal

The main problem with Miyagawa's proposal concerns the use of EPP without providing a clear definition of what can satisfy this principle. Following Alexiadou and Anagnostopoulou (1998), Miyagawa assumes parametric variation for checking EPP; however with the XP parameter, it is unclear what phrases can satisfy EPP. (10) demonstrated that a wh-PP can move to [Spec, TP] to satisfy EPP because the wh-PP agrees with the WH feature on T. A regular PP is unable to satisfy EPP because the PP lacks features that agree with features on T.

If a PP lacks features to agree with a feature on T, then answers to the following questions are needed to enhance the proposal: what features are necessary for an accusative or dative DP to move to [Spec, TP]? Is it enough that the DPs have a set or subset of phi-features, even if the feature values do not match those on T? The conditions for object scrambling (Miyagawa 2003) are independent of feature matching. Rather, object scrambling depends on languages that use V-to-T raising and have morphological Case-marking to create an equidistant structure. The mismatch for requiring feature matching for a regular PP and wh-PP, while an equidistant structure for object scrambling to satisfy EPP, seems inconsistent.

Although, equidistance may not be relevant problem in Turkish scrambling since Turkish satisfies the EPP differently than in Japanese. As Turkish has a rich verbal agreement morphology, V-to-T raising satisfies the EPP (Miyagawa 2004). Where would the sentence-initial constituents like the nominative DP (7), accusative DP (5a), dative DP (6a), and locative PP (12) be located? Would they be located in [Spec, TP] or in some other position? As EPP is satisfied via head movement, phrasal movement to [Spec, TP] may be optional.

(12)   *Locative in [Spec, TP]*
        Gündüz-de   biz-ø   okul-a   gitme-ye   karar   yer-dik.
        Daytime-Loc we-Nom   school-Dat go-Dat   decision in-Sub
        'During the day we decided to go to school.'

Rather than questioning equidistance, more data is needed to support the conclusion that head movement can satisfy EPP on T in Turkish. In Japanese only wh-PPs can move to [Spec, TP] to satisfy EPP (10), but not regular PPs. If phrases can be located in [Spec, TP] and if the head movement parameter is correct, any phrase in Turkish should be able to move into [Spec, TP]. If there are restrictions on any phrase, then the head movement parameter to satisfy EPP is incorrect and would show only phrasal movement can satisfy EPP.

A second problem is that EPP is a trigger of A-scrambling, but several examples of scope ambiguity in Japanese and Turkish indicate that A'-scrambling has also occurred. The scrambling of the accusative object to a specifier above TP is A'-scrambling, as indicated by the existence of the secondary reconstructed scopal reading, which is predicted to be unavailable if this is analyzed as A-movement. The proposal in (8) would need to be modified to account for the reconstructed interpretations in (9b) and (10).
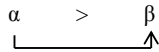
3. An Alternative Analysis

The goal of my analysis is to provide clearer definitions of how the EPP is satisfied and what phrases are involved in A'-scrambling. Among my elicitations, only A'-scrambling was present. A-scrambling exists in Turkish, but my elicitations failed to show this fact.

3.1 Theoretical Background

I assume a 'Minimalist' syntactic approach (Chomsky 2000) including the definition of probe/goal Agree (13). The DP that valued the phi-features on T, simultaneously has its Case valued by T. Then once Case is checked/deleted, the remaining features that cannot be checked via long-distance agree, mainly the EPP feature, drive phrasal movement to a specifier.

(13)   *Definition of Agree*
$$\alpha \quad > \quad \beta$$
Agree (α, β), where α is a probe and β is a matching goal, '>' is a c-command relation and matching uninterpretable features of α and β are checked/deleted.

(Chomsky 2000, p.122)

EPP features are located on a variety of syntactic heads (Chomsky 2000)—specifically on T, Force, Focus, and Topic within my analysis. The EPP features can only be satisfied through phrasal movement to the respective specifier. EPP on T will only attract the nominative DP because it matches all of the phi-features and is the closest to T. EPP features on the other heads will attract the XP that matches the feature (+Q, +Foc, +WH) borne by the head.

(14)   *Expanded CP Layer (Rizzi 1997)*
[ForceP [TopP [FocP [TopP [FinP TP]]]]]

The use of Force, Focus, and Topic originates from Rizzi's (1997) expanded CP layer proposal (14). Within Rizzi's analysis, Focus and the wh-phrase are incompatible— meaning that wh-phrases move to or through [Spec, FocP], but cannot remain in this syntactic position. I assume that X cannot move to a position P, if X is not allowed to remain in P. Since the wh-phrase is not allowed to remain in [Spec, FocP], the wh-phrase cannot move to [Spec, FocP]. Non-wh-phrases marked with +Foc can move into [Spec, FocP] to satisfy EPP. Wh-phrases are marked with +WH and move to [Spec, TopP].
I assume that a null wh-operator (OP) is generated in the specifier of a wh-phrase (İşsever 2008). This operator has the [+Q] feature and moves to [Spec, ForceP] to covertly satisfy

EPP, which maintains my assumption of wh-in-situ languages (section 1.1)[5]. The movement of the wh-OP also assumes that the feature movement model (Miyagawa 2004) is correct, since the wh-OP morphologically separates from the wh-phrase.

3.2 Scrambling Derivations

Following the theoretical framework proposed above, I will show a derivation without scrambling, two examples of one constituent scrambling, and one example of two constituents scrambling within a single clause wh-question. Only unvalued feature that have been valued through agree (13) are shown in the trees for readability. Inherent features listed on the heads have been omitted. Lastly, all trees show the movement of the OP to [Spec, ForceP] to satisfy the EPP on the Force head, although this movement occurs covertly in LF (Cheng 1997).

(15) *Wh-Question without* A'-*scrambling*

```
                    ForP
                   /    \
                 OPₐ     For'
                 /\     /    \
                Ø     TP      For
                     /  \     [+Q]
                   DPᵢ    T
                   /\    /      \
                 Top   vP    Vᵦ+vᵥ+T
               [+NOM]  /\   Vur+Ø+PAST
                      tᵢ   v'
                          /  \
                        VP    t_c
                         |
                         V'
                        /   \
                      PP     V'
                       |    /   \
                       P'  DP    tₕ
                      / \  Deniz-i
                    DP   P
                    /\  [+ABL] [+ACC]
                  tₐ  D
                      Ne
```
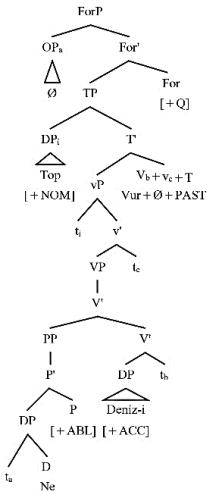
(15) is the S-structure analysis of (7) and shows that no scrambling has occurred. I assume agree (13) precedes any phrasal movements to ensure feature matching for EPP. For example in (15), T will probe into its c-command domain and find the DP located in [Spec, vP]. The DP will value T's phi features, and simultaneously its Case feature will be valued with nominative Case. Since the complete set of phi-features match, the nominative DP targets [Spec, TP] thereby satisfying EPP. The same applies for the unchecked Q feature on Force. Force will probe into its c-command domain and find its goal, *neden*, which is marked as [+Q]. As long as phrasal movement follows agree (13), no other stipulations are

---

[5]It is possible that the null wh-OP in Turkish could move overtly in the narrow syntax. With this phonologically null element, it is difficult to show which option should be used as opposed to the other.

required. If phrasal movement to a specifier does not occur, then the derivation will crash due to an unchecked EPP feature.

I define A'-scrambling as phrasal movement to a specifier motivated by EPP features on Foc or Top within the expanded CP layer. (16a) shows the movement of the accusative DP to [Spec, FocP], while (16b) shows the movement of the dative wh-phrase to [Spec, TopP].

(16) *A'-Scrambling: One Constituent*
  a. Accusative DP   b. Dative DP
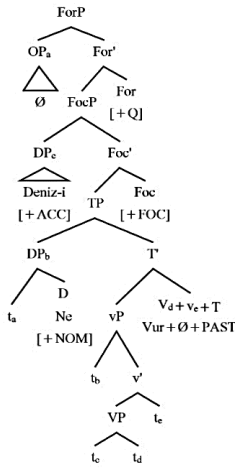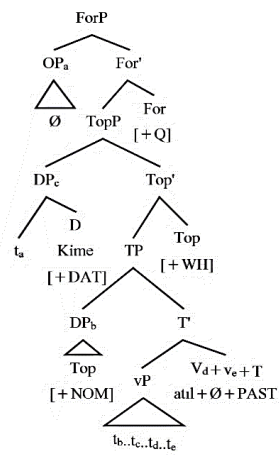
(16a) is the representation (5a). The accusative DP, *Deniz'i*, bears the +Foc feature. Once agree (13) occurs to check/delete features and Case, then movement to satisfy EPP via phrasal movement follows. The nominative wh-phrase moves to [Spec, TP] because its phi-features match the phi-features on T. The accusative DP will move to [Spec, FocP] because it matches the focus feature located on the Focus head. (6a) matches the syntactic tree in (16b) and is nearly identical to (16a). The key difference between (16a) and (16b) is the scrambling of a non-wh- versus wh-phrase. The dative wh-DP is marked with a +WH feature, which agrees with the unvalued WH feature on Top. Following agree, the dative wh-DP targets [Spec, TopP] to satisfy EPP on Top. Within both derivations the wh-OP is generated in the specifier of the wh-phrase. The wh-OP bears the [+Q] feature and will value this feature on Force. The EPP feature on Force is satisfied by covert phrasal movement of the wh-OP to its specifier.

Two phrases, TopP and FocP, appear within the expanded CP layer to account for A'-scrambling. The EPP features on these categories motivate movement to the respective specifiers. Both phrases are required above TP to permit the scrambling of two separate constituents within the same wh-question (17).

(17) *A'-Scrambling[6]: Two Separate Constituents*

    a. [Ne-den]$_j$ [Deniz'-i]$_k$ [Top-ø]$_i$ t$_i$ t$_j$ t$_k$ vur-du-ø?

       [What-Abl]$_j$ [Deniz-Acc]$_k$ [Ball-Nom]$_i$ t$_i$ t$_j$ t$_k$ hit-Past-3SG

      'Why did the ball hit Deniz?'

    b.
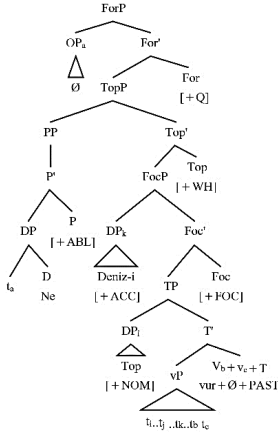
The Foc head merges to TP and the Top head merges directly to FocP in (17b). After agree (13) checks and deletes Case and phi-features, movement to satisfy EPP can occur. The XPs target specifiers of phrases, which match the features on the XP. The nominative DP can only move to [Spec, TP] because it is the closest DP that matches the entire set of phi-features (person, gender, and number). The accusative DP moves to [Spec, FocP], while the ablative wh-phrase moves to [Spec, TopP]. In LF, the wh-OP moves covertly to satisfy EPP on Force.

(18) *Low Focus Projection*

    [Top-ø]$_b$ [Deniz'-i]$_a$ t$_b$ ne-den   t$_a$ vur-du-ø?

    [Ball-Nom]$_b$ [Deniz-Acc]$_a$ t$_b$ what-Abl t$_a$ hit-Past-3SG

    'Why did the ball hit Deniz?'

Most of the preceding data indicate that A'-scrambling only occurs within the expanded CP layer; however, (18) demonstrates that a low focus projection is needed. The accusative DP is merged above the vP, but below [Spec, TP]. A low focus position for A'-scrambling would fit within phase-based theory and phase-edges. Before the phase-head complement is transferred, the accusative DP would move to the phase edge. The accusative DP can stay in this low focus position, or the accusative DP can move out of the phase-edge into [Spec, FocP]—creating a high focus position within the expanded CP layer. More research is needed to determine the extent of the empirical adequacy of such an approach.

---

   [6]Another instance of double A'-Scrambling is *Deniz'i neden Top vurdu?* The expanded CP layer is slightly different from (18a). The TopP would merge directly to TP, instead of directly to FocP as in (18b).

4. Conclusion

Miyagawa's proposal on scrambling in Japanese may extend to Turkish, but the EPP and A'-scrambling need to be more clearly defined in his approach. Within my proposal, I focused on providing a clear syntactic analysis of A'-scrambling in Turkish. I defined A'-scrambling as phrasal movement to [Spec, FocP] or [Spec, TopP] motivated by EPP features on Foc or Top. Phrasal movement to satisfy EPP can only occur if the head of the phrase and the XP share a feature or a set of phi-features (person, gender, and number).

The use of FocP and TopP within the expanded CP layer (Rizzi 1997) accounts for single (16) and multiple (17) instances of scrambling within a single wh-question. Only non-wh-phrases marked with +Foc can move into [Spec, FocP] to satisfy EPP. Wh-phrases are marked with +WH and move to [Spec, TopP]. The Force head bears an unvalued question feature and an EPP feature. A null wh-OP, generated in the specifier of a wh-phrase, values and deletes the Q-feature. The OP satisfies the EPP in LF by covertly moving to [Spec, ForceP].

I assume that agree (13) precedes phrasal movement, but no other stipulations are needed or required. If an unchecked EPP feature remains in the derivation due to a lack of phrasal movement, then the derivation will crash.

References

Alexiadou, A. & E Anagnostopoulou.1998.Parameterizing AGR: Word Order, V-Movement & EPP-checking. *Natural Language & Linguistic Theory,*16: 491-539.
Carnie, A. 2007. *Syntax: A Generative Introduction*. Malden, MA: Blackwell.
Cheng, L. 1997. *On the Typology of Wh-questions*. New York: Garland.
Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
Chomsky, N. 2000. Minimalist Inquiries: the Framework. In *Step by step: essays on minimalist syntax in honor of Howard Lasnik*, ed. by R. Martin, D. Michels, and J. Uriagereka, 89- 155. Cambridge, Massachusetts: MIT Press.
İşsever, S. 2008. EPP-driven Scrambling and Turkish. In *Ambiguity of Morphological and Syntactic Analyses*, 27-41.
Kornfilt, J. 1997. *Turkish*. London: Routledge.
Miyagawa, S. 1997. Against Optional Scrambling. *Linguistic Inquiry,* 28:1-26.
Miyagawa, S. 2001. The EPP, Scrambling, and Wh-in-situ. In *Ken Hale: A Life in Language*. ed. by Michael J. Kenstowicz. Cambridge, MA: MIT, 293-338.
Miyagawa, S. 2003. A-Movement Scrambling and Options. In *Word Order and Scrambling*. ed. by Simin Karimi. Malden, MA: Blackwell Pub., 177-200..
Miyagawa, S. 2004. Wh-in-Situ and Scrambling in the Context of Comparative Altaic Syntax. *MIT Working Papers in Linguistics,* 46: 15-32.
Öztürk, B. 2005. *Case, Referentiality and Phrase Structure*. Philadelphia: John Benjamins.
Rizzi, L. 1997. The Fine Structure of the Left Periphery. In *Elements of Grammar. Handbook of Generative Syntax*. Kluwer: Dordrecht, 281-337.
Ross, J. R. 1986. *Infinite Syntax*. Norwood, NJ: ABLEX.

Tamarae Hildebrandt
tamhil@umich.edu

Reanalyzing Indo-Iranian "stems": A case study of Adıyaman Kurmanji[*]

Laura Kalin & Ümit Atlamaz

University of Connecticut, Rutgers University

1.   Introduction

Indo-Iranian languages present several challenges for recent theoretical approaches to morphology (Halle & Marantz 1993, *et seq*) and agreement/case splits (Coon 2010, *i.a.*). First, the Indo-Iranian verbal system revolves around two verb "stems"—traditionally labeled the "present stem" and the "past stem"—whose semantics are difficult to pin down, leading to the claim that the stems have a morphomic distribution, appearing in fixed constructions (Haig 2008). Second, the form of the stems is highly opaque and irregular, making decomposition into morphemes a challenge. Finally, these stems form the basis of various sorts of case and agreement splits, despite the lack of a clear syntactic or semantic trigger.

In this paper we undertake a case study of one Indo-Iranian language, Adıyaman Kurmanji (AK), a dialect of Kurdish spoken in the town of Adıyaman in southeastern Turkey. We closely examine both the semantics and morphology of verb stems in AK, and argue that the above challenges are only apparent. In particular, we propose (i) a morphological analysis of these stems that opposes a ∅-marked form ("present stem") with an overtly-suffixed form ("past stem"), §2, and (ii) a semantic analysis of this morphology, with the null suffix as nonpast tense, and the overt suffix elsewhere (past tense, nonfinite tense), §3.

This compositional analysis, combined with a variety of morphosyntactic evidence, leads us to several surprising observations. The puzzle we take up in detail in §4 is that T behaves morphologically and syntactically as though it were below Asp in AK. Two additional puzzles that we note but do not explore in detail are that AK exhibits an extremely rare morpheme order within the complex verb, Asp-V-Tns (Julien 2002), and that split ergativity in AK is conditioned by tense, which has been claimed to be unattested (Salanova 2007, Coon 2013). We explore various possible explanations for the special position and role of tense in AK, but ultimately leave these new puzzles open for further research.

2.   Morphological breakdown

AK is an SOV but predominantly head-initial language with frequent *pro* drop (Atlamaz 2012). Like other Indo-Iranian languages, the verbal system of AK revolves around two so-called verb "stems". The term "stem" here carries with it two main implications (Aronoff 1994, 2012, *i.a.*). First is the implication that these verb forms are stored in the lexicon, i.e., stems are not broken down into smaller pieces. Second, stems—since they are stored in the lexicon—can have a distribution that is purely morphologically-determined, or "morphomic"; in other words, the choice of stem can be completely disconnected from syntactic and semantic factors. The specific claim with respect to Indo-Iranian languages is that the "present stem" and the "past stem" are chosen from the lexicon as bases for further verbal morphology, with the choice of stem conventionalized across constructions, with the synchronic distribution due to diachronic factors (Haig 2008).

In this section, we pursue a morphological analysis of the verb stems in AK by isolating a piece of morphology that differentiates the stems. We postpone a discussion of the semantics of this stem-differentiating morpheme to §3, instead using the stems' traditional labels—"present stem" and "past stem"—throughout this section.

At first glance, there seems to be no consistent morphological relation between the present stem and the past stem. Some common verbs come in suppletive stem pairs, (1), while other stem pairs differ in unpredictable phonological material, (2), and yet others do not differ at all, (3).

(1)

	"present stem"	"past stem"
'go'	her	çü
'say'	we	go
'see'	wun	di
'come'	e	hot

(2)

	"present stem"	"past stem"
'eat'	x	xor
'catch'	g	geşt
'give'	d	do
'do'	k	kır

(3)

	"present stem"	"past stem"
'sew'	drü	drü
'die'	mır	mır
'chew'	cü	cü
'cry'	gri	gri

Looking further than these common and irregular verbs, however, two more consistent patterns emerge. A number of verbs form the past stem by adding *-i* to the present stem, (4), and a number of verbs form the past stem by adding *-t* to the present stem, (5).

(4)

	"present stem"	"past stem"
'boil'	kel	keli
'rain'	bor	bori
'steal'	dız	dızi
'buy'	kırr	kırri

(5)

	"present stem"	"past stem"
'milk'	do	dot
'hear'	biz	bist
'hold'	gr	gırt
'want'	xaz	xast

As is evident from a quick glance through (1)–(5), the grouping of verbs into different types of stem pairs is not semantically or phonologically conditioned.

We can make a number of concrete observations based on these verb stem pairs. First, whenever there is a clear and non-null phonological relation between the stems, the past stem always builds on the present stem, with the additional segments in the past stem following the present stem. The two most regular and common strategies for deriving the past stem from the present stem are adding *-i* or *-t*. We conjecture, then, that the present stem reveals the phonological form of the verb root, while the past stem consists of the verb root plus a suffix, which we will take to be *-i* in its default/elsewhere form.[1] For semantic reasons discussed in the following section, we also posit that the present stem bears a null suffix; both verb stems can thus be broken down into a root and a suffix. A number of representative vocabulary items are given in (6):[2]

(6)   *Some relevant vocabulary items (first pass)*

"present" $\rightarrow$ -$\emptyset$

"past" $\rightarrow$ -t   /]$_V$___ ; where V = {$\sqrt{\text{MILK}}$, $\sqrt{\text{HEAR}}$, $\sqrt{\text{WANT}}$ ...}

"past" $\rightarrow$ -$\emptyset$ /]$_V$___ ; where V = {$\sqrt{\text{SEW}}$, $\sqrt{\text{DIE}}$, $\sqrt{\text{CHEW}}$ ...}

"past" $\rightarrow$ -or   /]$_V$___ ; where V = {$\sqrt{\text{EAT}}$}

"past" $\rightarrow$ -eşt   /]$_V$___ ; where V = {$\sqrt{\text{CATCH}}$}

"past" $\rightarrow$ -i

$\sqrt{\text{BOIL}}$ $\rightarrow$ kel

---

[1]It is difficult to determine with any certainty whether *-i* or *-t* is the elsewhere form of the suffix that forms the past stem, since both exponents appear frequently and neither picks out a natural class of verbs on semantic or phonological grounds. We have chosen *-i* as the elsewhere allomorph based on the fact that the suffixation of *-t* is often accompanied by small phonological changes in the verb root, while *-i* is typically not accompanied by such changes, cf. (4)–(5). Nothing hinges on this choice.

[2]We treat the suppletion cases, (1), as portmanteaux, because these verbs never co-occur with additional (overt) past stem morphology. We analyze portmanteaux as resulting from insertion at non-terminal nodes (Caha 2009, Radkevich 2010), but nothing hinges on this; we could instead take root suppletion to occur in the context of the "past stem" suffix, with a null allomorph of this suffix occurring with the suppletive root. Finally, if we are correct in taking both verb stems to contain the verb root and an additional piece, then it is important to note that there is nothing about (1) that tells us for certain whether it is the "past stem" or "present stem" that is the elsewhere (non-portmanteau) form of the verb root, or whether both are portmanteaux.

$\sqrt{\text{MILK}}$ → do
$\sqrt{\text{SEW}}$ → drü
$\sqrt{\text{EAR}}$ → x
$\sqrt{\text{CATCH}}$ → g
$\sqrt{\text{GO}}$ → her
$\sqrt{\text{GO}}$,"past"   → çü

As a first step, then, we have shown that the verb stems break down into predictable pieces.

3.   Semantic breakdown

In the previous section, we proposed that the two verb stems in AK consist of the verb root and a suffix: $\sqrt{\text{V}}$-∅ for the "present stem", $\sqrt{\text{V}}$-*i* for the "past stem" (plus a number of other allomorphs of this suffix, cf. (6)). We are now in a position to investigate the semantic range of these verb stems, in order to see whether a consistent semantics can be attributed to the different pieces we have identified, thereby nullifying the need for an appeal to "stems" as a theoretical primitive in the system. Our core observations in this section will be that the verb stems are distributed based on tense, but not aspect, corresponding closely (but not perfectly) to their traditional labels, as schematized in (7).

(7)   a.   "Present stem": V-TENSE.NONPAST   (abbrev: T.NONPST)
       b.   "Past stem": V-TENSE.ELSEWHERE   (abbrev: T.ELSE)

The null suffix in the present stem consistently expresses nonpast tense, while the overt suffix that forms the past stem spells out both past tense and nonfinite tense, hence it is the elsewhere exponent of T. For clarity and consistency, we adopt the proposed glosses in (7) (abbreviated as indicated) in the examples that follow.[3]

3.1   The "present stem"

There are three main uses of the present stem, (8).

(8)   a.   *Present habitual:* IMPF-V-T.NONPST
              Ez   çay-ê   **dı-kırr-∅-ım**.
              I.NOM tea-OBL IMPF-buy-T.NONPST-1SG
              'I buy tea.'

       b.   *Present progressive:* IMPF-V-T.NONPST-COP
              Ez   çay-ê   **dı-kırr-∅-ım-e**.
              I.NOM tea-OBL IMPF-buy-T.NONPST-1SG-COP
              'I am buying tea.'

---

[3]There is also a copula that will surface in the examples in this section. There are in fact two copulas in AK, the "present copula" and the "past copula". Just like the "present" and "past" labels for the stems, however, it is not clear that these are informative labels. Since we will only be dealing with the "present copula" in this paper, we will simply gloss it as COP.

c. *Future:* AUX + SBJN-V-T.NONPST

Ez   dıkê çay-ê   **bı-kırr-∅-ım**.   (dı-k-e = IMPF-do-COP)[4]
I.NOM AUX tea-OBL SBJN-buy-T.NONPST-1SG
'I will buy tea.'

The present progressive, (8b), can get a future interpretation in matrix contexts when it appears with future adverbials e.g., 'tomorrow' *sıwe*. However, the present stem in all of its forms is ungrammatical in matrix contexts with past tense adverbials, e.g., 'yesterday' *dhıni* and 'two hours ago' *dı saata ber ve*.

Thus far the best characterization of the present stem is that it encodes nonpast tense. Looking further to embedded clauses, we can see that the present stem in fact expresses a *relative* nonpast tense. Under a matrix past verb, the present stem is interpreted as cotemporal with the matrix verb phrase or temporally following the matrix verb phrase, (9). (Though not shown here for space reasons, this is also true for the other forms of the present stem.)

(9)   Mı   go-∅   [kı   çay-ê   **dı-kırr-∅-ım-e**   ].
      I.OBL say.T.ELSE-3SG   that tea-OBL IMPF-buy-T.NONPST-1SG-COP
      'I said that I was buying/was going to buy tea.' (*saying before/cotemp. with buying*)

Finally, as can be seen throughout (8), the present stem is always prefixed. Following Giorgi & Pianesi (1997) among others, we take present tense to be incompatible with perfective aspect, hence the obligatoriness of the imperfective prefix, *dı-*, which appears in (8a)–(8b). In addition, there seems to be a surface requirement in AK that verbs cannot bear more than one prefix: in environments where the subjunctive is required, as in (8c), the imperfective prefix cannot appear, and instead we see the subjunctive prefix *bı-*.

Our conclusion about the present stem, then, is that it expresses a relative nonpast tense. We therefore propose that the present stem consists of the verb root with a null suffix that contributes this relative nonpast tense.

3.2   The "past stem" in finite clauses

The default, context-free interpretation of the "past stem" is a simple past tense, (10).

(10)   *Simple past:* V-T.ELSE
       Mı   çay   **kırr-i-∅**.
       I.OBL tea.NOM buy-T.ELSE-3SG
       'I bought tea.'

This simple past seems to be interpreted as perfective, a point we will return to later. With the addition of the imperfective prefix, the interpretation is past imperfective, (11).

---

[4]It is not clear whether this decomposition of the future auxiliary is correct synchronically, or whether it is now an unanalyzed whole. Since the conveyed meaning is neither imperfective nor present tense, and since the auxiliary does not need to agree with the nominative subject (as present stems normally do), we simply gloss the future auxiliary as AUX, taking it to be morphologically simplex.

(11)   *Past progressive/habitual:* IMPF-V-T.ELSE
        Mı   çay   **dı-kırr-i-∅**.
        I.OBL tea.NOM IMPF-buy-T.ELSE-3SG
        'I was buying tea.' / 'I used to buy tea.'

The past stem, (10)–(11), is grammatical with past adverbials in matrix clauses, but not with future adverbials, the converse of the present stem.[5,6]

Just like nonpast tense morphology in AK, (9), past morphology expresses a relative tense. Under a matrix past tense, the past stem is interpreted relative to that past time, (12).

(12)   Mı   go-∅   kı çay **(dı-)kırr-i-∅**.
        I.OBL say.T.ELSE-3SG that tea IMPF-buy-T.ELSE-3SG
        'I said that I bought/was buying/used to buy tea.'   (*saying after buying*)

Our first hypothesis is that the suffix that forms the past stem encodes relative past tense.

A logical question to ask at this point is whether the past stem is in fact perfective, in addition to (or instead of) being past. However, while the (non-imperfective-marked) past stem is interpreted as perfective by default, cf. (10), it does not have the expected entailments of a true perfective, as shown in the felicitous continuations of (13a) in (13b–c).

(13)   a.   Ehmet dhıni   xanı çê **k-ır-∅**.
             Ahmet yesterday house good do-T.ELSE-3SG
             'Ahmet built a house yesterday.'

        b.   . . . hema, soğ   ne-k-ır-∅.
                  but   finish NEG-do-T.ELSE-3SG
             '. . . but (he) didn't finish (it).'   (*completedness not entailed*)

        c.   . . . hema, hin jı çê   dı-k-∅-ê.
                  but   still too good IMPF-do-T.NONPST-COP.3SG
             '. . . but (he) is still building (it).'   (*boundedness not entailed*)

While the entailment of completedness is not a reliable indicator of high perfective aspect crosslinguistically, the entailment of boundedness is (Altshuler 2015). The past stem in AK thus fails the crucial test in (13c), and so we conclude that it is not perfective.[7] We instead

---

[5]The plain past stem can also have a reading that is something like "about to" when combined with the present adverbial *nha* "now". This interpretation is restricted to predicates that can be construed to have a process and subsequent logical culmination. We take this to be a pragmatically-available interpretation, based on a culmination being so imminent that it can be said to already have happened. Thus, such sentences are only felicitous when the process is already underway. We thank Sabine Iatridou for helpful discussion.

[6]In counterfactuals, the verb takes the form of a past stem bearing a counterfactual suffix, and in this case the past stem can appear with future adverbials. We take this to be part of a larger pattern whereby past tense morphology is co-opted in counterfactuals (Iatridou 2000, Bjorkman & Halpert 2012, *i.a.*).

[7]When a finite verb lacks overt tense (i.e., in the present stem), the interpretation is present tense; we therefore take there to be a null present tense morpheme. When a finite verb lacks overt aspect, the interpretation is *not* perfective; we therefore do not posit a null perfective aspect morpheme.

take the default perfective interpretation of simple past clauses to be a result of implicature, triggered by the absence of the imperfective prefix.

Our preliminary conclusion is that the suffix that forms the past stem contributes a (relative) past tense in finite clauses. We will see in the next section, however, that this suffix does not uniquely expone past tense, but rather is underspecified for tense.

3.3   The "past stem" in non-finite clauses

There is a complication to the clean opposition of nonpast (null suffix) vs. past (overt suffix): all participles and nominalizations in AK are formed on the past stem. This is notable because these participles and nominalizations do not semantically encode tense in AK. Adjectival participles, (14), as well as nominalizations, (15), are compatible with both past and future readings.

(14)   *Past and future readings of adjectival participles:*
   a.   Me   beq-ê   **kuşt**-i   di
        we.OBL frog-EZ kill.T.ELSE-PART see.T.ELSE
        'We saw the killed frog.'   (*frog already dead at present*)

   b.   Em   dıkê beq-ê   **kuşt**-i   bı-wun-∅-ın
        we.NOM AUX frog-EZ kill.T.ELSE-PART SBJN-see-T.NONPST-PL
        'We will see the killed frog.'   (*frog may or may not be dead yet at present*)

(15)   *Past and future readings of nominalizations:*
   a.   Me   **fırr-i**-n-ê   Mehemed-ê   mezı k-ır
        we.OBL fly-T.ELSE-NMLZ-EZ Mehemed-OBL watch do-T.ELSE
        'We watched Mehemed's flying/flight.'   (*flying/flight in the past*)

   b.   Em   dıkê **fırr-i**-n-ê   Mehemed-ê   mezı
        we.NOM AUX fly-T.ELSE-NMLZ-EZ Mehemed-OBL watch
        bı-k-∅-ın
        SBJN-do-T.NONPST-PL
        'We will watch Mehemed's flying/flight.'   (*flying/flight in the future*)

Our conclusion based on the appearance of the past stem in nonfinite environments is that the exponent of past tense, *-i*, is not exclusively used for past tense. Rather, this exponent is the elsewhere morphological realization of T, (16).

(16)   *Updated vocabulary items*
   a.   T[PRES] → -∅
   b.   T   → -*i* (and other allomorphs, cf. (6))

Both T[PAST] and T[NONFIN] are realized with the same vocabulary item, namely, the suffix that forms the so-called past stem. Note that this is a marked morphological system,

as the underspecified/elsewhere morphology is richer (overt, with many allomorphs) than the non-underspecified morphology, which is consistently null.

3.4 Interim summary

We summarize the morphology of the core verb forms we have covered in (17) and (18):

(17) *Core verb forms built on the "present stem"*
    a. present habitual   IMPF-V-T.NONPST
    b. present progressive IMPF-V-T.NONPST-COP
    c. future   Aux + SBJN-V-T.NONPST

(18) *Core verb forms built on the "past stem"*
    a. simple past   V-T.ELSE
    b. past habitual/progressive   IMPF-V-T.ELSE
    c. participles/nominalizations V-T.ELSE-PART/NOMZN

The "present stem" and "past stem" approximate the right labels, but: (i) the morphology is fully compositional and predictable, so there is no need for "stems" as a theoretical primitive, nor for an appeal to a morphomic distribution of the stems; (ii) the "present stem" is nonpast; and (iii) the "past stem" represents the verb root bearing the elsewhere morphology of T, surfacing both in past tense and nonfinite environments.

4. Locating T in Adıyaman Kurmanji

We have now identified the core pieces of aspectual and temporal verbal morphology in AK: present tense T as -∅, elsewhere T as -i (etc.), imperfective aspect as dı-, and subjunctive mood as bı-. Combining this with a variety of morphosyntactic evidence points us to a rather surprising conclusion about the syntactic position of T in AK. In particular, T seems to be (syntactically and/or morphologically) below Asp, in conflict with the expected superiority of T over Asp. In this section, we present the morphosyntactic evidence for the low position of T, and we explore various ways of situating this finding theoretically.

4.1 Evidence for the low position of T

There are four pieces of evidence that suggest a close relationship between T and *v*+V. First, the choice of verb root affects the choice of an allomorph for T:

(19)

	V-T.NONPST	V-T.ELSE
'eat'	x-∅	x-or
'catch'	g-∅	g-eşt
'die'	mır-∅	mır-∅
'steal'	dız-∅	dız-i
'hear'	biz-∅	bis-t

As can been seen in (19), it is impossible to pick the right phonological form of T without knowing the identity of V. Second, for a number of verbs, the choice of tense in T affects the form of the verb:

(20)

	V.T.NONPST	V.T.ELSE
'go'	her	çü
'say'	we	go
'see'	wun	di

As seen in (20), it is impossible to pick the right phonological form of the verb without knowing the identity of T. Under the fairly standard assumption that there are locality restrictions on allomorph-conditioning (and portmanteaux), T and V must be local.

But how local exactly are T and V to each other? V and *v* seem to be spelled out together in the normal case and are inseparable as V+*v* (or *v* is typically null). But, when *v* is overtly pronounced separately from the verb root as a causative morpheme, *v* conditions the form of T rather than V doing so. Take, for example, the verb 'boil' *kel*. The regular past stem is *kel-i*. When the past stem of 'boil' bears the causative *v* suffix *-on/-ın*, the elsewhere T is instead spelled out as *-d* (determined by CAUS) and appears outside the causative morpheme, as seen in *kel-on-d* (**kel-on-i*, **kel-i-on*). Causative morphology thus shows us that T is separated from the verb root by (at least) *v*, both linearly and with respect to allomorph-conditioning.

Unlike causative morphology, however, V-T allomorphy is not blocked by an overt exponent for Asp. Imperfective aspect blocks neither the verb root from conditioning the form of T, (21), nor T from conditioning the form of V, (22), cf. (19)–(20).

(21)

	IMPF-V-T[PRES]	IMPF-V-T(ELSE)
'eat'	dı-x-∅	dı-x-or
'catch'	dı-g-∅	dı-g-eşt
'die'	dı-mır-∅	dı-mır-∅
'steal'	dı-dız-∅	dı-dız-i
'hear'	dı-biz-∅	dı-bis-t

(22)

	IMPF-V.T[PRES]	IMPF-V.T(ELSE)
'go'	dı-her	dı-çü
'say'	dı-we	dı-go
'see'	dı-wun	dı-di

Further, imperfective Asp never conditions V's form, nor does V condition Asp's form.[8]

It is worth pausing for a moment to fully understand the implications of this data. If Asp were between T and V (as is standard), then we would have the structure in (23).

---

[8]The only apparent exception to this generalization involves vowel-initial verb roots which trigger surface allomorphy of *dı*, e.g., "bring" *dı-un* → *tin*; "come" *dı-e* → *te*. We take this to be a purely phonological process, and so do not consider it to be a counterexample.

(23)   T > Asp > *v*+V

Taking as input the structure in (23), accounts of allomorph selection that rely on strict
adjacency (e.g., Bobaljik 2012) cannot account for any of (19)–(22), since T is separated
from *v*+V by Asp. Accounts that allow allomorph conditioning via "spans" (Svenonius
2012, Merchant 2015) also cannot account for any of (19)–(22), since Asp is "otiose"
(irrelevant for the choice of allomorph) and so cannot be included in a conditioning span.
Accounts that allow "pruning" of null nodes (e.g., Embick 2003, 2010, Calabrese 2012)
can account for (19)–(20), but not (21)–(22), since Asp is overt and thus cannot be pruned,
leaving T and *v*+V still non-local. These problems all disappear if Asp is in fact above T in
AK. We return to this point in §4.2.3.

The final piece of evidence for the low position of T comes from the morphology that is
(im)possible in nominalizations. In particular, we know from the overt suffix that surfaces
in participles and nominalizations that (nonfinite) T is present in nominalizations, cf. (14)–
(15). Further, this nonfinite T is obligatory, (24), and Asp is disallowed, (25).[9]

(24)   *Participles/nominalizations not allowed to lack T*
        a. *beq-ê **kuj**-i
            frog-EZ kill-PART
            Intended: 'the killed frog'

        b. ***x**-ın-ê   Mehemed-ê
            eat-NMLZ-EZ Mehemed-OBL
            Intended: 'Mehemed's eating'

(25)   *Participles/nominalizations not allowed to have Asp*
        a. beq-ê **(*dı-)kuşt**-i
            frog-EZ IMPF-kill.T.ELSE-PART
            'the killed frog'

        b. **(*dı-)x-or**-ın-ê   Mehemed-ê
            IMPF-eat-T.ELSE-NMLZ-EZ Mehemed-OBL
            'Mehemed's eating'

For nominalizations to include T but not Asp, there must be a syntactic constituent in the
clause that contains T and lacks Asp.

To summarize: (i) V conditions allomorphy on T; (ii) T conditions allomorphy of V;
(iii) this conditioning is not blocked by overt Asp; and (iv) nominalizations include T but
not Asp. All of these facts point to T being closer to V than Asp is at some point in the
syntactic or morphological components of the derivation.

---

[9]It is impossible to rule out the presence of an empty/null Asp in (25). However, participles and nominal-
izations in many languages preserve aspectual distinctions (Alexiadou 2001, Aikhenvald 2011, *i.a.*), so we
might expect Asp to be able to be imperfective, and thus overt, if the Asp projection were present here.

4.2 The puzzle and some solutions

Why is Asp being above T surprising or problematic? Compositional semantic analyses of tense and aspect crucially rely on Asp mediating between T and the VP event/situation (e.g., Klein 1994, Stowell 1995, Demirdache & Uribe-Etxebarria 2000, Iatridou et al. 2001), (26a). Simplifying somewhat, the role of Asp is to relate the time argument of the VP (the event time) to Asp's own time argument (the reference/assertion time), which T in turn relates to its time argument (the utterance time in matrix clauses). If Asp and T are structurally reversed, (26b), they are in the wrong relative positions to do their jobs.

(26) a. Standard hierarchy: T > Asp > V
      b. Reversed hierarchy: Asp > T > V

How can we reconcile the morphosyntactic evidence—which points to (26b)—with the semantic evidence (a standard tense/aspect system)—which points to (26a)?

4.2.1 A semantic solution

Perhaps the most obvious solution is to design a complex semantics for tense and aspect in AK, such that the apparent problem of the reversed hierarchy goes away. This solution would adopt (26b) and simply re-engineer the denotations of T and Asp in AK to get the clause-level semantics right. This is in principle possible, and essentially would involve allowing T and Asp to bypass their local time arguments for more distant ones, such that T could still relate the reference time to the utterance time, and Asp could still relate the reference time to the event time.

A number of problems arise for this solution. One problem is that it adds significant power to the semantics of tense and aspect, by taking away the restriction that each temporal or aspectual operator relate its own time argument in one of several fixed ways to that of its complement. Another problem is that we would expect languages to be able to freely choose between (26a) and (26b), rather than nearly always choosing (26a). Perhaps most importantly, under this analysis it is entirely accidental that the complex denotations of T and Asp in the reversed hierarchy in AK end up mimicking exactly the output of the (comparably) simple denotations of T and Asp in the standard hierarchy.

4.2.2 A syntactic solution

A very different sort of solution involves syntactic licensing, in the style of Stowell (2007). The basic idea here is that tense morphology need not surface in the same position as tense interpretation, but rather the morphology could appear in a low position, with its form conditioned/licensed by a temporal operator higher in the structure, in T. For Stowell (2007), this is the case in English for low past tense morphology on the verb licensed under a high Past operator in T. For AK, since the more specific morphology is nonpast tense (the present stem), it would have to be low nonpast tense morphology licensed under a

high Nonpast operator, with the elsewhere morphology (the past stem) occurring when the Nonpast operator is absent.

Such an analysis can maintain the standard hierarchy, (26a). However, one of the primary motivations for this type of analysis of tense morphology is the presence of sequence-of-tense effects. We have already seen in AK, however, that there are no sequence-of-tense effects, neither for present tense, (9), nor for past tense, (12). A Stowell-type analysis therefore does not fare well for AK.

4.2.3   Two morphological solutions

Finally, there are two routes towards a morphological solution. On the one hand, we could maintain both the standard hierarchy, (26a), and a strict locality condition on allomorph-selection (Bobaljik 2012), but invoke an operation that brings T closer to *v*+V in the postsyntax, before Vocabulary Insertion. In particular, we could invoke the Lowering operation (Embick & Noyer 2001, 2007), which allows a head to adjoin to the head of its complement. Operating on the standard hierarchy, (26a), one operation of Lowering would adjoin T to Asp, a position from which T would be local to *v*+V; this would allow T to condition allomorphy of *v*+V and vice versa.

On the other hand, we could loosen the locality restriction on allomorph-selection to allow (somewhat) non-local conditioning. This again maintains the standard hierarchy, (26a), while allowing T to condition the form of *v*+V across an overt Asp, without invoking any special morphological operations. For a revised version of the relevant locality condition, one option is to allow linear adjacency to be sufficient for allomorph-conditioning, regardless of structural distance, as proposed by Ostrove (2015). Another would be to allow conditioning of allomorphy to proceed (basically) freely within the same complex head, as proposed by Harley & Choi (2016), building on Bobaljik (2012); see also Bobaljik (2000).

The major benefit of a morphological solution to our puzzle is that nothing special needs to be said about tense/aspect semantics in AK; T and Asp appear in their standard positions, (26a), and have a standard semantics. Note, however, that none of these morphological solutions naturally explains the exclusion of Asp from nominalizations and participles, cf. (24)–(25) and fn. 9.

We leave it as an open question here which—if any—of these solutions to the low tense puzzle is the right one.

5.   Conclusion

In this paper, we have shown that so-called verb stems in Indo-Iranian need not be treated as morphomic stems and so are not generally problematic for recent theoretical approaches to morphology and case/agreement splits. In particular, we showed that verb stems can be decomposed into morphemes (exponents) that relate systematically to semantic tense.

Solving this puzzle has uncovered several more. First, tense seems to occupy an unusually low position in AK, as discussed in §4. This puzzle might have a semantic, syntactic, or morphological solution. Second, though not discussed further in this paper for space reasons, is that split ergativity in AK is triggered by a change in tense. The directionality of this

split aligns with the purported universal directionality of tense-based split ergativity, with ergative alignment in the past tense (Dixon 1994). However, more recently, Coon (2013), following Salanova (2007), has surveyed a number of purported cases of tense-based split-ergativity, and concludes that there are no true tense-based splits; rather, these splits are in fact aspectually-conditioned. If our analysis of AK verbal morphology is correct, then Indo-Iranian languages (or at least AK) have true tense-based splits. A final observation is that the morpheme order in AK verbs, Asp-V-T, is extremely rare (Julien 2002).

Taking into account these other ways that AK is exceptional—a tense-based split and a rare morpheme order—may help us differentiate among the possible analyses of low tense in AK, §4.2. For example, a semantic solution would solve all three puzzles in one fell swoop: the reversed order of heads (Asp > T) allows T to affect the case/agreement alignment and results in a rare morpheme order within the complex verb. Alternatively, if case and agreement are post-syntactic phenomena, a morphological Lowering account might also solve all three puzzles together. We leave this open for future research, as well as the question of whether our analysis can extend to other Indo-Iranian languages.

References

Aikhenvald, Alexandra. 2011. Word-class changing derivations in typological perspective. In *Languages at large: essays on syntax and semantics*, 221–289. Leiden: Brill.

Alexiadou, Artemis. 2001. *Functional structure in nominals*. John Benjamins.

Altshuler, Daniel. 2015. Does viewpoint aspect make reference to time? Paper presented at Workshop on Aspect, Cornell University.

Aronoff, Mark. 1994. *Morphology by itself: stems and inflectional classes*. Cambridge, MA: MIT Press.

Aronoff, Mark. 2012. Morphological stems: What William of Ockham really said. *Word Structure* 5:28–51.

Atlamaz, Ümit. 2012. Ergative as accusative case: Evidence from Adıyaman Kurmanji. Master's Thesis, Boğaziçi University.

Bjorkman, Bronwyn, & Claire Halpert. 2012. In an imperfect world: deriving the typology of counterfactual marking. Talk at Modality Workshop, University of Ottawa.

Bobaljik, Jonathan. 2000. The ins and outs of contextual allomorphy. In *University of maryland working papers in linguistics*, ed. Kleanthes K. Grohmann & Caro Struijke, volume 10, 35–71. College Park: University of Maryland, Department of Linguistics.

Bobaljik, Jonathan. 2012. *Universals in comparative morphology*. Cambridge: MIT Press.

Caha, Pavel. 2009. The nanosyntax of case. Doctoral dissertation, University of Tromsø, Tromsø.

Calabrese, Andrea. 2012. Allomorphy in the Italian passato remoto: a Distributed Morphology analysis. *Language and information society* 18:1–75.

Coon, Jessica. 2010. Complementation in Chol (Mayan): A theory of split ergativity. Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Coon, Jessica. 2013. TAM split ergativity (parts I–II). *Language and Linguistics Compass* 7:171–200.

Demirdache, Hamida, & Myriam Uribe-Etxebarria. 2000. The primitives of temporal relations. In *Step by step*, ed. Roger Martin, David Michaels, & Juan Uriagereka, 157–186. Cambridge, MA: MIT Press.

Dixon, R. M. W. 1994. *Ergativity*. Cambridge University Press.

Embick, David. 2003. Locality, listedness, and morphological identity. *Studia Linguistica* 57:143–169.

Embick, David. 2010. *Localism versus globalism in morphology and phonology.*. Cambridge, MA: MIT Press.

Embick, David, & Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 555–595.

Embick, David, & Rolf Noyer. 2007. Distributed Morphology and the syntax/morphology interface. In *The Oxford handbook of linguistic interfaces*, ed. Gillian Ramchand & Charles Reiss, 289–324. Oxford University Press.

Giorgi, Alessandra, & Fabio Pianesi. 1997. *Tense and aspect: from semantics to morphosyntax*. Oxford: Oxford University Press.

Haig, Geoffrey. 2008. *Alignment change in Iranian languages: A construction grammar approach*. Mouton de Gruyter.

Halle, Morris, & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In *The view from building 20*, ed. Kenneth Hale & Samuel Jay Keyser, 111–176. Cambridge, Massachusetts: MIT Press.

Harley, Heidi, & Jaehoon Choi. 2016. Syntactic domains for morphological rules: Suppletion, dissociation, spanning and complex heads in Korean honorification. Talk presented in the University of Connecticut colloquiums series.

Iatridou, Sabine. 2000. The grammatical ingredients of counterfactuality. *Linguistic Inquiry* 31:231–270.

Iatridou, Sabine, Elena Anagnostopoulou, & Roumyana Izvorski. 2001. Observations about the form and meaning of the Perfect. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 189–238. MIT Press.

Julien, Marit. 2002. *Syntactic heads and word formation*. Oxford: Oxford University Press.

Klein, Wolfgang. 1994. *Time in language*. London: Routledge.

Merchant, Jason. 2015. How much context is enough? two cases of span-conditioned stem allomorphy. *Linguistic Inquiry* 46:273–303.

Ostrove, Jason. 2015. On morphological locality, linear adjacency and the nature of Vocabulary Insertion. Ms. University of Santa Cruz.

Radkevich, Nina. 2010. On location: the structure of case and adpositions. Doctoral dissertation, University of Connecticut, Storrs.

Salanova, Andrés Pablo. 2007. Nominalizations and aspect. Doctoral dissertation, MIT.

Stowell, Tim. 1995. The phrase structure of tense. In *Phrase structure and the lexicon*, ed. Johan Rooryck & Laurie Zaring, 277–291. Kluwer.

Stowell, Tim. 2007. The syntactic expression of tense. *Lingua* 117:437–463.

Svenonius, Peter. 2012. Spanning. Ms., University of Tromsø, CASTL.

Laura Kalin, Ümit Atlamaz

laura.kalin@gmail.com, umit.atlamaz@rutgers.edu

<p style="text-align:center">**Turkish comitatives: The genuine and the apparent**[*]</p>

<p style="text-align:center">Jaklin Kornfilt</p>

<p style="text-align:center">Syracuse University</p>

1.   Introduction: The Comitative Issue, the Turkish facts, and older literature on comitative constructions

Turkish has an interesting construction involving, at first glance, what is traditionally called a comitative[1] DP (or PP[2]), and plural agreement on the predicate, instead of the expected singular agreement. The construction in question is illustrated in (1):

(1)   Hasan-**la**   sinema-ya   git-ti-k.
       Hasan-with   cinema-DAT go-PAST-1.PL

Some literature, e.g. Aissen 1989, claims that (1) is ambiguous between the following two[3] readings: (a) *We* went to the movies with Hasan and (b) *I* went to the movies with Hasan.

---

[1] The comitative marker in Turkish surfaces in two possible forms: As the free morpheme *ile*, and its postclitic variant *-lA*. (Note that I shall follow here the general Turkological practice of using capital letters for segments that assimilate; where vowels are concerned, the capitalized vowel will undergo the rules of vowel harmony.)

[2] Some scholars treat the comitative in Turkish as a case like all others; if so, a DP marked with the comitative would still be a DP (or a KP, i.e. a Case Phrase). Others argue that the comitative marker shares certain properties with postpositions — in particular, that it is never stressed, "throwing" the (otherwise final) word-level stress onto the syllable immediately preceding it. Under this view, the comitative constituent will be a PP.

It is the second reading that is of interest to us here. How can (1) be interpreted as having a singular subject, given the plural agreement on the predicate?

In what follows in this introductory section, I will mainly present the discussion of both readings, and particularly of the "surprising", second reading, in previous literature — especially in older literature where some relevant cross-linguistic observations were made, and proposals were discussed, some of which were more semantically oriented than is necessary, in my opinion, for a syntactic phenomenon. In sections 2 and 3, I will then discuss the "surprising reading" and related phenomena, and advance my coordination-based proposal. Section 4 will summarize and conclude.

Let us start with an account for the less problematic first reading.

Suppose the "comitative" constituent is a dependent of the VP — a daughter of VP or adjoined to the VP. If the subject is *pro* with the features of first person plural, the agreement facts and the interpretation of (1) under its reading in a. are explained in a straightforward way. Given that Turkish is a "Pro-Drop" (Null Subject) language, this is a plausible analysis which gains in credibility if one considers that (1) has a counterpart with an overt pronoun for the reading in a.:

(2)   Biz Hasan-la   sinema-ya   git-ti-k.
        we Hasan-with cinema-DAT go-PAST-1.PL[4]
        'We went to the movies with Hasan.'

How should the surprising reading in b. for (1) be analyzed?

In this introductory section, I shall give a brief overview of prior literature addressing the "surprising reading" for comitative constructions, at least of work which has mentioned Turkish, even if briefly, and work related to those sources.[5]
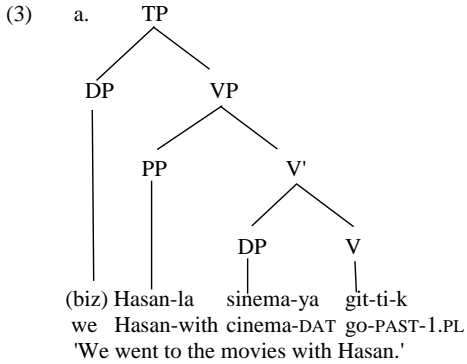
Aissen (1989) (henceforth A.) proposes the following two structures[6] to account for the two readings found in Turkish and other languages for the corresponding examples:

---

[3] As a matter of fact, I shall claim that the second reading of (1), given in b. below, is actually different, and that the correct form of it is: 'Hasan and I went to the movies.'
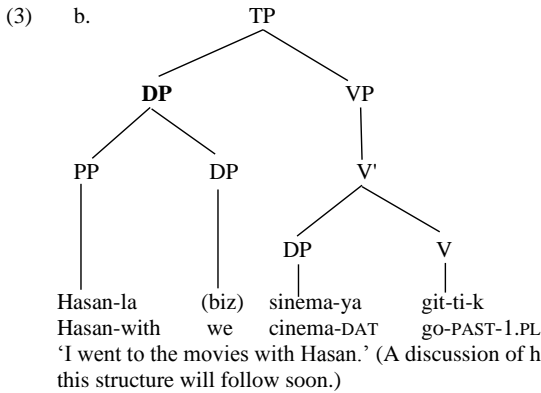
[4] The abbreviations used in this paper are as follows: 1.PL: first person plural; 1.SG: first person singular; 3.SG: third person singular; ACC: accusative; Act.Nom.: action nominalization; B: Boolean (head); BP: Boolean Phrase; CC: Coordinate Construction; DAT: Dative; Fact.Nom.: Factive Nominalization; GEN: Genitive; KP: Case Phrase; LOC: Locative; NumP: Number Phrase; Optat.: Optative; PASS: Passive; PL: Plural; PPC: Plural Pronoun Construction; Q: Yes/No question marker; SPC: Singular Pronoun Construction.

[5] Some more recent work, mainly on such constructions in Russian, has come to my attention only very recently and therefore can't be addressed in this paper; I hope to make up for this shortcoming in future work. Some of these studies on comitatives in Russian are Vassilieva & Larson (2005) and Ionin & Matushansky (2003). I am grateful to Sabine Iatridou for having given me access to these studies. Similarly, Demirok (2016) came to my attention only a few days before completion of this paper; I am thus unable to address that work on Turkish comitatives, other than noting that some observations and ideas, such as the person hierarchy mentioned here briefly, are also mentioned there anda re obviously proposed independently from my work.

[6] Aissen's structures are proposed for Tzotzil, but they are intended to carry over to other languages, such as Turkish, which display comitative structures with what I will continue to call the "surprising reading", i.e. with an understood singular subject under presence of plural agreement on the predicate, as in reading b. for (1). As already mentioned in the previous footnote, I shall claim that Turkish lacks the "surprising reading" in this word order, with the comitative phrase preceding the pronoun. Instead, I claim that the

(3)   a.

```
                    TP
                   /  \
                 DP    VP
                  |   /  \
                  | PP    V'
                  |  |   /  \
                  |  | DP    V
                  |  |  |    |
              (biz) Hasan-la  sinema-ya  git-ti-k
               we   Hasan-with cinema-DAT go-PAST-1.PL
              'We went to the movies with Hasan.'
```

                                                                   (adapted from A.'s (7a))

(3)   b.

```
                         TP
                        /  \
                      DP    VP
                     /  \   /  \
                   PP    DP V'
                    |    |  /  \
                    |    | DP    V
                    |    |  |    |
                Hasan-la (biz) sinema-ya  git-ti-k
                Hasan-with we  cinema-DAT go-PAST-1.PL
```

'I went to the movies with Hasan.' (A discussion of how to obtain this reading from
this structure will follow soon.)

                                                                   (adapted from A.'s (7b))

---

subject in the apparent "surprising reading" under this word order is actually a coordinate noun phrase. In the mirror image word order, which we will also consider, we do get a genuine "surprising reading", and I will attempt to explain that reading for the word order with the comitative expression following the pronoun, in section 3.

A. calls the bold-faced and enlarged DP (for her, an NP) in (3b) the "Plural Pronoun Construction (PPC)", and she schematizes the PPC as follows:

(4)   DP
         ┌────────┴────────┐
    (Adjunct)   (Head)
         │   │
        XP   Plural Pronoun   (adapted from A.'s 8)
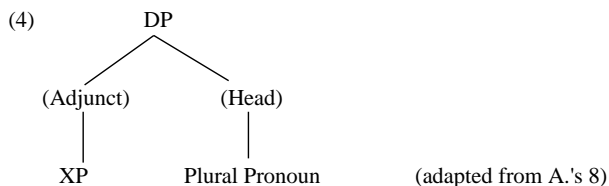
Crucially, the PPC has two daughters: A head, which is a plural personal pronoun, and an "Adjunct", which can take on various forms from language to language; in Turkish, it would show up as a comitative constituent.

   Given that Turkish is a Null Subject Language, the plural pronominal head of the PPC has the option of not being realized phonologically, thus yielding (1), under the reading shown in b. However, it should also be possible to have an overt pronoun in the head position of the PPC in (1), and A. offers the following example to confirm her prediction:

(5)   [Hasan-la   biz]$_{NP}$[7]   dükkân-a   git-ti-k.
         Hasan-with   we   store-DAT   go-PAST-1.PL
        'Hasan and **I** went to the store.'
                        (A.'s (9c), with her constituent analysis and her translation)

The schema for the PPC given in (4) would explain all facts (other than the singular interpretation of the plural pronoun, if such interpretation does exist — as already mentioned, I claim later that it does not in this word order, and under this constituent analysis), i.e. those of the plural agreement on the predicate as well as those of the order between the comitative and the pronoun: Turkish is head-final, hence any adjunct (which the comitative phrase is claimed to be under A.'s analysis) would precede a head, as it does in (5), under the assumption that the nominative pronoun is the head of the complex subject. Secondly, since the PPC would have a (plural) head, the (plural) features of that head would determine those on the predicate, presumably via feature percolation through the mother node — i.e. the boldfaced and enlarged DP (for A., an NP) in (3b).
   A. discusses the following semantic analysis of the PPC, based on Ladusaw (1988):

   The adjunct in the PPC imposes reference conditions on the head it modifies, without introducing any additional referents. In (5), the first person plural pronoun requires as a referent a set containing the speaker and at least one more person; the adjunct requires that this set contain *Hasan*. The speaker and *Hasan* together satisfy the plurality requirement of the head; hence, the PPC may refer to just two people: the (singular) speaker and *Hasan*. Thus, the adjunct refers to an individual *included* in the reference of the head rather than added to it. This

---
[7] Today's DPs were NPs when Aissen's article was written.

procedure of semantic interpretation would account for the "surprising" reading in (1b), which I shall call, following A., the "PP-included" reading — as opposed to the reading in (1a), which is the "PP-excluded" reading.

Given the schema in (4) for the PPC, reference to more than two individuals should also be possible, since the plural pronoun by itself may refer to more than two individuals. This, then, would be a second source of the "PP-excluded" reading, in addition to the source provided by the structure in (3a). This structural ambiguity is motivated in A. by examples such as (6):

(6)   Hasan-la   sinema-ya   git-ti-**m.**
       Hasan -with   cinema-DAT   go-PAST-**1.SG**
       'I went to the movies with Hasan.'   (adapted from A.'s (11c))

Under A.'s approach, (6) must have the structure in (3a), but with a *singular* subject and *singular* agreement, rather than the PPC structure in (3b), which would require a subject referring to at least two individuals and hence would be plural, determining plural agreement on the predicate.

   The type of analysis which this approach is designed to supersede has been called "Conjunct Movement" (cf. Hale (1975) for Navajo), "Conjunct Union" (Aissen 1987 for Tzotzil), and "Conjunct Splitting" (Grinder 1969, with a critique by Chung 1972, for Samoan); an early treatment for English is found in Lakoff and Peters 1969).

   All of these analyses have in common that they posit a coordinate structure in subject position at an early stage in the derivation, thus accounting for the plural agreement (via the simple arithmetic of adding instances of the number feature in each conjunct); they claim further that one of the conjuncts moves out of the coordinate structure, thus forming the comitative element. For our example in (1), the underlying source in such a framework for the PP-inclusive reading would be as follows:

(7)   [Hasan ve ben]_{DP}   sinema-ya   git -ti-**k.**
        Hasan and I   cinema-DAT   go-PAST-**1.PL**
       'Hasan and I went to the movies.'

The first person in the second conjunct would determine the first person agreement on the predicate, via a person hierarchy which would place the first person highest, followed by second person, with third person at the lowest end of the hierarchy.[8]

   Crucially, the coordinate nature of the subject would determine the plurality of the agreement under these (even) older approaches, i.e. at a minimum, each conjunct would represent a singular DP, and the two DP conjuncts would add up to a plural subject.

   Then, *Hasan* would move out of the higher DP in our example and, presumably, adjoin to the subject DP, or possibly to the VP; the conjunction marker would be deleted

---

[8] A hierarchy of this nature is probably needed in all accounts of this and related constructions; cf. A., p. 530, section 4.3. We shall see later that in coordinated DPs, a similar hierarchy dictates that the first conjunct has to be lower than the second conjunct, at least for many speakers. We shall return to this issue.

and the comitative marker would be inserted[9] in an appropriate position, thus creating the actual structure of (1) under its "inclusive" reading at Spellout (or, in the terminology of those older accounts, at S(urface)-Structure).

Note, however, that now the agreement on the predicate is not determined by any coordinate subject at Spellout, unless one allows a feature calculus based on heads and adjuncts, as well as one based on conjuncts in coordinate structures — an undesirable move. If the comitative expression is taken to have adjoined to VP rather than to the subject DP, the situation is even worse, because at Spellout, there wouldn't be any constituent that determines (plural) agreement. One would have to say that either the coordinate subject from a previous stage in the derivation determines agreement at Spellout, or else that a constituent consisting of the subject and the comitative *adjunct* at Spellout determines agreement. Neither proposal is a desirable move (although in recent models of grammar, morphology can be interspersed with syntax, thus perhaps weakening my first objection above), and both are rightly criticized by A. Perhaps even more seriously, the proposal based on "conjunct movement" has the additional problem of violating the Coordinate Structure Constraint.

Lastly, as mentioned earlier, one might adopt a semantic approach to the determination of plural agreement: It might be possible to claim that the "controller" of the agreement morphology on the predicate is determined not by the *syntactic* features of the subject, but by the features of the subject's *intended referent* (cf. Pollard and Sag 1988, as referred to in A.). Thus, suppose we did not follow A. in deriving the PP-inclusive reading for (1) from a PPC structure; rather, we would posit a first person *singular* pronoun in head position of the complex subject. According to Pollard and Sag's approach, although there would be no syntactic constituent that has the features of first person *plural*, the intended referent would be a set that includes the speaker and *Hasan*, hence *semantically* the crucial plurality is present and would determine the plural agreement.

A. succeeds in showing that, at least for Tzotzil and some of the other languages she considers, agreement — a *syntactic* phenomenon — is *syntactically* determined (namely by the PPC), and that it is not necessary to posit a semantically based explanation.[10] Furthermore, A.'s study also makes a strong case in favor of a "controller" of agreement that is a constituent at Spellout (for A., "S-Structure"), rather than a non-constituent string at Spellout, or else a string which was a constituent at a different level of representation.

My main purpose in this paper is to show that the analysis of the "comitative" construction with plural agreement and the PP-inclusive reading offered in A. is wrong for Turkish. However, A.'s central claims, made primarily for Tzotzil, do carry over nevertheless: The plural agreement is determined syntactically rather than semantically, and the determining entity is a constituent at Spellout.

---

[9] Obviously, such deletion and insertion operations of morphological markers are not available to us in more recent (morpho-syntactic) approaches; at any rate,they are unnecessary for an understanding of comitatives.

[10] However, A.'s approach is not *purely* syntactic, either, since her structure for the PPC has a complex subject whose head is syntactically plural, and where the singular interpretation of the plural head has to be based on semantic/pragmatic considerations, too (as based on Ladusaw 1988, as mentioned earlier).

2. Turkish Comitatives, Conjunctions, and the SPC

In this section, the properties of comitative subjects and of certain coordinate constructions will be discussed and compared.

2.1 Against a Turkish PPC

Let us return to (5) (A.'s (9c)), which was offered by A. as an overt instantiation of the PPC. For the sake of convenience, this example is repeated here as (8):

(8)   [Hasan-la   biz]$_{DP}$ dükkân-a   git-ti -k.
        Hasan-with we   store-DAT   go-PAST-1.PL
        'Hasan and **I** went to the store.' (A.'s bracketing and translation.)

As bracketed in (5/8), and with the brackets made clear by introducing a heavy intonational break after the putative subject (i.e. after the putative PPC), the only reading available for this example is not the PP-inclusive, but the PP-exclusive reading.[11] In other words, the translation given by A. and repeated here in (5) and (8) is wrong, and the only reading available for this example with the indicated bracketing (assuming, as I do, that the heavy intonational break does reflect the complex subject's right-hand boundary) is the one in (9):

(9)   Hasan and **we** went to the store.

However, the same string can have a PP-inclusive reading, if a heavy intonational break is introduced after the "comitative" constituent:

(10)   [Hasan-la]   [biz]   dükkân-a   git-ti -k.
         Hasan-with we   store-DAT   go-PAST-1.PL

(10) is ambiguous between a PP-inclusive reading (that is, '*I* went to the store with Hasan.') and a PP-exclusive reading ('*We* went to the store with Hasan.').
    The availability of the second reading, i.e. of the PP-exclusive interpretation for (10), is not surprising, if we realize that (10), under this reading, is a version of the PP-exclusive reading for (1), already discussed. The basic structure of (1) and of (10), under their PP-exclusive interpretation is given in (11):

(11)   [biz]   [Hasan-la]   dükkân-a   git-ti-k.
         we   Hasan-with   store-DAT   go-PAST-1.PL
         'We went to the store with Hasan.'

Here, the comitative constituent would be an adjunct of the VP, and *biz*, the first person plural pronoun, would be the sole, simple (and complete) subject. Given that Turkish is

---

[11] The grammaticality judgments reported here are mine and those of seven other native speakers, acknowledged in the special acknowledgment footnote.

word-order free to a rather high degree, I would claim that (10) is derived from (11) by scrambling the comitative VP-adjunct over the subject; the reading would, of course, be preserved.

We are faced with two questions, however. First, why does (5/8) not permit a PP-inclusive reading? Second, how can we account for the PP-inclusive reading of (10)?

My answer to the first question will be that this is simply because Turkish does not have the PPC — at least not in the way conceived in A.

Before attempting to answer the second question, however, I shall turn to yet another construction, which I will call — at least for the time being — the *Single* Pronoun Construction (SPC).[12]

2.2   Complex Subjects and Singular Pronouns

Note the properties of the construction in (12):

(12)   **[Hasan-la   ben]**_{DP} sinema-ya   git-ti-**k.**
    Hasan-**and I**   cinema-DAT   go-PAST-**1.PL**
    'Hasan and I went to the movies.'

Note that here, even though there is no unique, i.e. non-complex, first person *plural* subject that determines the first person plural agreement on the verb, we might explain the plural agreement by making either one of the following two assumptions:

A.   The complex constituent in subject position has a head; that head is the first person *singular* pronoun, whose person feature determines the verbal agreement in part. The factor determining the plurality on the verb is not syntactic, but semantic: The intended referent of the set denoted by the subject is a set consisting of *Hasan* (a comitative *PP* adjunct within the complex subject) and the speaker, and is therefore plural. Let me call this the *semantic* view of the *Single Pronoun Construction (SPC)*.

B.   The constituent in subject position is a kind of coordinate structure. At this point, we have to ask ourselves what the internal structure of such a coordination looks like. There are two main options: 1. The coordinate structure is completely parallel, with the two conjuncts merged at the same level, i.e. the traditional approach to coordination; 2. The coordinate structure is headed, in conformity with X'-syntax. I shall adopt the second option in both instances: The complex constituent in subject position is a coordinate structure, and that structure does conform to X'-syntax, at least up to a point, as we shall see very soon.

---

[12] It would be interesting to see if some of the other languages considered in A. to have the PPC might instead be more similar to Turkish and have a SPC which actually is a coordinate structure instead, as I shall claim later.

There are different options for the latter approach, with the coordinate structure in conformity with X'-syntax; I shall briefly consider two.

One approach to analyzing coordinate structures as headed is to view the conjunction morpheme as the head of the coordinate structure, with the second conjunct the complement, and the first conjunct the specifier of that head. In other words, what I had glossed earlier as 'with', i.e. as the comitative morpheme, is actually a conjunction marker in the SPC, as illustrated in (12). Note that there, I glossed the morpheme in question, *-lA*, as 'and', a conjunction marker, rather than as a comitative marker meaning 'with'. This is essentially the analysis proposed for coordinate structures in Munn (1993).

However, this is somewhat problematic, because in a head-final structure, the zero-level head, i.e. the conjunction, should be in an absolutely final position; if *–lA* is that head, it should be at the rightmost edge of the coordination, and it should cliticize to the second conjunct. However, as we have seen with the examples so far, it cliticizes to the first conjunct. Kayne (1994) addresses this problem (with respect to coordination in general, especially in head-final languages, rather than with respect to comitatives), by speculating that there could be two conjunction markers, with the head of the coordination showing up after the second conjunct (which would be its complement), and the first one after the first conjunct, the specifier of the coordination.

If this analysis is applied to the SPC, one would need to say that the conjunction which would be the head of the entire coordinate structure would be systematically silent in all of its occurrences, while the second instance of the conjunction marker would, again systematically, show up on the specifier. I find this problematic.[13]

Munn (1996) proposes another singly-headed analysis for coordinate structures. In this proposal, which was made for English, the first conjunct (=XP_2) acts as the phrasal head to the entire coordinate structure, while the second conjunct (=XP_3), itself headed by the coordinate conjunction, is adjoined to the first conjunct:

(13)

$$XP_{1(=2)}$$

$$XP_2 \qquad BP$$

$$B \qquad XP_3$$

(BP: Boolean Phrase; B: the Boolean conjunction)

---
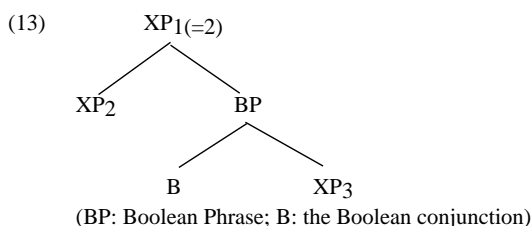
[13] Alternatively, one could say that the conjunction marker which is the head of the entire coordination and is thus at the rightmost edge of it moves to the first conjunct at PF and cliticizes to it, skipping over the second conjunct. This would be an undesireable move; what would the motivation be, and why wouldn't the coordination head not be able to cliticize to its (adjacent) complement, i.e. the second conjunct? Note further that *–lA* can never be realized overtly on both conjuncts. With respect to the above-mentioned speculations on the part of Kayne (1994) whether in head-final languages such doubling of conjunction markers is possible, with the marker showing up after both conjuncts: this possibility would support setting up a syntactic head position at the right edge. Turkish does have such a marker, *-DA*; however, it means, when doubled, 'as well as; too', rather than 'and', the meaning that *–lA* has in the SPC. As mentioned in the text, *-lA* shows up only once in this construction, and always on the first conjunct.

I propose to adopt this structure for Turkish, but in a mirror-image fashion with respect to the direction of the adjunction. Note that in head-final languages it is unusual to have adjunction to the right; typically, adjunction is to the left of the target (cf. also Chomsky 1995: 340); Munn's phrasal adjunction illustrated in (13) would result in (14) for head-final structures:

(14)

$$\begin{array}{c} XP_{1(=2)} \\ \diagup \quad \diagdown \\ BP \qquad XP_2 \\ \diagup \quad \diagdown \\ XP_3 \qquad B \end{array}$$
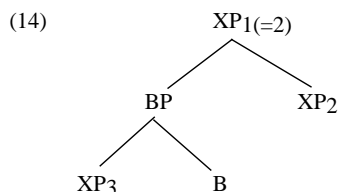
I propose here that, at least for head-final languages, adjunction in coordinate structures is leftward in general. Further, the B-element is a marker of the adjunct (which, in the terms of Kayne 1994, can also be viewed as a specifier, given that in that approach, there is no distinction between the two). This is why it shows up with the adjunct phrase, i.e. the "first conjunct", rather than at the rightmost edge of the entire coordination.

    We now have a structure which conforms to X'-syntax somewhat, in that it has a head which projects its features to the mother node. However, that head is phrasal, rather than being a zero-bar level category. The structure is quite similar to the structure proposed by A. for the PPC (or the SPC) cross-linguistically; crucially, the XP_2, i.e. the phrasal head, is singular here, so that its singular semantics can be read off the structure directly, without needing recourse to semantic or pragmatic considerations. Furthermore, the first "conjunct" is crucially not a PP, but rather a Boolean Phrase, thus determining the coordination-related properties of the entire coordination. We shall see later that *ile/-lA*, too, has different properties as B, from when it is a P in Turkish.

    In (12), the relevant kind of example for the comitative coordination construction with a singular phrasal head, it is the person feature of the first person singular pronoun that determines the person feature of the verbal agreement. Also, I mentioned earlier that it is the nature of the subject as a coordinate structure which determines the subject's plurality and consequently the plural feature of the verbal agreement. Let me call this the *Coordinate Construction (CC)* view of the Singular Pronoun Construction (SPC). As pointed out earlier, I have indicated my commitment to this view by glossing the suffix -*la* in (12) as 'and', rather than as 'with'.

    Before turning to evidence in favor of this view, let me mention briefly that A. does mention the possibility that some languages might have a SPC, but she does not mention any specific languages of this type. Turkish, then, might be claimed to be one such language, rather than a language with the PPC. The SPC, then, would provide UG with a second source for the PP-included readings of examples like (1) cross-linguistically. This would be in addition to the PPC; crucially, syntactic rather than semantic properties of the subject would determine the plural agreement on the predicate, as already pointed out earlier.

2.3   Against Conjunct Movement

First, I would like to advance some data that differentiate the structure underlying the PP-excluded reading from the structure underlying the PP-included reading; while these data will be neutral between a PPC, or a coordination-based SPC structure for the "PP-included" reading, they will serve to argue against the older view based on "Conjunct Movement", and they will at the same time support a systematically different analysis for the "adjunct" comitative when it is a PP, from when it is a BP.

First of all, note that the difference in function between *-lA* as a genuine comitative marker, i.e. as a P, and as a conjunction marker, i.e. as a B, goes along with a difference in (basic) word order. In the structure with the PP-excluded reading, illustrated in (11) for a simple plural pronominal subject, as well as in (15) below with a singular pronoun as the simple subject, the comitative constituent follows the subject in the basic word order:

(15)   Ben   Hasan-la   sinema-ya   git-ti -**m.**
         I   Hasan -with   cinema -DAT   go-PAST-**1.SG.**
         'I went to the movies with Hasan.'

Note the singular agreement on the verb, showing that *ben* 'I' is the sole subject. The word order is as expected, since *Hasan-la* 'with Hasan' is an adjunct of VP.

As we saw in (10)/(11), *Hasan-la*, the comitative PP, as well as *ben*, the subject, can be scrambled around, when the comitative PP is a VP-adjunct; importantly, these two expressions do not form a constituent and thus can easily be separated. A few examples follow, showing how freely such scrambling can apply:

(16)   **Hasan-la   ben**   sinema-ya   git-ti-**m.**
         Hasan-with   I   cinema-DAT   go-PAST-**1.SG**
         Same as in (15), with sentential focus on the dative DP.

(17)   **Ben**   sinema-ya   **Hasan-la**   git-ti-**m.**
         I   cinema-DAT   **Hasan-with**   go-PAST-**1.SG.**
         Same as in (15), with sentential focus on the comitative PP.

(18)   **Ben**   sinema-ya   git-ti-**m**   **Hasan-la.**
         I   cinema-DAT   go-PAST-**1.SG.**   Hasan-**with**
         Same as in (15), with sentential focus on the dative DP, and with the comitative PP backgrounded, as a post-verbal constituent .

(19)   **Hasan-la**   sinema-ya   **ben**   git-ti -**m.**
         Hasan-with   cinema-DAT   **I**   go-PAST-**1.SG.**
         Same as in (15), with sentential focus on the subject, and the comitative PP topicalized.

(20)   **Hasan-la**   sinema-ya   git-ti-**m**   **ben.**
       Hasan-with   cinema-DAT   go-PAST-**1.SG.**   **I**
       Same as in (15), with sentential focus on the dative DP, with the subject
       backgrounded as a post-verbal constituent, and the comitative PP topicalized. [14]

On the other hand, where -*lA* functions as a conjunction marker rather than as a
comitative marker (a P), the order is rigid: The constituent marked with -*lA* has to
precede the second element of the larger constituent, i.e. the second conjunct under the
view defended here; we see this order in (12), repeated here as (21), for the reader's
convenience:

(21)   [Hasan-**la**   **ben**]$_{DP}$ sinema-ya   git-ti-**k.**
        Hasan-**and**   **I**   cinema-DAT   go-PAST-**1.PL**
        'Hasan and I went to the movies.'

Furthermore, the constituent marked with -*lA* cannot move out of a complex DP, nor can
the second element in that DP move out; as we just saw in (16) through (20), both of
those elements are free to scramble within the clause if they do not form a constituent
with each other. In contrast, the following examples constitute unsuccessful attempts to
scramble either element out of the complex subject DP in (12)/(21), which I have claimed
is a coordinate structure:

(22)   *[[e]$_i$ **ben**]$_{DP}$ sinema-ya   git-ti-**k**   [Hasan-**la**]$_i$
               I   cinema-DAT   go-PAST-**1.PL**   Hasan-**and**
        Same intended reading as (21).

(23)   *[Hasan-**la**   [e$_i$]]$_{DP}$ sinema-ya   git-ti-<u>k</u>   [**ben**]$_i$
         Hasan-**and**   cinema-DAT   go-PAST-**1.PL**   **I**
        Same intended reading as (21).

(24)   *[**Ben**]$_i$ [Hasan-**la**   [e$_i$]]$_{DP}$ sinema-ya   git-ti   -**k**
         I   Hasan-**and**   cinema-DAT   go-PAST-**1.PL**
        Same intended reading as (21).

Note also that the entire complex subject, which I claim is a coordination, can scramble
as a single unit:

(25)   sinema-ya   git-ti-**k**   [Hasan-**la**   **ben**]$_{DP}$
       cinema-DAT   go-PAST-**1.PL**   Hasan-**and**   **I**
       Same reading as (21), with the entire coordinated subject backgrounded.

---

[14] An alternative reading exists, where the comitative PP is a contrastive focus, and where the dative
constituent is not focused. Similar contrastive focus readings are available for the subject in (16) and (18),
likewise only possible when the verb-adjacent constituent is not focalized. The issue of contrastive focus in
positions other than the one preceding the verb is a complicated one which this paper will not address.

These facts are easily explained by appealing to the Coordinate Structure Constraint, however formalized: The structure in (21), with a singular pronominal second conjunct following the apparently comitative adjunct posits a coordinate structure for the subject, thus not allowing movement of either part of the coordinate structure to move out of the coordination.[15] The plural agreement on the verb is explained by the nature of the subject as a coordination; here, each DP conjunct is singular, and, together, they form a plural coordination. The singular reading of the second conjunct is obviously read off directly from the syntax, given that this conjunct is an overt singular pronoun. This point will be important later, when looking at similar structures with a phonologically unrealized second conjunct in such constructions with plural agreement.

2.4   The Coordinate Construction (CC) nature of the Single Pronoun Construction (SPC) Structure: Additional evidence

I now turn to some additional evidence that favors the coordination nature of the SPC structure, i.e. in favor of the view that the structure underlying the PP-included readings is a coordinate structure rather than the view that the complex subject is headed by a singular person pronoun with a PP-adjunct.

First of all, Turkish does not, in general, allow direct modification of a head by a PP; in most instances, a verb is needed that, together with the PP, forms a relative clause modifier of the head:

(26)   a. * [[Chomsky tarafından]$_{PP}$ bir kitap]$_{DP}$ al -dı-m.
        Chomsky by a book buy- PAST-1.SG
      Intended reading: 'I bought a book by Chomsky'

    b. [$_{DP}$[$_{CP}$[$_{PP}$ Chomsky tarafından] yaz-ıl-mış] bir kitap] al-dı-m.
           Chomsky by write-PASS-Participle a book buy-PAST-1.SG
      'I bought a book written by Chomsky.'

(27)   a. *Bu, Ali için bir kitap-tır.
       this Ali for a book-Assertive
      Intended reading: 'This is a book for Ali.'

    b. Bu, Ali için yaz-ıl-mış bir kitap-tır.
       this Ali for write-PASS -Participle a book-Assertive
      'This is a book written for Ali.'

---

[15] In addition to the movement-based examples which illustrate the workings of the Coordinate Structure Constraint and thus support the analysis of examples such as (12)/(21) as having complex subjects with a coordinate structure, there are intonation-based examples which make the same point: In (12)/(21), no intonational break can be placed between the two conjuncts:

(21)'   *[Hasan-**la //** **ben]**DP sinema-ya git-ti-**k.**
      Hasan-**and** **I** cinema-DAT go-PAST-**1.PL**

The reason is clear: the break forces a constituent boundary between the two conjuncts; now, we have no coordinate structure any longer, and ben 'I' is the simple and sole subject of the utterance; its singular number feature conflicts with the plural agreement on the verb.

Of course it would be possible to claim that the postposition *ile* or its cliticized form *-lA* are somewhat exceptional and do allow the PP that is headed by either form to act as a nominal modifier. Note, however, that this would weaken the (non-CC) SPC hypothesis considerably.

Secondly, in the rare instances where a PP can immediately modify a nominal, it can move out of the complex DP:

(28) a. [[Sen-in   kadar]$_{PP}$ bir   çocuk]$_{DP}$ gör-dü-m.
          you-GEN   "as much as" a   child   see-PAST-1.SG
       'I saw a child as old/tall/... as you.'

      b. [[e]$_i$ bir çocuk]$_{DP}$ gör-dü-m [$_{PP}$ sen-in kadar]$_i$

(29) a. [[San-a   göre]$_{PP}$ bir okul]$_{DP}$ bul-du-m.
          you-DAT suitable a   school   find-PAST-1.SG
       'I found a school appropriate for you.'

      b. [[e]$_i$ bir okul] bul-du-m [san-a göre]$_i$

Note that the result of scrambling the PP-modifier out of the complex NP is fine, illustrated by (28b) and (29b). However, crucially, DPs marked with *-lA* cannot scramble out of their complex DP when they are in a CC-structure, as we saw in (22) - (24). In those examples, the complex DP is a subject; but the same behavior is also found where a corresponding DP is an object:

(30) a. Bugün saatlerce [fizik-**le**   kimya]]   çalış -tı-m.
          today   for hours   physics-**and** chemistry   study-PAST-1.SG
       'I studied physics and chemistry for hours today.'

      b.* Bugün saatlerce [[e]$_i$ kimya]] çalış-tı-m [fizik-**le**]$_i$

All of these facts are easily explained if the complex DP is a conjunction: Any movement of a conjunct out of the coordinate structure will be ruled out by the Coordinate Structure Constraint. I therefore take these facts to establish the coordinate nature of the complex constituent, which is an object here. As a consequence, the complex subject that determines plural agreement for the inclusive readings and has a single pronominal head is a coordinate structure.

Under this account, the PP-included reading of (1) is due to the same CC-structure underlying (12), with the difference that the second conjunct is a *pro* (a phonologically empty pure pronominal in the terminology of Chomsky 1982 and much subsequent work), with the features of first person singular for the PP-included reading and with the features of first person plural for the PP-excluded *conjunctive* reading.

2.5   Brief summary of the readings of (1)

For the reader's convenience, I repeat example (1), which the paper started off with, and then summarize the readings that it has:

(1)   Hasan-**la**   sinema-ya   git-ti-k.
         Hasan-**with**   cinema-DAT   go-PAST-1.PL

(1) is actually three-ways ambiguous, rather than just two-ways as stated in the introduction:

A.   [Hasan and *I*] went to the movies ("PP-included" (actually, "BP-included"); the subject is a *complex* DP: a coordinate structure with a *singular pro* as one conjunct, and *Hasan* as the other conjunct — under the analysis adopted here for coordinations, *Hasan* is a BP-adjunct of the singular *pro*);
B.   [Hasan and *we*] went to the movies; the subject is a *complex* DP: a coordinate structure with a *plural pro* as one conjunct and *Hasan* as the other conjunct = a BP-adjunct within the coordinate structure;
C.   *We* went to the movies *with Hasan* (PP-excluded); the subject is a *simple* DP: a *plural pro*; the comitative PP-constituent is external to the subject.

Both a. and b. have a structure with a coordinate subject and differ only with respect to the plurality of the *pro* conjunct within the subject. Reading c. is related to the structure in (3a.), with a simple subject and a genuine comitative PP which is a VP-adjunct.

I now turn to some apparent counterexamples against the CC Hypothesis for the BP-included reading (as I shall be calling the "PP-included reading" from now on) for examples such as (1).

2.6   Challenging the CC-Hypothesis

One type of potential argument against the CC-Hypothesis is based on the limitations of the distribution of the "comitative" marker when used as a conjunction marker.

The marker *ile*/*-lA* us used to conjoin only two elements, while there is no such restriction on the number of conjuncts when the conjunction *ve* 'and' is used:

(31)   a.* Dün   sinema-da Zeynep-**le** Mehmet-**le** Hasan-ı   gör-dü-m.
               yesterday cinema-LOC Zeynep-**and** Mehmet-**and** Hasan-ACC   see-PAST-1.SG
               Intended reading: 'Yesterday I saw Zeynep and Mehmet and Hasan at the movies.'

         b.? Dün   sinema-da Zeyneb-i   **ve** Mehmed-i   **ve** Hasan-ı
               yesterday cinema-LOC Zeynep-ACC **and** Mehmet-ACC **and** Hasan-ACC
               gör-dü-m.
               see-PAST.1
               'Yesterday I saw Zeynep and Mehmet and Hasan at the movies.'

The awkwardness of (31b) is stylistic and comes from the repetition of the conjunction marker. Crucially, (31b) is grammatical, while (31a) is not.[16]

Secondly, where the complex DP is in a case-marked position, only the second conjunct receives the appropriate case; the first conjunct, marked with *ile/-lA*, does not:

(32) a. Dün   sinema-da   Zeynep-**le**   Hasan-**ı**   gör-dü -m.
          yesterday cinema-LOC   Zeynep-**and**   Hasan-**ACC**   see-PAST-1.SG
      'I saw yesterday Zeynep and Hasan at the movies.'

      b.* Dün   sinema-da   Zeyneb-**i-yle**   Hasan-**ı**   gör-dü -m.
          yesterday cinema-LOC   Zeynep-**ACC-and**   Hasan-**ACC**   see-PAST-1.SG

In coordinate structures with the general conjunction marker *ve*, on the other hand, each conjunct can receive the appropriate case, as we saw in (31b).

These limitations on *ile/-lA* would follow, if this element is not a conjunction marker, but a postposition; that postposition would be expected to show up on one element (i.e. its complement) only, and it would block case assignment to the postpositional object by another case marker (in this instance, by the verb).

If we accept this analysis of *ile/-lA*, we would also have to adopt an analysis which views the comitative adjunct as a PP rather than as a BP in examples like (12) (and also in examples like (1), with the second conjunct a small *pro*, but otherwise with an identical structure), i.e. accept an analysis of the second conjunct as a singular pronoun to which a PP rather than a BP has been adjoined. This, in turn, would force us to posit semantic determination of the plural agreement. (Or, conversely, an analysis of the BP as a PP in examples like (1) would force us to posit semantic rather than syntactic determination of the singular reading of the silent second conjunct despite the presence of plural subject agreement on the verb.)

I would like to suggest here that, at least for Turkish, the cost is too high for this move, and that the CC analysis of the inclusive reading and a singular interpretation of the second conjunct can be maintained in spite of the problems just mentioned.

2.7   In Defense of the CC-Hypothesis

We saw earlier, namely in the examples (16)-(20) and in (21)-(24) that the putative PP-adjunct marked with *-lA* in the complex DP behaves differently from other PP-adjuncts, and we further saw, via examples (26) and (27), that PP-adjuncts within DPs are extremely limited in Turkish. These facts argue strongly against a SPC-approach which would not be based on a CC.

Furthermore, the arguments against the CC-hypothesis presented above lose some of their force when additional evidence is considered.

First of all, the case assigned to a complex DP does not have to spread to the conjuncts in Turkish (as it apparently must in other languages like Russian), when that complex NP is a coordinate structure; witness the behavior of a coordinate structure with *ve* 'and', whereby either both conjuncts are marked with the accusative suffix (as also

---

[16] Actually, (31a) is OK, when the three potential conjuncts are arranged in a pair-wise fashion, i.e. under the reading: 'Yesterday, I saw Zeynep and [Mehmet and Hasan] at the movies.'

seen earlier in (31b)), but where the option exists of only the second conjunct bearing that suffix:

(33) a. Dün   sinema-da   [Ahmed-**i**   ve   Hasan-**ı**]   gör-dü-m.
          yesterday cinema-LOC   Ahmet-**ACC** and Hasan-**ACC**   see-PAST-1.SG
          'Yesterday I saw Ahmet and Hasan at the movies.'

      b. Dün   sinema-da   [Ahmet   **ve**   Hasan]-ı   gör-dü-m.
          yesterday cinema-LOC   Ahmet   **and** Hasan-**ACC** see-PAST-1.SG
          'Yesterday I saw Ahmet and Hasan at the movies.'

Note that in (33b), the first conjunct does not bear the Accusative marker; only the second conjunct does. From this point of view, the behavior of the coordinate object in (33b) is closer to that of the complex DP with *-lA* in (32a) than it is to the coordinate object with *ve* in (31b).

I would like to suggest the following approach to the conjunction facts with respect to case marking. In examples like (31b), where each conjunct bears the case marker appropriate for the mother node of the coordination, each conjunct is a KP (i.e. a case phrase). In examples like (33b), where only the last conjunct bears the case marker, each conjunct is a DP. The reason why the case marker attaches only to the last conjunct in such examples is not the status of that conjunct as the head of the phrase; rather, the case is really at the level of the mother node from a syntactic point of view; but since the case marker happens to be a suffix morphologically, it attaches to the right-hand periphery of the phrase, which happens to be the end of the last conjunct.[17]

If this is the correct analysis for conjunctions formed with *ve* in examples where the conjuncts do not bear case themselves, there is no reason why the same analysis should not hold for conjunctions formed with *-lA*, as in (31a). There, too, the last (i.e. second) conjunct will be the only one to carry the case marker; this will be due that conjunct's phrase-final position and not because of its putative head status.

The fact still remains that *ile/-lA* can conjoin only two elements, while *ve* is not restricted in this way. But we can simply say that this is a lexical property of this item; why shouldn't languages have the option of marking in some special way coordinations consisting of two items only? The alternative is to say that the constituent marked with this element is a PP, and we saw earlier that it does not behave like a regular PP adjunct within an DP. Hence, the account positing a coordinate structure for the complex DPs in question is the preferable one.

There is one more restriction on the distribution of *ile/-lA*: It can conjoin only nominal elements, while *ve* is not so restricted. This, at first glance, would make *ile/ -lA* seem more like a case marker or a postposition rather than a conjunction marker.

However, some more detailed observation shows that *ile/-lA* can conjoin not only DPs, but also nominal entities of a lower phrasal level. An example of N^0-level (or perhaps NumP-level) conjunction with *ile/-lA* follows.

---

[17] This phenomenon of a single case marker (or of other bound morphemes) distributing over conjuncts in a coordination is often referred to in the literature as Suspended Affixation. For this phenomenon in Turkish, see, for example, Erdal 2007, Kornfilt 2012, and Lewis 1975.

(34) Dün   deniz-de [DP [NP kocaman [N ıstakoz-lar-**la**   balık-lar]]] gör-dü-m.
      yesterday sea-LOC   huge   lobster-PL-**and** fish-PL   see-PAST-1.SG
      'Yesterday I saw huge lobsters and (huge) fish in the sea.'

Note that under the reading sketched in the translation, the adjective *kocaman* 'huge' can
have scope over both conjuncts, i.e. it can be external to the coordination with *–la* under
this reading.

    Since case is assigned to DPs and not to lower level phrases, and since postpositions
select for DP objects (unless they select for clausal objects), it appears that *ile/-lA* as a
conjunction marker is neither a case nor a postposition.

    Note also that *ile/-lA* exhibits some interesting behavior where it appears with some
types of nominal clauses. It can freely attach to infinitivals and clauses whose predicate is
a so-called action nominal, but not to clauses with so-called factive nominals[18]:

(35) [[Ahmed-in   yarış-ı   kazan-**ma**-sı] -**yla**   [Mehmed-in
       Ahmet-GEN race-ACC   win-**Act.Nom.**-3SG-**and**   Mehmet-GEN

      iş-in-den   kov-ul-**ma**-sı]]   aynı   gün-e   rasla-dı.
      work-his-ABL fire-PASS-**Act.Nom.**-3.SG   same   day-DAT   meet-PAST
      'Ahmet's winning the race and Mehmet's getting fired from his job happened on
      the same day.'

(36) * [[Ahmed-in   yarış-ı   kazan-**dığ**-ı]-**yla**   [Mehmed-in
       Ahmet-GEN race-ACC   win-**Fact.Nom.**-3.SG-**and**   Mehmet-GEN

      iş-in-den   kov-ul-**duğ**-un]]-u   yeni   duy-du-m.
      work-his-ABL fire-PASS-**Fact.Nom.**-3. SG-ACC   new   hear-PAST–1.SG
      Intended reading: 'I've heard recently that Ahmet won the race and Mehmet was
      fired from his job.'

It is likely that an explanation of this difference lies in the somewhat more nominal
nature of the "action" nominal complements as opposed to "factive" complements; the
main independent evidence is that action nominal complements have no marking of
independent tense at all, while in factive nominals, the marker *-DIG* denotes non-future;
future is expressed by the special factive nominal marker
-AcAG.

    Note, however, that factive nominal complements are nevertheless nominal enough to
accept, indeed require, case marking — otherwise a property of DPs; furthermore, they
show up in syntactic positions otherwise occupied by DPs, e.g. as objects of verbs as well
as of postpositions. Hence, if factive complements are DPs on independent grounds, the
inability of *ile*/-*lA* to conjoin such complements must be due to some lexical

---

[18] This terminology for different types of nominalization goes back to Lees 1963. As has been shown more
recently, these terms do not always fit the context, given that the so-called factive nominalization can show
up in non-factive contexts, and the so-called action nominalization can show up for non-actions. For some
discussion see Erguvanlı Taylan 1998, among others.

idiosyncracy[19] of the conjunction marker rather than (solely) to the defective nominality of the complement clause.

It should also be noted that the conjunction marker *ve*, while more general in its distribution, cannot be used everywhere, either. For example, *ve* sounds very stilted when conjoining sentences and VPs:

(37) ?? [Ahmet yarış-ı   kazan-dı] **ve** [Hasan iş-in-den   kov-ul-du].
         Ahmet race-ACC win-PAST **and** Hasan job-his-ABL fire-PASS-PAST
         'Ahmet won the race and Hasan was fired from his job.'

(38) ? Ahmet [yarış-ı   kazan-dı] **ve** [vatan-ın-a   dön-dü].
         Ahmet   race-ACC win-PAST   **and** country-his-DAT   return-PAST
         'Ahmet won the race and returned to his home country'

It sounds much better to use a construction for sentential coordination whereby the conjunction marker follows the first item of the last conjunct:

(39)   [Ahmet yar-ış-ı   kazan-dı], [Hasan **da**   iş-in -den   kov-ul -du].
        Ahmet race-ACC win-PAST   Hasan **and** job-his-ABL   fire-PASS-PAST
        'Ahmet won the race and Hasan was fired from his job.'

The best-sounding VP-coordination pattern, on the other hand, is one where the tense/aspect and agreement markers of the first conjunct are replaced by the suffix *–(y)Ip* (if tense/aspect and agreement are to be interpreted as being the same as those of the second conjunct):

(40)   Ahmet [yarış-ı   kazan-**ıp**]   [vatan-ın-a   dön-dü].
        Ahmet race-ACC win-**and**   country-his-DAT return-PAST
        'Ahmet won the race and returned to his home country.'

It appears, therefore, that Turkish has a variety of different coordination markers, each one with its own idiosyncratic properties. Thus, rather than claiming that *ile/ -lA* is a postposition in these coordination structures, it would be more appropriate to claim that it is a special conjunction marker. This will enable us to maintain a purely syntactic explanation of the plural agreement facts for the so-called PP-included readings, which are, as mentioned earlier, actually BP-included readings.

I now turn to some differences between *ile/-lA* as a clear-cut postposition in the PP-excluded construction and *ile/-lA* as a conjunction marker.

---

[19] This idiosyncracy probably has to do with the difference in the height of the clausal architecture at which different coordination markers attach; the general coordination marker *ve* clearly attaches very high, at the KP level which is above the nominalized clausal level, while *ile/-(y)la* attaches lower (since case markers can't show up when this marker is present), but probably as high as TP, given that it can distinguish between a T-head which has genuine tense features from a completely defective T without any tense features; for a treatment of nominalized clauses with different levels at which "nominalization" applies, and for the idea of defective tense in such clauses, see Borsley & Kornfilt (2000), Kornfilt & Whitman (2011), among others.

2.8   Differences between the Postposition and the Conjunction

The first difference is a semantic one. In its purely comitative use, *ile/-lA* denotes togetherness. On the other hand, togetherness is a possible, but not a necessary ingredient of a situation denoted by a coordinate structure. Hence, different truth conditions obtain for the following two examples:

(41)   Ben   dün   Hasan-**la**   sinema-ya   git-ti-**m**.
       I   yesterday   Hasan-**with**   cinema-DAT   go-PAST-**1.SG**
       'I went to the movies with Hasan yesterday.'

Hasan-la is a regular comitative constituent here, and the construction has the PP-excluded reading; it makes good sense to view *ile/-lA* as a postposition in this usage. Note that (41) is true only if Hasan and I went to the movies together; it would be false, for example, if we went to different movie theaters, even if we did so at the same time.
    Now compare (41) with (42), a BP-included construction and, as I claim, with *ile/-lA* as a conjunction marker rather than as a comitative postposition:

(42)   Hasan-**la**   ben   dün   sinema-ya   git-ti-**k**.
       Hasan-**and**   I   yesterday   cinema-DAT   go-PAST-**1.PL**
       'I went to the movies with Hasan yesterday.'

Although the primary reading of (42) is that Hasan and I went to the movies together, it is possible to imagine situations where we went separately or even to different cinemas. Suppose, for example, that Hasan and I are part of a group of diligent students preparing for an exam. Yesterday, Hasan and I were the only ones who wanted to escape work; (42) could then be uttered successfully. If we went separately, (42) would be true in such circumstances, but (41) would be false. If so, the "conjunctive" use of *ile/-lA* is not comitative, and the morpheme should not be glossed as 'with', but with 'and', in this particular usage.
    Another difference between *ile/-lA* as a genuine comitative postposition and as a conjunction marker comes from dialect data.
    ile/-lA is one of a small number of postpositions that assign Genitive case to their objects, if the object is a pronoun (more specifically, a personal pronoun, a demonstrative, or the [+human] WH-element *kim* 'who'; otherwise, no overt case is assigned to the complement, i.e. the object of this element:

(43)   Ben   sen-**in-le**   sinema-ya   gid-eceğ-**im**.
       I   you-**GEN-with**   cinema-DAT   go-FUT-**1.SG**
       'I will go to the movies with you.'

(44)   Ben   Hasan-**la**   sinema-ya   gid-eceğ-**im**.
       I   Hasan -**with**   cinema-DAT   go-FUT-**1.SG**
       'I will go to the movies with Hasan.'

(45) *Ben   Hasan-**ın-la**   sinema-ya   gid-eceğ-**im.**
      I   Hasan-**GEN-with**   cinema- DAT   go-FUT -**1.SG**

There is a tendency in modern Turkish to omit the Genitive assigned by postpositions to pronouns in the spoken, colloquial style.[20]

     More conservative speakers do tend to retain the Genitive — a tendency which is enforced by normative influences via grammar textbooks, grammar instruction in schools etc. However, in spite of such pressure, even the conservative dialect allows (as a matter of fact, even prefers) for "Genitive-Drop" in the structures which I am claiming are due to coordination and to *ile/–lA* as a conjunction marker:

(46)   a.*/?? Gel, [sen-**in-le**   ben] bu akşam sinema-ya   gid-elim.
             come you-**GEN-and** I   this evening cinema-DAT go-OPTAT+1.**PL**
             'Come on, let's you and I go to the movies tonight!'

        b.   Gel, [sen-<u>l</u>e ben] bu akşam sinema-ya gid-elim.

This strong tendency to drop the Genitive on a pronominal which is the first conjunct in a coordination formed with the cliticized form *–lA* becomes a prohibition of grammar when it is the first conjunct rather than the second which determines, via the Person Hierarchy mentioned earlier, the person feature of the agreement marking on the predicate:

(47)   a. * [Sen-**in-le**   o]   dün   sinema-ya   mı git-ti-**niz**?[21]
            you-**GEN-and** he/she/it yesterday cinema-DAT   Q   go-PAST-2.**PL**
            'Did you and she go to the movies yesterday?'

        b.(?)[Sen-**le** o] dün sinema-ya mı git-ti-<u>niz</u>?[22]

---

[20] The Genitive found elsewhere, e.g. with subjects of embedded nominalized clauses or on possessors in possessive noun phrases is not subject to this trend to omit it. This trend is observable not only for the comitative, but also for other postpositions that share the property of assigning the Genitive to pronouns only—although this tendency is much more pronounced with the clitic version of *ile/–lA*.

[21] It is important to note the plural rather than singular agreement here, in order to understand the relevance of the ungrammaticality of this example. The almost identical counterpart of (47a), with the same word order and with the Genitive marking on the pronoun, is perfectly grammatical when the agreement is singular (note that the agreement marker for third person singular is zero in the verbal agreement paradigm):

(i)   Sen-**in-le**   o   dün   sinema-ya   mı git-ti?
        you-**GEN-with** he/she/it yesterday cinema-DAT   Q   go-PAST
        'Did she go to the movies with you yesterday?'

Here, the third person singular pronoun *o* is the sole subject of the sentence and thus determines the agreement on the predicate. There is no complex subject here; rather, the DP marked with **-le** is simply a comitative PP which has been scrambled to sentence initial position from its original position within the VP; the source for this PP-excluded construction is (ii):

(ii)   O dün sen-**in-le** sinema-ya mı git-ti?

In order to understand these facts better, it might be helpful to realize that overt conjunction markers are a rather late development in Turkish; Turkologists report that the oldest documents of Old Turkic (runic inscriptions from the 7th century AD) do not show any evidence of conjunction markers. Even in Modern Standard Turkish it is possible to omit conjunction markers in many syntactic contexts:

(48)   a.   Ahmet, Mehmet **ve**   ben
             Ahmet Mehmet **and**   I
             'Ahmet Mehmet, and I'

       b.   Ahmet, Mehmet, ben
             Same translation as in (49a.)

We see the general conjunction marker *ve* in (49a); this element was borrowed into the language from Semitic. (49b) shows the possibility of dispensing with a conjunction marker.

Note that the very existence of examples like (47b) is strong evidence in favor of the CC-based analysis. If the complex subject in such constructions were not a coordinate structure, as it would not be under any hypothesis which views the "comitative" phrase as a PP-adjunct for inclusive readings, we would not expect for such an adjunct (i.e. the DP marked with *ile/–lA*) to be able to determine agreement. Note that it would be curious enough for an adjunct to determine plurality by its mere existence; it would be all the more curious for a postpositional object within a PP which is itself a mere adjunct to be able to pass on its person features to the "mother node" of the subject. On the other hand, these two phenomena are perfectly predictable behavior for a conjunct in a coordinate construction; hence, the CC analysis must be the correct one.

Note also that the rejection of the Genitive marker which is otherwise found on pronominal complements of the "comitative" marker is expected under the CC analysis, while it is surprising for any hypothesis which views the comitative phrase as a PP-adjunct within the complex subject. There is no reason to expect a PP to have any different (morpho-phonological) properties when it is a VP-adjunct(or is in a chain with such a VP-adjunct position) on the one hand and when it is part of a complex DP, on the other. However, if we are not dealing with a PP adjunct but rather with a straightforward conjunct, i.e. with a BP, we would expect it to have different properties from a PP.

The fact that rejection of the Genitive by a pronoun is even stronger when that pronoun (when it is a first conjunct) determines agreement (via the Person Hierarchy) is also expected under the CC hypothesis. In such a situation, the first conjunct is a stronger determining entity than the second conjunct and thus takes on "head-like" properties, revealing the coordinate nature of the complex subject in a clear-cut manner, since Turkish has otherwise no left-headed constructions. The Genitive marking, even if it is only a semblance of case and not genuine case, would render that coordinate nature

---

[22] The (half a) question mark of (47b) can be explained by the preference (for some speakers) for having the conjunct with the higher ranking on the Person Hierarchy as the second, rather than the first, conjunct. i.e. as the phrasal head of the coordination.

opaque; rejection of the (surface) Genitive marker makes the structure transparent with respect to its true nature as a conjunct within a coordination.

   Now that we have unraveled the mysteries of the CC-nature of the SPC construction, it is clear that this structure is the source for the related PP-included reading of examples like (1), which is an instance of a CC construction with an appropriate *pro*-head, whereby the *pro* would have the features first person and singular for the "surprising reading" (reading A. in the summary of section 2.5), and the features first person and plural (reading B. in the summary of section 2.5).

3.   The CC as a Parenthetical

This brings us back to a question that was raised earlier in the paper and which has been left unanswered so far. This question was for an account of the PP-included reading for (10), which is repeated below as (49):

(49)   [Hasan-**la**]   biz   dükkân-a   git-ti-**k**.
        Hasan-**and**   we   store-DAT   go-PAST-**1.PL**
       'Hasan and I went to the store.'

As mentioned in the introductory discussion of (10), this example is ambiguous between a BP-included and a PP-excluded interpretation. The latter is not surprising and is simply due to scrambling of the comitative constituent to the beginning of the sentence. But if all BP-included readings are due to the CC-construction in Turkish, yet if the subject *biz* 'we' does not form a constituent with the (apparent) comitative element, how can we account for this reading?

   I would like to claim that (49) is a particular version of (1), namely of a construction including the CC with a (singular) *pro*-head. The CC is a parenthetical and has been scrambled from its original position after the plural subject to sentence initial position. The source for (49) would be (50):

(50)   Biz   [Hasan-**la**]   dükkân-a   git-ti-**k**.
        we   Hasan-**and**   store-DAT   go- PAST-**1. PL**
       'Hasan and I went to the store.'

To see the PP-included translation here might be surprising; we have seen similar examples earlier, claiming that they were typical examples of a PP-excluded reading, with a unique subject and a VP-dependent comitative PP in their basic word order. As a matter of fact, the surface string of (50) *is* ambiguous and does also have a PP-excluded reading. A more appropriate representation of the PP-included reading of (50) is given in (51):

(51)   Biz**,**   [Hasan-la   **pro**]**,**   dükkân-a   git-ti-**k.**
        we   Hasan-and   [+1.PRS +SG]   store-DAT   go-PAST-1.**PL**
       'We, that is Hasan and I, went to the store.'

It is important that under this analysis of the "comitative" as a conjunct within a parenthetical coordination, the plural subject pronoun which agrees with plural agreement on the verb does not form a constituent with the "comitative".

Given that this is a rather subtle issue, the judgments of four native speakers (a subset of those acknowledged in the acknowledgment footnote) were elicited. It turned out that whenever the agreement on the verb is plural, a sequence of a pronoun and a "comitative" (under either order) is well-formed only if the sequence is interrupted by a heavy intonational break, thus making it clear that the sequence does not form a constituent. This is true for both the BP-included and the PP-excluded readings.

Furthermore, it is possible to have an appropriate overt singular pronoun in the place of *pro*; while the juxtaposition of these personal pronouns, with the parenthetical pronoun referring to a subset of the non-parenthetical pronoun, is not completely felicitous, obvious downgrading of the parenthesis by intonational means or syntactic material makes the utterance better:

(52)   Biz**,** yani [Hasan-la   ben]**,** dükkân-a   git-ti -k.
         we   that is Hasan-and I   store-DAT   go-PAST-1.**PL**
         'We, that is Hasan and I, went to the store.'

(53)   Biz , (*sotto voce*: Hasan-la ben), dükkân-a git-ti-k.
         Same reading as in (52).

If this view is correct, (50)/(51) would be related to (52) and (53); (49), in turn, would be a scrambled version of (50)/(51), with the parenthetical material scrambled.

As a matter of fact, additional scrambled versions of (50)/(51) suggest that this is indeed the correct analysis of the construction in (49) under its BP-included interpretation.

The sentence-initial position in Turkish is usually the site of a topic or of shared information; thus, a better translation of (49) than the one given there would be:

(54)   As for Hasan and me, we went to the store.

Another syntactic position suited for a parenthetical is post-verbal, where afterthoughts or presupposed material are found:

(55)   Biz dükkân-a git-ti-k [Hasan-**la** pro].
         'We went to the store—that is, Hasan and me'

However, one would not expect for a parenthetical to be the sentential focus; indeed, this is not possible:

(56)   a.* Biz—[HASAN-**LA** pro]—git-ti-k dükkân-a.
         b.* Biz dükkân-a—[HASAN-**LA** pro]—git-ti-k

The immediate pre-verbal position is where focused elements are found in Turkish. The surface string depicted in (56) is grammatical, however not under the analysis outlined in

(56), but rather under the PP-excluded reading only. In other words, the ambiguity of (49) between a BP-included and PP-excluded reading is maintained under scrambling of the parenthetical to unstressed positions, but not under focusing. This is just as expected if, for the BP-included reading, we assume that the constituent marked with _ile_/-_lA_ is contained within a CC which is itself a parenthetical.

4.   Summary and conclusions

I hope to have shown in this paper that certain instances of plural agreement without an overt plural antecedent in Turkish are due to the existence of complex subjects which consist of coordinate structures rather than of a pronominal head with a PP-adjunct. In those instances where there is an overt antecedent, I have argued that a coordinate structure of the sort described here is present, as well, as a parenthetical. One of the conjuncts within such complex subjects (or within a parenthetical to a plural subject) can be an overt singular pronoun, or it can be a phonologically empty pure pronominal, namely *pro*, with the appropriate features, and crucially interpreted as singular. This would give us the reading A. in the summary in section 2.5. (Such a *pro* can also be interpreted as plural — reading B in the summary in section 2.5.)

I have also argued that Turkish does not have a PPC of the sort argued for in A. I take the existence of the PPC in languages like Tzotzil to have been convincingly established by A., as well as by other scholars for Russian, as mentioned here in passing. Its lack in Turkish just shows that constructions like (1) and (12) with their "surprising" plural agreement may have more than one source cross-linguistically. It might be interesting to investigate whether any given language that has a construction like (1) has more than one source for its PP-included reading; I leave this question open for future research.

My main concern here has been to show that in Turkish, the PPC is not a possible source for the PP-included reading for (1); furthermore, I hope to have shown that the determining elements for agreement are syntactic rather than semantic, and that they have to form constituents at Spellout. More specifically, I have proposed syntactic analyses for all three readings of (1) as summarized in section 2.5, whereby all readings can be read off straightforwardly from a syntactic structure; the *pro*, the phrasal head of the relevant coordinate structures as subjects of utterances with plural agreement, was shown to correspond in a transparent manner to the overtly singular as well as plural pronominal phrasal heads in coordinate subjects that co-occur with such plural agreement. Thus, the Turkish comitative marker has a dual nature: 1. As a genuine comitative, when it is a P and heads a PP adjunct of VP, and 2. As a special coordination marker, B, heading a BP adjunct of a coordination with a pronominal second conjunct, which is the phrasal head of the coordination. As we have seen, that phrasal head can be overt (a singular or a plural pronominal), as just explained, or it can be small *pro* (with the option of singular or plural features); in either instance, the person feature of that phrasal head will be projected to the entire coordination, thus determining the person feature of the predicate; the coordinate nature of the complex subject (whether with an overt phrasal head or with small *pro* as that head) determines the predicate's plural agreement.

References

Aissen, Judith. 1989. Agreement controllers and Tzotzil comitatives. *Language* 65. 518-536.

Borsley, Robert & Jaklin Kornfilt. 2000. Mixed extended projections. In *The Nature and Function of Syntactic Categories* ed by R. Borsley, San Diego, CA: Academic Press, 101–131.

Chomsky, Noam. 1982. *Some concepts and consequences of Government and Binding.* Linguistic Inquiry monograph 6, Cambridge, Massachusetts: MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.

Chung, Sandra. 1972. On Conjunct Splitting in Samoan. *Linguistic Inquiry* 6. 510-6.

Demirok, Ömer. 2016. On *ile* in Turkish. Ms., MIT.

Erdal, Marcel. 2007. Group inflexion, morphological ellipsis, affix suspension, clitic sharing. In *Combat pour les langues du monde — Fighting for the world's languages*, ed by Fernandez-Vest. Hommage à Claude Hagège. Paris: Harmattan, 177-190.

Erguvanlı Taylan, Eser. 1998. What determines the choice of nominalizer in Turkish nominalized complement clauses? In *Proceedings of the XVIth International Congress of Linguists* ed by Bernard Caron Oxford: Pergamon. (Paper no.220, CD Rom).

Grinder, John.1969. Conjunct Splitting in Samoan. *Linguistic Notes from La Jolla* 2.

Hale, Kenneth. 1975. Counterexamples and explanations in Navajo linguistics: Syntax. *Navajo Language Review 2.29-60.* Washington, DC: Center for Applied Linguistics.

Ionin, Tania & Ora Matushansky. 2003. DPs with a twist: A unified analysis of Russian comitatives. In *Formal Approaches to Slavic Linguistics #11: The Amherst Meeting 2002* ed by Wayles Browne, Ji-Yung Kim, Barbara H. Partee, and Robert A. Rothstein, Michigan Slavic Publications, Ann Arbor.

Kayne, Richard. 1994. *The Antisymmetry of Syntax*. (Linguistic Inquiry Monographs.) Cambridge, MA: The MIT Press.

Kornfilt, Jaklin. 2012. Revisiting 'suspended affixation' and other coordinate mysteries. In *Functional Heads: The Cartography of Syntactic Structures*, vol. 7 ed by L. Brugé, Cardinaletti, G. Giusti, N. Munaro & C. Poletto (eds Oxford: Oxford University Press, 181–196.

Kornfilt, Jaklin & Whitman, John. 2011. Afterword: Nominalizations in linguistic theory. *Lingua* 121(7). 1297–1313.

Ladusaw, William.1988. Group reference and the Plural Pronoun Construction. *Papers on the Plural Pronoun Construction and Comitative Coordination.* Syntax Research Center, University of California, Santa Cruz, 1-7.

Lakoff, George and Stanley Peters.1969. Phrasal conjunction and symmetric predicates. In *Modern studies in English*, ed. by David Reibel and Sanford Schane, 113-142. Englewood Cliffs: Prentice Hall.

Lees, Robert. 1963. *The Grammar of English Nominalizations*. Indiana University Research Center in Anthropology, Folklore, and Linguistics 12. The Hague: Mouton.

Lewis, Geoffrey. 1975. *Turkish Grammar*. (Corrected edition; originally published in 1967.) Oxford: Oxford University Press.

Munn, A. (1993) *Topics in the syntax and semantics of coordinate structures*. PhD dissertation, University of Maryland.

Munn, A. (1996) *First Conjunct Agreement Without Government: A Reply to Aoun, Benmamoun and Sportiche*; ms., Michigan State University.

Pollard, Carl, and Ivan Sag.1988. An information-based theory of agreement. *Papers from the 24th Annual Regional Meeting of the Chicago Linguistic Society, part two: Parasession on agreement in grammatical theory*, ed. by Diane Brentari, Gary Larson, and Lynn MacLeod, 236-57. Chicago: Chicago Linguistic Society.

Vassilieva, M and R. K. Larson. 2005. The semantics of the plural pronoun construction. Natural Language Semantics 13: 101 – 124.

Jaklin Kornfilt
kornfilt@syr.edu

What Turkish conditionals can teach us about the question particle[*]

Sabine Laszakovits

University of Connecticut, Storrs

1.   Introduction

This paper is concerned with an asymmetry between two instantiations of the defective copula verb *i-* in Modern Turkish: the past form *i-di* and the conditional form *i-se*. The suffixes *-di* and *-se* attach to *i-* as to any other verbal stem, and they also take the same pronominal suffixes, cf. (1) and (2).

(1) a. i   -di -k
          COP -PST -1PL
          'we were'

(2) a. i   -se -k
          COP -COND[1] -1PL
          'if we were'

(1) b. gel -di -k
          come -PST -1PL
          'we came'

(2) b. gel -se   -k
          come -COND -1PL
          'if we came'

They furthermore share the property that they like to attach to their host as clitics, in which case the initial [i] disappears after consonants, (3b), but remains visible as [j] after vowels, (4b).

(3) a. yorgun i   -di -m
          tired   COP -PST -1SG

(4) a. hasta i   -di -m
          sick COP -PST -1SG

(3) b. yorgun -∅   -du -m
          tired   -COP -PST -1SG
          'I was tired'

(4) b. hasta -y   -dı -m
          sick -COP -PST -1SG
          'I was sick'

[1]See Iatridou (2013) for arguments that *-sA* is not a conditional suffix but a correlative. I will gloss it as COND for simplicity and because I am not concerned about its semantics.

When *-DI*[2] and *-sA* attach to the verb stem, they seem to take the same morphological slot as participial tense/aspect/mood (TAM) suffixes, such as *-Iyor* PROG, *-(y)AcAG* FUT, *-(I/A)r* AOR, *-mIş* PERF, a.o. (Göksel & Kerslake's (2005, §8.2.3) position 3). As TAM suffixes, they cannot be separated from the verb stem by the polar question particle, (5), but as copular clitics, the question particle can intervene between V+TAM and *i-di*, (6a). Interestingly, this intervention is not possible when followed by *i-se*, (6b).

(5) a. *gel -mi -di -k
         come -Q   -PST -1PL

     b. gel   -di -k   -mi
         come   -PST -1PL -Q
         'Did we come?'

(6) a. gel -ecek -mi -y -di -k
         come -FUT -Q -COP -PST -1PL
         'Were we going to come?'

     b. *gel -ecek -mi -y -se -k
         come -FUT -Q -COP -COND -1PL

One might think that the reason for the unacceptability of (6) is that conditionals are islands and do not allow embedding of question elements. However, this is easily shown to not be the right analysis. In (7), a conditional clause contains a wh-word, and in (8), *-mI* is inside a conditional but attaches to a different host.

(7) Ali [Merve NE oku -r   -∅   -sa   ] mutlu ol -acak?
     A   M   what read -AOR -COP -COND   happy be -FUT
     'What is such that Ali will be happy if Merve reads it?'

(8) [BU KİTAB -ı   -mı oku -r   -∅   -sa   -m] mutlu ol -acak -sın?
     this book -ACC -Q read -AOR -COP -COND -1SG   happy be -FUT -2SG
     'Is it this book that you will be happy if I read it?'[3]

This paper is structured as follows: In section 2, I will present evidence for a movement analysis of the question particle *-mI*, according to which *-mI* can only attach to constituents that can independently undergo syntactic movement. In section 3, I use this movement test to account for (6b), whereby I try to account for a substantial amount of speaker variation.

---

[2]Capital letters indicate phonological variation: *A*=[a], [ɛ] or [e], *D*=[d] or [t], *G*=[j] ⟨ğ⟩, [ɣ] ⟨ğ⟩ or [k], *I*=[i], [ɯ] ⟨ı⟩, [u] or [y] ⟨ü⟩. Sounds in brackets indicate insertion/deletion in order to avoid hiatuses. All other transcriptions follow Turkish orthography, except that I will treat the question particle *-mI* as a suffix because it forms a prosodic unit with its host, rather than as a separate word as in standard orthography.

[3]I use clefts in the translation to disambiguate which consituent is being questioned, rather than adding prosodic information to the English translation. I do not mean to imply that *-mI*-questions are syntactically underlying clefts.

I will weaken the claims made in section 2 for certain properties of the verbal complex. Section 4 concludes. In section 5, I briefly present two further differences between *i-di* and *i-se*, but leave their analysis open for future work.

2.   The diagnostic: *-mI* as a trigger for movement

-mI can be roughly described as attaching to the constituent whose focus alternatives it introduces (Kamali 2015). In (9a), it attaches to the subject, in (9b) to a direct object, and in (9c) to an adjunct.

(9)   a.   Ali ⌐-mi¬ dün   yemek yap -tı?
             A -Q   yesterday dinner make -PST
             'Was it Ali who made dinner yesterday?'
             Focus alternatives: {Ali, Markus, Merve, … }

       b.   Feyza Derya -yı   ⌐-mı¬ öp   -tü.
             F   D   -ACC kiss -PST
             'Was is Derya who Feyza kissed?'
             Focus alternatives: {Derya, Ömer, Ahmet, … }

       c.   Ali dün   ⌐-mü¬ yemek yap -tı?
             A yesterday -Q   dinner make -PST
             'Was it yesterday that Ali made dinner?'
             Focus alternatives: {yesterday, today, last monday, … }

Apart from subjects and adjuncts, possible hosts for *-mI* include predicates (10a), possessors (10b), and for some speakers numerals (10c).

(10)   a.   Ali hasta ⌐-mı¬?
              A sick -Q
              'Is Ali sick?'

        b.   Emre [Ali -nin ⌐-mi¬ araba -sın -ı   ] al -dı?
              E   A -GEN -Q   car -3SG -ACC buy -PST
              'Was it Ali whose car Emre bought?'

        c.   %Emre [üç   ⌐-mü¬ araba] al -dı?
              E   three -Q   car buy -PST
              'Is three how many cars Emre bought?'

However, *-mI* cannot directly attach to demonstratives (11a), attributive adjectives (11b), and complements of postpositions (11c).

(11)   a.   *Emre [bu ⌐-mu¬ araba -yı] al -dı?
              E   this -Q   car -ACC buy -PST
              Intended: 'Was it THIS car that Emre bought?'

b. *Emre [beyaz -mı̄ araba -yı   ] al   -dı?
    E   white -Q   car -ACC   buy -PST
    Intended: 'Was white the color that Emre bought a car of?'

c. *Emre [Merve -mi için] araba al   -dı?
    E   M   -Q for   car buy -PST
    Intended: 'Was it Merve who Emre bought a car for?'

Instead, *-mI* has to attach distantly at a phrase that is able to host it, and must indicate its scope by prosodic means.

(12) a. Emre [BU araba -yı   ] -mı̄ al   -dı?
          E   this car -ACC   -Q buy -PST

      b. Emre [BEYAZ araba -yı   ] -mı̄ al   -dı?
          E   white car   -ACC   -Q buy -PST

      c. Emre [MERVE için] -mi araba al   -dı?
          E   M   for   -Q car buy -PST

    The observation is that those constituents that *-mI* cannot attach to are the same ones that are not able to undergo movement, while the constituents that *-mI* does attach to, can undergo movement. As an exemplification for syntactic movement to the left edge, I will use topicalization.

    Topicalization is possible for subjects (13a), objects (13b), adverbials (13c), possessors (13d), and for some speakers numerals (13e). Crucially, the speakers who like topicalization for numerals are the same ones that allow *-mI* attaching to numerals. Topics are separated by a small pause from the rest of the clause, which is indicated by a comma.

(13) a. Ali, dün   yemek yap -tı.
          A   yesterday dinner make -PST
          'Ali, he made dinner yesterday.' [4]

      b. Dün,   Ali yemek yap   -tı.
          yesterday A   dinner make -PST
          'Yesterday, Ali made dinner then. '

      c. Derya -yı, Feyza öp   -tü.
          D   -ACC F   kiss -PST
          'Derya, Feyza kissed her.'

      d. Ali -nin, Emre [t araba -sın -ı   ] al   -dı.
          A   -GEN E   car   -3SG -ACC   buy -PST
          'Ali, Emre bought his car.'

─────────

[4]I use hanging topics for the translation without meaning to imply that fronted topics in Turkish are hanging topics.

    e. %Üç, Emre [*t* araba] al -dı.
        three E   car   buy -PST
        'Three, Emre bought this many cars.'

Topicalization is not possible for demonstratives (14a), attributive adjectives (14b), and the complements of postpositions (14c).

(14)   a. *Bu, Ali [*t* araba -yı   ] al -dı.
           this A   car -ACC buy -PST

     b. *Beyaz, Ali [*t* araba -yı   ] al -dı.
           white A   car -ACC buy -PST

     c. *Merve, Ali [*t* için] araba al -dı.
           M   A   for   car buy -PST
           Intended: 'Merve, Ali bought a car for her.'

I propose to interpret this correlation between hosting *-mI* and being able to topicalize as an underlying causal connection. See also Özyıldız (to appear) for a similar conclusion.

3.   On lexical focus in conditionals

This section is concerned with the difference between *-DI* and *-sA* described in section 1: *-mI* can separate the copula from V+TAM, but *-sA* cannot. A minimal pair exemplifying this is repeated below:

(15)   balığ -ı   HAŞLA -r   =(6a)
       fish -ACC boil   -AOR -Q   -COP -PST -1SG
       'Is it boil what I would (do to) the fish?'

       balığ -ı   HAŞLA -r   [-mı] -y   [-dı] -m

(16)   *balığ -ı   HAŞLA -r   [-mı] -y [-sa]   -m   =(6b)
       fish -ACC boil   -AOR -Q   -COP -COND -1SG
       Int'd: '(So you don't like fried fish, but) if I BOILED it, (would you eat it?)'

Given the flexibility of *-mI* to attach to the constituent that it modifies, one might expect (16) to be felicitous at least in those contexts where what is questioned is the lexical content of the verb. The intended meaning of (16) indicates that even then, (16) is unacceptable. Instead, *-mI* must attach distantly and indicate its scope by prosody, (17). This is parallel to the distant attachment seen with demonstratives, attributive adjectives, and complements of postpositions, (12).

(17)   balığ -ı   HAŞLA -r   -∅ [-sa]   -m [-mı]
       fish -ACC boil   -AOR   -COP -COND -1SG -Q
       'Is it if I BOIL the fish that …'

Some speakers accept another repair, namely replacing the defective copula *i-* with the non-defective copula *ol-* that also carries the meaning 'become'. With *ol-*, *-mI* can attach below *-sA*, (18).

(18)   %balığ -ı   HAŞLA -r   -mı ol   -sa   -m
       fish   -ACC boil   -AOR -Q   COP -COND -1SG

The meanings of (18) and (17) are not completely identical. Using the auxiliary *ol-* with an aorist gives rise to a somewhat habitual reading (Ersen-Rasch 2012, §15.3). One consultant also describes an 'ironic' flavor.

   In light-verb (LV) constructions, *-mI* can attach to the nominal part under a lexical focus interpretation instead of joining the verbal complex, (19).

(19)   Ali Merve -yi   DAVET   -mi et -ti?
       A   M   -ACC invitation -Q do -PST
       'Is it invite that Ali did Merve?'

This is not excluded in conditionals:

(20)   Ali Merve -yi   DAVET   -mi ed -er   -∅   -se
       A   M   -ACC invitation -Q do -AOR -COP -COND
       'If it is invite that Ali does Merve, . . . '

(18) and (20) are another indication that in principle *-mI* is not blocked in the scope of *-sA*, but that the impossibility of (6b) and (16) must be due to a combination of idiosyncrasies of *-mI*, *-sA*, and *i-*.

   Given the diagnostic established in section 2, we make clear predictions about topicalizations of V+TAM when *-mI* can attach directly to this complex, resp. about topicalizations of the nominal part of LVCs. Speakers pattern into (at least) two groups with respect to this prediction. What I will label 'dialect A' conforms to this prediction, cf. section 3.1. The majority of speakers that I have consulted, however, do not speak dialect A. I will label their grammar 'dialect B' and discuss properties of their *-mI* in section 3.2.[5]

3.1   Dialect A: *-mI* as a trigger for movement

Given that *-mI* can attach to V+TAM in (15), where it is followed by *i-di*, the diagnostic in section 2 predicts that V+TAM should be able to undergo topicalization. This is not perfectly acceptable, but not deemed unacceptable by speakers of dialect A.

---

[5]My informants are 10 speakers of Turkish who have grown up in different areas of Turkey. Out of those 10, 2 were speakers of dialect A. I am not aware of any geographic or social factors that distinguish dialect A from dialect B. I have encountered much micro-variation in both dialects as indicated at various occasions throughout this paper. A detailed investigation about the extent of the variation and its factors remains to be undertaken.

(21) ??oku -r, kitab -ı *t* i | -di | -m
      read -AOR book -ACC   COP -PST -1SG

Given that *-mI* cannot attach to V+TAM in (16), where it is followed by *i-se*, we predict that V+TAM should not be able to undergo topicalization if it strands *i-se*. This is borne out for dialect A.

(22) *oku -r, kitab -ı *t* i | -se | -m
      read -AOR book -ACC   COP -COND -1SG

Replacing the copula *i-* with the copula *ol-* makes V+TAM topicalization better. Compare (23) with (22).

(23) ???oku -r, kitab -ı *t* | ol | | -sa | -m
       read -AOR book -ı   COP -COND -1SG

Regarding LVCs, we predict that the nominal part should be able to undergo topicalization given that it can host *-mI*. For some speakers, this holds, but others reject this construction.

(24) %davet, ablam Cihan -ı *t* et -ti
       invitation my.aunt C -ACC do -PST

Speakers who find (24) acceptable fall within our predictions. For those speakers of dialect A who don't, I propose to extend Kamali's (2011) account of *-mI* as a second-position clitic from VPs to Vs.

Kamali (2011) argues that *-mI* attaches to the constituent carrying the sentential main accent. In wide focus sentences, the main accent sits on the left-most element inside the VP, i.e. low adverbs and internal arguments, including subjects of unaccusatives, but not high adverbs or subjects of unergatives. When *-mI* modifies a VP, Kamali takes *mI*P to attract the left-most constituent inside the VP into its specifier position so that *-mI* becomes a second-position clitic seemingly inside the VP.

I propose to extend this second-position-clitichood to Vs, and I take LVCs to be complex V heads. Predicate focus in LVCs will then appear as *-mI* attaching to the nominal element, (25).

(25) Ali Bahar -ı davet | -mi | et -ti?
       A B -ACC invitation -Q do -PST
       'Is it invite that Ali did Bahar?'

This finds independent support from disambiguating the scope of predicate-final *-mI* in simplex Vs. (26), as reported by Kamali (2011, ex. (12)) has two readings, either predicate focus or verum focus.

(26)   Ali dün   iskambil OYNA -DI -mı?
       A yesterday cards   play -PST -Q
       Predicate focus: 'Did Ali PLAY this time?'
       Verum focus: 'Did Ali play cards yesterday, (as discussed before)?'

In LVCs we observe a different PF for predicate focus and verum focus. Predicate focus has
-mI as a second-position clitic on V, thereby attaching to the nominal element as in (25).
Verum focus has -mI in a higher position (in the CP-domain, by assumption), thereby
following the entire verbal complex. With simplex V, this is syncretic, but with complex V,
we predict -mI to attach to the end of the verbal complex only. This is borne out, (27).

(27)   Ali Bahar -ı   DAVET et -ti   -mi ?
       A B   -ACC invitation do -PST -Q
       Verum focus: 'Did or didn't Ali invite Bahar?'
       #Predicate focus: 'Is it invite that Ali did Bahar?'

   Let's return to example (20), repeated below as (28a). If -mI here does not attach to
davet but to the complex V *davet et-*, we predict *davet ed-er* to be able to topicalize. Cru-
cially, we predict it to be able to topicalize in contexts where simplex V cannot host -mI
or topicalize, such as before *i-se*. This is borne out. In (28b), *davet ed-er* more or less suc-
cessfully undergoes topicalization. (29) is the control with simplex V. (29a) shows that -mI
cannot attach to simplex V, and (29b) that simplex V cannot undergo topicalization. (Recall
that these data hold for a very small group of speakers, namely speakers of dialect A who
do not allow topicalization of the nominal part of LVCs.)

(28)   a.   Ali Merve -yi   DAVET -mi ed -er   -∅   -se   =(20)
            A M   -ACC invitation -Q   do -AOR -COP -COND
            'If it is invite that Ali does Merve, . . . '

       b.   ??[davet   ed -er   ], Ali Merve -yi *t* i   -se
            invitation do -AOR   A M   -ACC   COP -COND

(29)   a.   *oku -r   -mu -y   -sa   -m   =(6b)
            read -AOR -Q   -COP -COND -1SG

       b.   *oku -r,   kitab -ı   *t*   i   -se   -m   =(22)
            read -AOR book -ACC   COP -COND -1SG

Being unaware of further variation within dialect A, I hope to have thus shown that -mI
does indeed only attach to elements that can move, but that when it attaches to VP and V,
it is hosted by the first constituent inside the VP/V rather than by the entire phrase.

3.2 Dialect B: *-mI* as a second position clitic

Speakers of dialect B do not fulfill the prediction that V+TAM should be able to topicalize when it can host *-mI* — i.e., they accept (30a), but not (30b).

(30) a. kitab -ı   oku -r   -mu -y   -du -n   =(6a)
          book -ACC read -AOR -Q   -COP -PST -2SG

      b. *oku -r,   kitab -ı   *t*   i   -di   -n   =(21)
          read -AOR book -ACC   COP -PST -2SG

As established in section 1, there is a contrast between (30a) and (31), which only differ in the TAM suffix, *-DI* vs. *-sA*.

(31) *oku -r   -mu -y   -sa   -n   =(6b)
      read -AOR -Q   -COP -COND -2SG

At this point, I have nothing further to say than to stipulate that *-DI* allows reordering of *-mI* to its left and that *-sA* doesn't — in Özyıldız's (to appear) framework, *i-di* allows movement of its complement but *i-se* doesn't. Interestingly, reduction of the copula to *-∅/y* also seems to play a role as (32) with the full copula *i-* does not allow this reordering.

(32) *oku -r   -mu i   -di   -n
      read -AOR -Q   COP -PST -2SG

Note that this seems to be consistent with *oku-r* not being able to topicalize in (30b) since it would have to strand *i-di*. However, (30b) would not be improved by stranding the cliticized copula *-y-dı* instead, as shown in (33).

(33) *oku -r,   kitab -ı   *t*   -y   -dı -n
      read -AOR book -ACC   -COP -PST -2SG

Further evidence for the flexibility of *-mI* inside the verbal complex that is not connected to movement abilities of its host, comes from predicate focus under the ability suffix *-(y)Abil*. Historically, and cross-Turkically, *-(y)Abil* is a compound verb consisting of the converb *-(y)A* that still exists independently, and the main verb *bil-* 'to know'. Turkish has lexicalized most of its other compound verbs, but *-(y)Abil* is to this day fully productive.

(34) oku -ya   bil -ir -sin
      read -CONV know -AOR -2SG
      'you can read'

Given this morphological split between -(y)A and bil-, we might expect this construction to also be an instance of complex V as discussed for dialect A. We then predict that -mI should be able to attach to only the first element under predicate focus. However, while -mI can appear in this position, it does not give rise to a predicate-focus reading. My informants describe (35) as 'sarcastic' and as expressing surprise at someone's ability while mocking their previous inability. Note that in English, this context would be expressed in declarative form.

(35)   oku -ya   -mɪ bil -ir   -sin
        read -CONV -Q   know -AOR -2SG
        'You can READ?' (I thought you only ever watched movies.)

But I will set aside sarcastic expressions and leave them for future research.

4.   Conclusions

In this paper, I have argued for a movement account of the polar question particle -mI resulting in that the constituent that it attaches to as an enclitic must be able to undergo movement. I tentatively assume that mIP is left-headed and that the hosting phrase needs to move to Spec,mIP, effectively making -mI a second-position clitic. This is an extension of Kamali (2011), who assumes this for predicate-focus questions, but not for narrow-focus questions targeting specific constituents other than VP, and of Özyıldız (to appear), who assumes a similar syntax of -mI in general but does not account for the attachment pecularities of predicate focus.

I have furthermore used this account of -mI to explain an asymmetry between the two tense/aspect/mood (TAM) suffixes -DI (past tense) and -sA, which marks conditional clauses among other things. The asymmetry consists of -mI being able to attach below -DI in the verbal complex, but not below -sA. I have found that the speakers I consulted patterned into two major groups: group A fulfills the predictions made by this account alone, namely allowed other kinds of syntactic movement of the verbal complex that -mI attaches to. Group B did not allow this, but as I have argued, their grammar is compatible with a second-position clitic approach to -mI. An independent, currently unmotivated assumption is needed saying that -sA blocks the second-position clitichood of -mI.

5.   Appendix: More differences between i-di and i-se

Apart from the different behavior of -mI under i-di and i-se that was discussed in this paper, two other differences have come to my attention. I will briefly give the data here, but at this point I have no analysis to offer.

5.1   No conditional of a conditional

The second asymmetry concerns double occurrences of -DI and -sA in a single verbal complex. If -DI attaches to the verbal stem, it can attach again as a copular clitic in order

to form a past perfect, (36a). But *-sA* cannot attach to a verbal stem that already has *-sA* on it, (36b).

(36)   a.   gel -di -y -di -k
               come -PST -COP -PST -1PL
               'we had come'

        b.   *gel -se   -y   -se   -k
               come -COND -COP -COND -1PL

-sA can, however, attach to *-DI*, (37a), and *-DI* can attach to *-sA*, (37b).

(37)   a.   gel -di -y -se   -k
               come -PST -COP -COND -1PL
               'if (as you imply) we have come'

        b.   gel -se   -y -di -k
               come -COND -COP -PST -1PL
               'if we had come' (counterfactual)

The same pattern is attested when the personal agreement attaches to the main verb instead of the copula:

(38)   a.   gel -di -k -∅ -ti
               come -PST -1PL -COP -PST

        b.   *gel -se   -k -∅ -se
               come -COND -1PL -COP -COND

(39)   a.   gel -di -k -∅ -se
               come -PST -1PL -COP -COND

        b.   gel -se   -k -∅ -ti
               come -COND -1PL -COP -PST

Grammarians report a slight difference in meaning between the two attachment sites of the agreement suffix: Ersen-Rasch (2012, p. 177, p. 209) reports that attachment to the main verb continues a previous discourse, while attachment to the copula does not. Göksel & Kerslake (2005, p. 88) also report these suffix ordering possibilities.

5.2   Double ϕ-agreement depending on *-mI*

A further difference between *-DI* and *-sA* has been reported to me by Kadir Gökgöz (p.c.) that I have not found discussed in the literature.

While the pronominal suffix can attach either to the main verb or to the auxiliary, it usually cannot attach to both at the same time. This is not unexpected. This pattern holds for the past perfect (40), and for the counterfactual (41).

(40)   a.   söyle -di   ⟨-m⟩ -∅/y -di   ⟨-m⟩ ⁶
            say   -PST -1SG -COP -PST -1SG

       b.   *söyle -di   -m   -∅   -di   -m
            say   -PST -1SG -COP -PST -1SG

(41)   a.   söyle -se   ⟨-m⟩ -∅/y -di   ⟨-m⟩
            say   -COND -1SG -COP -PST -1SG

       b.   *söyle -se   -m   -∅   -di   -m
            say   -COND -1SG -COP -PST -1SG

However, under the realis past (42), double occurrence of ϕ-agreement is not out for all speakers. ((43) is given for completeness but presumably ruled out by independent reasons as discussed above.)

(42)   a.   söyle -di   ⟨-m⟩ -∅/y -se   ⟨-m⟩
            say   -PST -1SG -COP -COND -1SG

       b.   %söyle -di   -m   -∅   -se   -m
            say   -PST -1SG -COP -COND -1SG

(43)   a.   *söyle -se   ⟨-m⟩ -∅/y -se   ⟨-m⟩
            say   -COND -1SG -COP -COND -1SG

       b.   *söyle -se   -m   -∅   -se   -m
            say   -COND -1SG -COP -COND -1SG

Returning to the past perfect and the counterfactual, we find a contrast between (40b) and (44) for the past perfect, and between (41b) and (45) for the counterfactual: insertion of *-mI* into the verbal complex seems to ameliorate double agreement.

(44)   a.   *söyle -di   -m   -∅   -di   -m   =(40b)
            say   -PST -1SG   -COP -PST -1SG

       b.   ?söyle -di   -m   -mi -y   -di   -m
            say   -PST -1SG Q   -COP -PST -1SG
            "Had I said it?"

---

⁶Groups of angles brackets ⟨x⟩…⟨x⟩ indicate that the acceptability judgment applies if any single one of these occurences is realized.

(45) a. *söyle -se   -m   -∅   -di -m   =(41b)
           say   -COND -1SG   -COP -PST -1SG

      b. ?söyle -se   -m   -mi -y   -di -m
           say   -COND -1SG -Q   -COP -PST -1SG
           "What if I said it?"

I leave open the question of whether and how these properties of *-mI*, *i-*, *-sA*, and *-DI* can be unified in more general terms.

References

Ersen-Rasch, Margarete. 2012. *Türkische Grammatik. Ausführlich und verständlich. Lernstufen A1 bis C2*. Harrassowitz.

Göksel, Aslı, & Cecilia Kerslake. 2005. *Turkish. A Comprehensive Grammar*. Routledge Comprehensive Grammars. Routledge.

Iatridou, Sabine. 2013. Looking for free relatives in Turkish (and the unexpected places this leads to). In *Proceedings of the 8th Workshop on Altaic Formal Linguistics (WAFL8)*, ed. Umut Özge, number 67 in MITWPL, 129–152. MIT Press.

Kamali, Beste. 2011. The question particle in Turkish: Consequences for the interfaces. In *WCCFL 28 Online Proceedings*, ed. Mary Byram Washburn, Sarah Ouwayda, Chuoying Ouyang, Bin Yin, Canan Ipek, Lisa Marston, & Aaron Walker. URL `https://sites.google.com/site/wccfl28pro/papers/kamali%202010.pdf`.

Kamali, Beste. 2015. Clitic placement in Turkish yes/no questions: from pragmatics to morphosyntax. URL `https://www.academia.edu/14899076/`, ms, July 2015.

Kornfilt, Jaklin. 1997. *Turkish*. Descriptive Grammars. Routledge.

Özyıldız, Deniz. to appear. Move to mI, but only if you can. In *Proceedings of WAFL 11*.

Sabine Laszakovits
sabine.laszakovits@uconn.edu

Iconic templates in Turkish

Filiz Mutlu

McGill University

1.   Introduction

Iconic roots are bound in Turkish (Hatipoğlu 1971) and they become visible to syntax by means of two word formation processes: affixation and reduplication with or without ablaut (Baturay 2010). I claim these word formation processes create templates (cf. Inkelas 1993), and these templates are recursive. In contrast with the observed trend of iconicity to have more flexible phonotactics than the rest of the language (Samarin 1971, Childs 1988, Dingemanse 2012), I observe Turkish iconicity to be much more restrictive. I claim that Turkish iconic roots and templates reflect the core grammar of the language.

1.1   Iconicity

It was pointed out that the linguistic sign was arbitrary and seemingly non-arbitrary association of meaning and form such as found in onomatopoeic words was marginal (Saussure 1916/1998). However, though iconic tokens may be fewer in number than arbitrary ones, they belong in a system of their own rather than being marginal exceptions (cf. Diffloth 1979, Dingemanse 2012). A case in point is Turkish.

   A well-observed quality of iconic words is that they have more flexible phonotactics (Samarin 1971, Childs 1988, Dingemanse 2012). They can resist historical sound change, have speech sounds the rest of the language does not have or more possibilities for the combination of sounds the language has (Diffloth 1979). Turkish is a surprising exception in that the phonology of iconicity is more restrictive, with a fewer number of sounds and sound combinations.

   Reduplication and ablaut are two morphological tools used iconically in Turkish (Baturay 2010) like other languages (Doke 1935, Anderson 1998). Reduplication refers to a word formation process that repeats all or part of a word or phrase (Urbanczyk 2006). Reduplication can be transparent ("more of the same") or opaque Michelucci et al (2011). Ablaut (apophony) is a context-free non-arbitrary vowel alternation with grammatically distinctive character (Rieder & Schenner 2001). The apophonic path is a common pattern across languages (Guerssel & Lowenstamm 1996).

Iconic reduplication can follow the apophonic path (i) or it can be complete (ii):

(1)   i.   chit-chat, slip-slop, tick-tock (English, Rastall 2004)
                 piŋ1 piŋ1 paŋ1 paŋ1 'crash', ɕi1 li xua1 la1 'storm' (Mandarin, Meng 2012)

         ii.   hoep-hoep 'a bird' (Afrikaans, Huyssteen 2004),
                 pupu 'owl' (Cavinena, Guillaume 2008)
                 murmurizo 'murmer' (Greek, Fischer 2011)

2.   Turkish Iconic Roots

Turkish has bound iconic roots such as *güm-* 'bang', *löp-* 'soft and heavy', *tın-* 'slow'. The examples in this paper are taken from the onomatopoeic dictionary of Zülfikar (2005) which lists hundreds of roots and multiple words derived from each root.

The possible shape of a root is $(C_1)V(C_2)(C_3)$, like non-iconic ones. The most common form is C_1VC_2. The initial consonant of the root can be any consonant but [n, r] and the final any consonant but [h] (Özgenç 1980). Zülfikar notes [ʒ] cannot be C_1. Interestingly, these restrictions hold for verb roots as well, which are also bound (Bayırlı 2012).

Iconic roots cannot bear nominal or verbal inflection. They must undergo reduplication and/or affixation to become words. Reduplication of bound roots can be complete *çat çat*, apophonic *çat çut* or show consonant alternation *çat pat* (Baturay 2010). The output of reduplication is a nominal (Hatipoğlu 1981). Derivational suffixes on iconic roots yield both verbs and nominals.

Free roots are attested but few in number: *yırt 'to rip', pıs 'to be frightened' lop 'soft, roundish, kof 'empty, flimsy', gür 'strong, of a current'* (Standard Turkish); *çır 'to rip', mız 'to be a spoilsport', mış 'give up on', pıt 'to slip out of hand'* (dialect).

3.   Templates

There is a set of morphological operations a root goes through until it is lexicalised. I observe that these operations can be parasitic on one another or mutually exclusive. It looks like every operation comes with a slot in a template and operations requiring the same slot are incompatible with one another (cf. Inkelas 1993).

Reduplication results in creating nominals. Suffixation, however, can be expressive or derivational. I take derivational to mean that the suffix assigns a new lexical category or creates a categorical change in meaning. I will call the non-derivational suffixes extenders (cf. Ido 1999). Extenders intensify or expand the meaning of the stem in a gradient way.

3.1   Template I

Template I is made up of three extenders, a nominal deriving suffix, a verb deriving suffix and reduplication.

(2)   *Template I*

root	C3 insertion	Extender {-IR}	Extender {-t}	Der. Suffix
güm	gümb	gümbür		gümbürtü
				gümbürde
			gümbürt	

The second slot *gümb* and the fourth slot *gümbürt* are parasitic on the the third slot *gümbür*. *gümb* cannot be reduplicated or quoted with an auxiliary verb. The derivational suffixes {-tI} and {-dA} select the same base as the extender {-t} so they cannot exist together.

3.1.1   The extender {-IR}

A root can be affixed by {-Ir}/ {-Il} *tak-ır* 'rapping' (Özgenç 1980) attaching to both bases in reduplicated forms *tak-ır tak-ır* (Baturay 2010). Özgenç (1980) calls {-Ir, -Il} onomatopoeic substantive derivational suffixes, Ido (1999) calls them extenders and Zülfikar (1995) describes them as a means of deriving secondary forms of the roots. {-IR} does not derive substantives. A form like *gümbür* cannot bear nominal inflection or project into a phrase. {-IR} must be an intensifier/extender. The distribution of [-Ir] is more restricted than that of {-Il}: It does not follow [r] final roots. This does not follow from a phonotactic constraint: *üf-ür-ür* 'puff-{-IR}-AOR' is grammatical. It is a morpheme specific restriction, part of the morphological template itself.

3.1.2   The extender {-t}

The {-Ir} affixed stem can in turn be fortified by {-t}: *tak-ır-t* (Ido 1999). In fact, this is the only stem {-t} selects: *güm-b-ür-t, *güm-t.* {-t} does not select the [l] form of this extender: **pır-ıl-t.* 'sparkle'. This, however, is not a phonological restriction: *pır-ıl-tı.* 'sparkle, NOM.' is grammatical. The restriction has only to do with the selectional properties of a particular morpheme.

3.1.3   The extender {-C3}

When affixed with {-IR}, nasal and lateral final roots are fortified by a plosive to yield a homorganic sequence: *taŋ-g-ır 'crash', zım-b-ır 'dissonant sound', hal-d-ır 'rush'.* Note that the sequences *mır-* 'murmer', *lar-* *'suddenly'* are iconic roots and they are phonologically well-formed. Insertion is not a phonological requirement. It serves to fortify the iconic meaning: It is a morphological operation parasitic on the following

extender {-IR} in this template and {-I/A} in Template II. In fact, even if the final consonant is one that can stand alone, it optionally bears C_3 insertion in the form of gemination: *takır ~ takkır* 'rap', *vızır ~ vızzır* 'buzz'.

3.1.4 Reduplication

The root, the {-IR} stem and the {-t} stem can be reduplicated to yield a nominal: *fır fır* 'whirling, twirly' *fırıl fırıl* 'whirling fast', *gümbürt gümbürt* 'loud repeated banging'. Reduplication can repeat: *çat çat çat* 'repeated cracking', *tık tık tık 'repeated light rapping'*, *tin tin tin* 'pottering'. Triplication can only yield an adverb, not an adjective or a noun: Compare lexicalised *firfir* 'frill' to *fır fır fır* 'whirling fast'. It looks like the more iconic the use of reduplication, the fewer options of lexicalisation it has.

3.1.5 Derivational Suffixes

In this template, there are two derivational suffixes {-dA}, {-tI} which select only iconic roots and {-I/Ak} which is still active only on iconic roots. The verb deriving suffix {-dA} and the noun deriving {-tI} select only the {-IR} stem: *şangırda* 'to clang', *şangırtı* 'clang'. Bauer (2003, p. 85) notes that {-dA} semantically selects onomatopoeic words and phonologically those ending in a liquid. However, {-dA} never selects liquid final iconic roots: *gürde but *gürle* 'roar'. What {-dA} selects is not a sound but a base bearing only one specific morpheme: It is a morphological selection. Göksel and Kerslake (2004) mention that {-tI} is added to onomatopoeic stems to form nouns, but they do not mention that the only stem it selects is the {-IR} base.

The {-IR} stem and verbs derived from iconic roots can be suffixed with {-I/Ak} to yield a nominal: *fin-g-ir-de-k* 'flirty', *şın-g-ır-da-k* 'rattling', *çın-g-ır-ak* 'baby's rattle'. {-I/Ak} conceivably was a productive nominaliser in the history of the language, as a great number of nominals end in it, such as *dur* 'stop' and *durak* 'station'. {-I/Ak} currently seems to lead an insulated life of its own within iconic templates. All kinds of iconically derived verbs can be nominalised with {-I/Ak}: *pof-la-k* 'empty', *pört-le-k* 'bulging', *üf-ür-ük* 'breath', *cır-la-k* 'shrill', *tü-kür-ük* 'spit'. This suffix is active in templates but no longer so in plain morphology.

3.2 Template II

This template has four interdependent extenders (two of them with low/high vowel variants), the derivational suffix {-I/Ak} and reduplication. The fortifying function of low vowels contrast with the diminutive function of high vowels.

(3)   *Template II*

root	C3 insertion	Extender {I/A}	Extender {-dA}	Extender {-nA}	Der. {-I/Ak}
hop	hopp	hoppa			hoppak
			hoppada		hoppadak
				hoppadana	hoppadanak
			hoppala		hoppalak
		hoppi	hoppidi		hoppidik

The second slot *hopp* is parasitic on the third one, *hoppa/hoppi*. Every other extender is parasitic on the preceding one. (-I/Ak) selects a base with at least one of the extenders.

3.2.1   The extender {I/A}

A fourth extender, {-I/A} starts a different template. It supports the C_3 insertion slot on its left like {-IR} does: *güm-b-e*. C_3 insertion and reduplication cannot both occur with an {-I/A} stem: *pata pata* 'thumping', *lapa lapa* 'large flakes (of snow)' but **gümbe gümbe* 'bang'. It looks like the combination of C_3, following extender and reduplication is reserved for {-IR}: *gümbür gümbür* 'banging'.

3.2.2   The extenders {-dI/A}, {-nA} and {-lA}

The {-I/A} form can host the extender {-dI/A}: *gümbede, hoppidi 'hopping, sudden'*. The two extenders cannot have vowels of different height. Note that this is not for phonological reasons: Other suffixes and roots freely combine with no restrictions on vowel height. With these extenders height has an iconic function such that low vowels convey strong impressions and non-high vowels convey weaker impressions (cf. Jespersen 1922/2013).

      The resulting base in turn can host a sixth extender, {-nA}: *gümbedene, şappadana* 'splashing'. {-nA} can only attach to {-AdA} but not {-IdI}: *hoppadanak* but **hoppidinik*. I believe this is because {-nA}, as an extender building on three preceding extenders, has a fortifying function and it is semantically incompatible with {-I} which has a diminutive function. Similarly, {-A} but not {-I} can take a seventh extender {-lA}: *hoppala* 'sudden, surprising', *harala gürele* 'bustling, chaotic' but **hoppili*.

3.3.3   Reduplication and Nominalisation

Any slot in the template can be reduplicated to form a nominal (except the second slot, *hopp)*: şa*ppada(na) şappada(na), şappada(na) şuppada(na), şippidi şippidi, hoppala hoppala*. {-lA} is only reduplicated with ablaut: *çatala çutala* 'crashing' but **çatala çatala*. Any slot in the template can be nominalised with {I/Ak}: *hoppak, hoppalak, hopidik, hoppadak, hoppadanak*.

3.4   Further operations

3.4.1   The extender {-Im}

There is an eighth extender suffix, {-m}: *inim inim inle* 'moan' and *sızım sızım sızla* 'ache'. {-m} does not support C3 insertion on its left. The roots that take {-m} are always reduplicated. In other words, {-m} is parasitic on reduplication. Such roots are relatively few in number and mostly selected by a fixed verb: *tırım tırım ara(n)* 'search high and low', *mırın kırın et* 'make up excuses'.

3.4.2   Derivational suffixes {-KIR}, {-lA} and {-Ir}

Zülfikar (1995) mentions the verb-deriving suffixes {-sIr, -şIr, -kIr, gIr} *aksır, tıksır, hapşır* 'sneeze', *öksür, tohkur* 'cough', *püskür, püfkür* 'burst', *höykür, haykır* 'roar, cry' and provides a metathesis analysis for them. He takes the original suffix to be {-KIr}, which he believes historically underwent metathesis when added to sibilant final roots: *as-kır > aksır* 'sneeze'. The resulting {-SIr} then became productive by analogy. The roots these suffixes select cannot be reduplicated: **ak/s ak/s aksır, *püs püs püskür*. The suffix itself is iconic and have a connotation of bursting, explosion.

{-lA} is a verb deriving suffix (Göksel and Kerslake 2004): *tuz* 'salt*' tuzla* 'to salt'. It selects iconic and nominal roots: *gümle* 'to go bang', *fıs fısla* 'spray'. Extender-bearing stems cannot take {-lA}: **gümble *gümbürle, *gümbürtle*

Roots can be affixed with {-Ir} to form verbs: *üfür* 'blow', *bağır* 'yell'. {-Ir} is not very productive, but it exists in lexicalised forms.

4.   Discussion

4.1   Recursivity

The templates are recursive in that template II can take the extender-bearing output of Template I as input. The output *gümbürt* and the root that yielded it, *güm*, are both input to the same operation.

(4)   *Template I as input to Template II*

Template I Extenders				Template II Extenders			
root	C3	{-IR}	{-t}	Root	C3	{I/A}	{-dA}
				güm	gümb	gümbe	gümbede
güm	gümb	gümbür	gümbürt	gümbürt	gümbürtt	gümbürtte	gümbürttede

Remember that C$_3$ insertion is parasitic on the following slot that can host either {-IR} from Template I or {-I/A} from Template II, but not both: **gümbüre, *gümber*. The two suffixes occupy the same slot in the template. However, the whole string *gümbürt* can return to the root position to the left of C$_3$ insertion and go through Template II.

Previously, recursion with the inflectional suffix {-ki} has been noted by Hankamer (2004). His analysis of the phenomenon is that it "can best be understood in terms of its syntactic functions". That is, recursion still happens in syntax. Extender recursion is interesting because it is completely invisible to syntax. Extender bearing forms are not words. They cannot bear inflection or project into phrases. So the recursion must belong strictly to morphology.

In plain morphology, at least one derivational suffix has the potential to take a base bearing itself: *kitap-lık* (book.DER) 'a place to store books, bookcase'. Consider the possible but non-lexicalised form: *kitap-lık-lık* 'a place to store bookcases.'

4.2   Parasitic Slots

Most of the slots in the templates are parasitic on others: C_3 insertion only happens if followed by the extenders {-IR} or {-I/A}. The extender {-nA} is parasitic on the extender {-dA}. {-nA} and {-lA} follow only extenders with low vowels. {-t} is parasitic on {-IR}.

Any root can be used with *diye* 'like thus' or *et* 'do, go' as an instance of language mention (cf. Göksel, 2015). In fact, even extralinguistic signs such as a hand gesture can be mentioned with *diye/et*. However, some parasitic slots cannot be mentioned: **gümb diye/et, *gümbe diye/et*. These slots are completely invisible to any operation but those they are parasitic on.

4.3   Extension of templates to plain morphology

Non-iconic roots can be run through the iconic templates to create expressive morphology (cf. Zwicky and Pullum 1987). The extender {-Im} is recycled with verb roots: *erim erim eri-* 'melt-INTENSE'. It is mostly used with cognitive/psychological verbs or with emotional connotations: *sarım sarım sarıl* 'snuggle up to', *üzüm üzüm üzül* 'be sad'

{-IR} and reduplication can act on a sub-morphemic (meaningless) part of regular nominals: *orul orul oralet* 'hot drink-INTENSE', *börül börül börülce* 'vegetable, PLEASANT'.

4.4   Roots as the cognitive core of the language

Bayırlı (2012) argues that Turkish verbs (and in fact, verbs universally) are roots and not phrases until they are nominalised by some operation, which may be phonologically null. Turkish verbs display the same strict phonotactics as the iconic roots: Verbal suffixation is always harmonic and verb roots do not have clusters of a larger set than sonorant+obstruent stop. Neither verbs nor iconic roots contain the sound [ʒ]. Neither verbs nor iconic roots begin with a [r] or [n] or end in [h].

The roots, whether iconic are not, are the core of the language where the restrictions on the shape of objects arise. One important difference between verbal and iconic roots is that the former is a closed class but the latter is not. Novel iconic roots are produced on the go. While the strict phonotactics of verbs could be attributed to the fact that verbs as a

closed class have a frozen grammar, the productive nature of iconicity means that it is not a historical accident but synchronic restrictions at play.

The core patterns of a language are observable in the way roots are utilised: For English, zero-derivation is 'a particularly productive process' (Lieber 2005). I believe it is the default process with novel words: google > to google, but not for instance *googlise. The latter would have a different meaning than 'to search using google' even if it existed. Such patterns reveal different linguistic strata without any historical information.

It looks like Turkish iconic roots and verbs have similar qualities by virtue of being roots not words. However, verbal roots are inflectable: *gel-di* 'come-PAST' but iconic roots are not *puf-tu* 'puff-PAST'. English iconic roots are inflectable just like verbal ones: *whooshed*. This raises the question: Are there levels of rootness or boundness in Turkish? Considering cases where the bound stem is not even mentionable the way extra-linguistic objects are, such as *gümb diye*, some parasitic slots are also even more 'bound' than the root itself.

5.   Conclusion

Turkish iconic roots are lexicalised through templates. Morphological operations come with a number of designated slots that have to be filled in by other operations. For instance, the suffix {-Im} comes with an empty slot that must be filled by reduplication. C$_3$ insertion comes with a slot on its right that must host either {-Ir} or {I/A} but not both. Those extenders use the same slot and cannot exist together.

The templates are recursive; the output of one template can be the input of another. Recursion happens at a level invisible to syntax and it is purely morphological.

Roots, both iconic and verbal, belong to a deeper stratum and have the most restrictive phonotactics. Cross-linguistically, iconicity is observed to have more flexible phonotactics. Turkish iconicity is exceptional in being less flexible. The templatic nature of derivation and the phonotactics must be related as the roots are used through these templates, not merely quoted or mentioned as extra-grammatical objects.

Iconic roots are produced on the fly every day and can refer to sound, motion, shape, emotion. They are the sole group of objects selected by a whole set of morphological operations, and support only a subset of the phonotactic possibilities of the language. I believe they form a lexical category like the ideophones of African languages or Japanese mimetics.

<div align="center">References</div>

Anderson, Earl R. 1998. *A grammar of iconism.* Fairleigh Dickinson University Press.
Baturay, Semra. 2010. Apophony in Turkish Onomatopoeic Reduplication. Master's Thesis, Boğaziçi University.
Bauer, Laurie. 2003. *Introducing linguistic morphology.* Edinburgh University Press.
Bayırlı, İsa Kerem. 2012. On suffixhood and verbalness: A Mirror Theoretic approach. Master's Thesis, Boğaziçi University.
Childs, George Tucker. 1988. The phonology of Kisi ideophones. *Journal of African Languages and Linguistics* 10(2): 165–190.

Diffloth, Gérard. 1979. Expressive phonology and prosaic phonology in Mon-Khmer. In *Studies in Mon-Khmer and Tai Phonology and Phonetics in Honor of Eugénie J. A. Henderson,* ed. T.L. Thongkum, V. Panupon, P. Kullavanijaya, & M.R. Kalaya (Eds.), pp. 49–59. Bangkok: Chulalongkorn University Press.

Dingemanse, Mark. 2012. Advances in the cross-linguistic study of ideophones. *Language and Linguistics Compass*, 6(10), 654–672.

Doke, Clement M. 1935. *Bantu linguistic terminology.* London: Longmans, Green, and Co.

Göksel, Aslı. 2015. Phrasal compounds in Turkish; Distinguishing citations from quotations. *STUF - Language Typology and Universals*, 68(3), pp. 359–394.

Göksel, Aslı, & Kerslake, Celia. 2004. *Turkish: A comprehensive grammar*. Routledge.

Guerssel, Mohand. & Jean Lowenstamm. 1996. Ablaut in Classical Arabic measure I active verbal forms. In *Studies in Afroasiatic grammar*, ed. Lecarme, J., J. Lowenstamm & U. Shlonsky (eds.), Holland Academic Graphics, The Hague, pp. 123–134.

Guillaume, A. 2008. *A Grammar of Cavineño*. New York: Mouton de Gruyter.

Hankamer, J. 2004. An ad-phrasal affix in Turkish. *MIT Working Papers in Linguistics*, 46, pp. 289–299.

Hatipoğlu, V. 1971. *Türk Dilinde İkileme.* Türk Dil Kurumu Yayınları.

Hatipoğlu, V. 1981. *Türk dilinde ikileme*. TDK tanıtma yayınları.

Huyssteen, G Van. 2004. Motivating the composition of Afrikaans reduplications: A cognitive grammar analysis. *Studies in Linguistic Motivation*, 269.

Ido, S. G. 1999. Turkish mimetic word formation. *Asian and African Studies*, 8, 67–73.

Inkelas, S. 1993. Nimboran Position Class Morphology. *Natural Language & Linguistic Theory* 11(4), pp. 559–624.

Jespersen, O. 2013. *A Modern English grammar on historical principles: Part II - Syntax (First Volume)*. Routledge.

Kaye, J., Lowenstamm, J. & Vergnaud, J. R. 1990. Constituent structure and government in phonology. *Phonology*, 7(2), pp 193–231.

Meng, Chenxi. 2012. A description of ideophonic words in Mandarin Chinese. Master's Thesis, Universiteit Leiden.

Michelucci, P., Fischer, O. & Ljungberg, C. (eds.). 2011. *Semblance and Signification.* John Benjamins.

Özgenç, F. M. 1980. Onomatopoeia in Turkish: A prosodic analysis. Master's Thesis, Boğaziçi University.

Rastall, Paul. 2004. Playful English: kinds of reduplication. *English Today*, 20(4), pp. 38–41

Rieder, E. & Schenner, A. 2001. Ablaut and Umlaut: what they have in common, what distinguishes them. In *Proceedings of ConSOLE X*, ed. M. van Koppen, J. Sio, & M de Vos.

Samarin, W. J. 1971. Survey of Bantu ideophones. *African Language Studies*, 12, pp. 130–168.

Saussure, F. 1998. Nature of the Linguistic Sign. In *The Critical Tradition: Classic Texts and Contemporary Trends*. ed. by David H. Richter. Boston: Bedford/St. Martin's Press, 832–835.

Štekauer, P., & Lieber, R. (Eds.). 2006. *Handbook of Word-formation* (Vol. 64). Springer Science & Business Media.

Urbanczyk, S. 2007. Reduplication. In *The Cambridge Handbook of Phonology*, ed. Paul de Lacy, pp. 473–494, Cambridge University Press.

Zülfikar, H. 1995. *Türkçede Ses Yansımalı Kelimeler.* Ankara: TDK Yayınları.

Zwicky, A. M., & Pullum, G. K. 2011. Plain morphology and expressive morphology. In *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* (Vol. 13).

Filiz Mutlu
filizmutlu@yandex.com

A Locality Restriction on Indexical Shift: Evidence from Turkish*

Matthew Tyler

Yale University

1.   Introduction

In English, 1st and 2nd-person pronouns, or *indexical* pronouns, always refer to the speaker or hearer of the utterance respectively. But in some languages, an indexical pronoun embedded under a propositional attitude verb may refer instead to the author or addressee of the proposition, as in (1), from Zazaki (Anand & Nevins 2004).

(1)   Rojda ne va kɛ **mɨ** kes   paci kɛrd.
        Rojda not say that I   anyone kiss did
        "Rojda_i didn't say that **she_i** kissed anyone."

There has been much recent work on the syntax and semantics of indexical shift in various languages (Schlenker 2003, Anand & Nevins 2004, Podobryaev 2014, Shklovsky & Sudo 2014), and much of this work has come to the conclusion that there is a particular operator responsible for shifting the context on which indexical pronouns rely for their reference. However, it has generally been assumed that when an indexical pronoun finds itself in the scope of a context-shifting operator, it is obliged to pick up its reference from the context introduced by that operator, and it does so with no further restrictions. In this paper, I show that, at least for some speakers of Turkish, (a) not all pronouns in the scope of the operator must shift, and (b) the ability of a pronoun to shift can be blocked by an intervening unshifted pronoun. Such patterns are problematic for analyses of indexical shift that make use solely of context-shifting operators. I propose that at least for some speakers, indexical shift should be considered a special case of regular pronominal binding, and that intervention effects are instances of the *De Re Blocking Effect* in action.

The paper is organized as follows. Section 2 discusses the status of indexical shift in Turkish, and lays out the empirical contribution of this paper – the existence of violations of *Shift Together*, a supposedly universal principle governing the distribution of shifted

---

indexicals. The data appear to show a kind of 'intervention effect'. Section 3 provides an analysis in which indexical shift is considered pronominal binding, and the intervention effects are accounted for in terms of binding competition (Fox 2000). Section 4 provides cross-linguistic evidence for the proposal from Tsez, and Section 5 concludes.

2.   Indexical shift in Turkish

Indexical shift in Turkish has previously been discussed in two works: Şener & Şener (2011) and Özyıldız (2012). They are largely compatible, although they differ in one major respect: Şener & Şener claim that only null pronouns may be shifted, while Özyıldız claims that both null and overt pronouns may be shifted, providing examples such as (2).

(2)   İnan Ayşe-'ye [**sen-inle ben-i** nere-ye   ata-yacak-lar]   de-miş?
       Inan Ayşe-DAT [you-COM I-ACC where-DAT appoint-FUT-3PL] say-DUB

       "Where did Inan$_i$ say to Ayşe$_j$ that they would appoint {you and me/**him$_i$ and her$_j$**}?"

The judgments shown in the remainder of this paper come from a speaker of the most permissive variety of Turkish, with respect to indexical shift. The speaker allows shifting in finite clauses embedded under both *sanmak* 'believe' and *demek* 'say', he allows both 1st and 2nd-person pronouns to be shifted, and he allows both null and overt pronouns to be shifted.[1] In the rest of this section, we see that the speaker also reports a particular set of very revealing judgments that shed light on the nature of indexical shift.

2.1   *Shift Together* and its violations

Anand & Nevins (2004), who investigate the behavior of indexical shift in Slave and Zazaki, propose the generalization in (3).

(3)   *Shift-Together Constraint*
       All indexicals within a *speech-context domain* must pick up reference from the same context.

That is, where multiple indexical expressions find themselves in the scope of a the same attitude predicate, either they *all* receive shifted interpretations, or none of them receive shifted interpretations. 'Mixed' interpretations, where clausemate indexical pronouns are interpreted according to different contexts, are ruled out by (3).

In Anand & Nevins's analysis, Shift Together is a natural consequence of how indexical shift works, and so we should not expect to find it violated in any language. This has largely been corroborated in studies of other languages with indexical shift, including Amharic

---

[1]There is no doubt much more variation in indexical shift in Turkish. The results of some preliminary surveys show that there is variation in terms of which predicates license shifting: speakers pattern differently in whether they allow shifting under *sanmak* 'believe', *demek* 'say', and other attitude predicates. There may also be variation in whether or not 2nd-person pronouns may be shifted.

(Schlenker 2003), Uyghur (Shklovsky & Sudo 2014), Mishar Tatar (Podobryaev 2014) and Tsez (Polinsky 2015), all of which adhere to Shift Together.[2]

Özyıldız (2012), however, presents the paradigm in (4), which shows a Shift Together violation that is grammatical in his variety of Turkish. In (4), reading (i) involves no shifting. Reading (ii) shifts both pronouns, meaning that it obeys Shift Together and so is acceptable, as predicted. Reading (iv) violates Shift Together and is, unsurprisingly, disallowed. But the interesting data point is reading (iii), which is judged acceptable despite violating Shift Together.

(4) Tunç Ayşe-'ye [**ben sen-i**   nere-ye   götür-eceğ-im] de-miş?
     Tunç Ayşe-DAT [I   you-ACC where-DAT take-FUT-1SG] say-DUB

   "Where did Tunç$_i$ say to Ayşe$_j$ that...   i.   I would take you?"
                                                ii.   he$_i$ would take her$_j$?"
                                                iii. he$_i$ would take you?"
                                                iv.   *I would take her$_j$?"

The generalization seems to be that the lower (earlier) pronoun may shift *only* if the higher (later) pronoun shifts. We can show that it is not a simple asymmetry between 1st and 2nd-person pronouns by considering the sentence in (5), where the 1st and 2nd-person pronouns have been switched. Here, we see again that the lower pronoun may shift only if the higher pronoun shifts (the equivalent readings (i) and (ii) are left out, but they are both grammatical).

(5) Tunç Ayşe-'ye [**sen ben-i** nere-ye   götür-eceğ-im] de-miş?
     Tunç Ayşe-DAT [you I-ACC where-DAT take-FUT-1SG] say-DUB

   "Where did Tunç$_i$ say to Ayşe$_j$ that...   iii.   she$_j$ would take me?"
                                                iv.   *you would take him$_i$?"

We can also show that this is not simply an asymmetry between subject and non-subject – when the direct and indirect object are both indexical pronouns, we see the same pattern as before, where the lower pronoun may shift only if the higher pronoun shifts:

(6) Tunç Ayşe-'ye [patron **ben-i sen-inle** nere-de   tanış-tır-acak]   de-miş?
     Tunç Ayşe-DAT [boss I-ACC you-COM where-LOC meet-CAUS-FUT] say-DUB

   "Where did Tunç$_i$ say to Ayşe$_j$ that the boss would introduce   i.   me to you?"
                                                                     ii.   him$_i$ to her$_j$?"
                                                                     iii. ?him$_i$ to you?"
                                                                     iv.   *me to her$_j$?"

This pattern can be stated another way: an indexical pronoun may only shift if there is no *un*shifted pronoun intervening between the indexical pronoun and the DP serving as

---

[2]It is worth noting that in Mishar Tatar, apparent Shift Together violations are possible if one pronoun is null, and therefore a member of the class of shiftable pronouns, while the other pronoun is overt, and therefore a member of the class of *un*shiftable pronouns. However, such examples do not undermine Anand & Nevins's analysis as, according to Podobryaev (2014), the unshiftable pronouns in Mishar Tatar are assigned reference by a distinct mechanism.

its referent (where 'intervene' may, for now, be interpreted as involving structural or linear intervention). So in (6), *seninle* 'with you' may only shift when *beni* 'me' is shifted – this ensures that there is no unshifted indexical pronoun intervening between *beni* 'me' and the DP that serves as its referent, *Tunç*.

It is also important to reiterate once again that these judgments are not shared by all, or even most, Turkish speakers. These are the judgments of an individual with the most permissive possible constellation of properties related to indexical shift – however, I predict that all speakers who allow Shift Together violations should find the (iii) readings of (4-6) more natural than the (iv) readings. The next section presents an analysis of the grammar that generates this pattern.

3.   Analysis

In the previous section we saw that for speakers who allow Shift Together violations, an indexical pronoun may be shifted only if there is no *un*shifted pronoun that intervenes between the shifting pronoun and the DP serving as its referent (with 'intervene' yet to be refined). This pattern bears a strong resemblance to the *De Re* Blocking Effect (Anand 2006), indicating that indexical shift may be collapsible with regular pronominal binding.

I first compare the asymmetry found in indexical shift to the *De Re* Blocking effect, before moving onto how shifting and binding might be given a unified analysis.

3.1   The *De Re* Blocking Effect
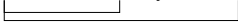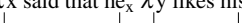
The *De Re* Blocking effect (Anand 2006) states that no obligatory *de se* anaphor can be c-commanded by a *de re* counterpart. One consequence of this is the well-documented asymmetry between the 1st-person pronouns in (7), from Lakoff (1972). The only possible interpretation is the one in which the object pronoun is interpreted *de re* (as George), and the subject *de se* (as Brigitte). If the object was interpreted *de se* instead, then the subject would necessarily be *de re*, and thus we would end up with a configuration in which a *de se* pronoun is c-commanded by a *de re* pronoun – a configuration explicitly ruled out by the Blocking Effect.

(7)   [George:] I dreamt I was Brigitte Bardot and I kissed me.

    a.   In the dream Brigitte kisses George.

    b.   # In the dream, George kisses Brigitte.

Another phenomenon that Anand argues to follow from the *De Re* Blocking Effect is Dahl's puzzle (Dahl 1973). The puzzle is that of the four logically possible interpretations of the ellipsis site in (8), only three are available. This can be restated as a restriction against a pronoun construed *strictly* c-commanding a pronoun construed *sloppily*. Anand (2006), following Fox (2000), argues that reading (d) is unavailable thanks to the *De Re* Blocking Effect, albeit indirectly.

(8)   John said that he likes his mother. Bill did too.

    a.   Bill said that John likes John's mother.

    b.   Bill said that Bill likes Bill's mother.

    c.   Bill said that Bill likes John's mother.

    d.   # Bill said that John likes Bill's mother.

The reasoning is as follows. Fox's *Rule H* (which I refer to henceforth as *binding locality*) states that where a pronoun could be bound by multiple possible antecedents, and the possibilities would be 'semantically equivalent', that pronoun must be bound by the *most local* possible antecedent. This means that if the pronoun *his* in (8) is construed sloppily (i.e. requires a binder), only the embedded subject pronoun *he* may serve as its binder – the matrix subject *John* is not the most local possible antecedent, and so cannot directly bind *his*. The possible and impossible binding configurations of the first sentence in (8) are shown in (9), from Anand (2007).

(9)   a.   * John λx said that he$_x$ λy likes his$_x$ mother.

    b.   John λx said that he$_x$ λy likes his$_y$ mother.

The binding relations in the elided VP would then have to match those in the anteceding sentence, thanks to a condition enforcing Parallelism. This explains why reading (d) is ruled out – in order to get this interpretation, the lowest pronoun *his* would have to be bound by the matrix subject, with the strictly-construed pronoun *he* intervening between them.

Crucial to Anand's explanation is the notion of binding locality. Anand proposes that the underlying motivation for binding locality is the same as the underlying cause of the *De Re* Blocking Effect – obligatory *de se* pronouns (of which *me* in (7) and *his* in (8) are both instances) are marked for binding by an operator, and it is this binding operation that cannot be intervened.

Phenomena that Anand also attributes to the *De Re* Blocking Effect include intervention of long-distance reflexives by indexical and deictic pronouns in Mandarin (Zushi 1995, Anand & Hsieh 2005), and pronoun obviation in the presence of logophoric pronouns in Yoruba (Adesola 2005). In each of these phenomena, the banned or blocked configuration is the same: an unbound, antecedentless element intervenes between a variable-like pronominal element and its binder.

The similarities between the *De Re* Blocking Effects outlined here and the indexical shift asymmetry in the previous section are clear: in both cases there is a relation between a pronoun and its binder or antecedent, and in both cases that relation is blocked by an intervening *unbound* (free) pronoun. Given that the *De Re* Blocking Effect is characteristic of operator-variable relations where the variable is a pronoun, this seems like good reason to attempt a unification of 'ordinary' pronominal binding and indexical shift. In the next section, I provide an analysis of indexical shift in (dialectal) Turkish that collapses it with 'ordinary' pronominal binding.

3.2   Indexical shift as binding

Following Anand (2006, 2007), I assume that there is a class of pronouns that need to be bound by an operator. For Anand, this is the class of *obligatory de se* pronouns – he argues that indexically-shifted pronouns acquire their referents via a different mechanism (see Anand & Nevins 2004). However, I propose that at least in some languages, pronouns that are indexically-shifted should be treated in the same way as obligatory *de se* pronouns.[3] That is, they both must be bound by an operator.

What exactly is this operator? Following earlier work in Tyler (2015), I propose that the operator responsible for binding the shifted indexical is the λ-operator associated with the subject or indirect object argument of the embedding attitude predicate.[4] The binding relation between a shifted indexical and its antecedent in the Turkish variety described here is schematized in (10), using English words.

(10)   İnan λx [I$_x$ clapped] said.

This operator-variable relationship is equally susceptible to intervention as the other binding relations Anand discusses. In (11), an unbound (and so unshifted) indexical pronoun intervenes between the bound (shifted) indexical and its antecedent, and the result is ungrammaticality. This is the configuration found in the unavailable reading (iii) in (4-6).

(11) * İnan λx [you [me$_x$ where take]] said.

However, there are two apparent difficulties with simply assimilating these to other cases of the *De Re* Blocking Effect. Recall that the effect is a consequence of binding locality, and under binding locality, binding configurations are only ruled out (a) if there is an intervening element which could also bind the variable and (b) if having the variable bound by one potential binder would be 'semantically equivalent' to having it bound by the other. Adopting this explanation would mean that (12a), equivalent to (11), is ruled out because (12b) is both available, and 'semantically equivalent' to (12a).

(12)   a.   * İnan λx [you λy [me$_x$ where take]] said.

          b.   İnan λx [you λy [me$_y$ where take]] said.

To assimilate the indexical shift asymmetries to other *De Re* Blocking Effects would therefore require not only that *you* can bind *I* and vice versa, but also that, for the purposes of evaluating binding configurations, indexically-shifted pronouns (e.g. *me* in (12a)) are considered 'semantically equivalent' to unshifted pronouns (e.g. *me* in (12b)).

---

[3]I do not claim that the mechanism proposed here should be adopted as an analysis for *all* indexical shift. Just as Anand proposes that there are multiple routes to *de se*-hood (see also Maier 2011), there may also be multiple ways to shift an indexical. See also the analysis of Tsez in Section 5.

[4]In the analysis in Tyler (2015), the person-feature discrepancy between the binder and the bindee is not a problem since the bound indexical is transmitted its 1st or 2nd-person feature from the attitude verb. For an alternative view, in which the attitude verb itself binds shifted indexicals, see von Stechow (2003).

Taking the first problem first, Charnavel (2015) uses the following data to argue that 1st-person pronouns can indeed bind 2nd-person pronouns and vice versa. In (13a), the interpretation of the pronoun in the ellipsis site is not strict (this gives rise to reading (13b)), nor is it clearly in a sloppy relation with an antecedent (this is because the only available antecedent is the 1st-person pronoun in subject position, with which *you* does not share person-features). Rather, the pronoun in the ellipsis site is interpreted as relying on a "dependent interpretation of *you* with respect to *I*".

(13)   [Romeo to Juliet:] I love you.
        [Juliet to Romeo:] I do too.

   a.   Juliet loves Romeo.

   b.   Juliet loves herself.

Interpretation (a) of the focus construction in (14) shows a similar dependency between *I* and *you*. I follow Charnavel in taking these to be instances of binding.

(14)   [Tom to Sue, in a ballroom dancing class:] Only I made you swirl.

   a.   No other dancer makes his partner swirl.

   b.   No other dancer makes Sue swirl.

Turning to the second problem, are shifted indexicals 'semantically equivalent' to their unshifted counterparts, for the purposes of evaluating binding configurations in a binding locality framework? (Anand 2006, 2007) shows that *de se* ascription is not taken into account during binding competition – this must be true in order for the dream report example in (7) to be derived via binding locality. However, Anand does maintain that in order to function as semantically equivalent, two pronouns do still need to be *denotationally* equivalent – that is, they must both refer to the same individual. I propose relaxing the definition of 'semantic equivalence' for binding competition, and allowing indexicals to be considered equivalent regardless of the individual they identify. I leave the justification and formalization of this notion to future research.

A consequence of the binding approach to indexical shift, in contrast to the approach based on context-shifting operators (Anand & Nevins 2004), is that it does not provide any explanation for Shift Together (3). While we must allow that not all speakers enforce Shift Together, we still require an explanation for why it seems so robust generally. I speculate that the route to indexical shift expounded here, in which shifting is a special case of binding and in which Shift Together is not predicted as a consequence, may co-exist alongside the route to indexical shift that involves a context-shifting operator, and in which Shift Together *is* predicted.

In the next section, I use Tsez data from Polinsky (2015) to provide further evidence that indexical shift may be a special case of pronominal binding, and that we may require two routes to indexical shift.

4.   Shifting as binding: evidence from Tsez

Polinsky (2015) shows that Tsez has indexical shift, as in (15a), and long-distance reflexives, as in (15b).

(15)   a.   *Indexical shift*[5]

Mariyat-ä   ʕarza boy-s   [**di**   magazine-y-ā-ɣor
Mariyat-ERG complain-PST [I.ABS(.II) store-OBL-LOC-VERS
y-ik'-ān=ƛin].
II-go-FUT.DEF=QUOT]

"Mariyat$_i$ complained that I/**she$_i$** must go to the store."

b.   *Long-distance reflexive*

Kid-b-ä   tungi   [**nelätow**   teł
girl-OBL-ERG jug.ABS.II [LD.REFL.LOC.ESS inside
zäw-ru]-zo   łi-d   esay-s.
be.PST-PST.PTCP]-ATTR.OBL water-INS wash-PST

"The girl washed the jug with the water that was in **itself**."

Polinsky also shows that long-distance reflexives may be based on 1st and 2nd-person pronouns, and that these may be bound by anteceding 3rd-person DPs. That is to say, long-distance reflexives may receive shifted interpretations (and indeed, they cannot receive unshifted interpretations in the absence of a binder that shares the same person-features):

(16)   Nes-ä   [**däqtow**   q'ʕanoquno ƛeb yoł=ƛin]   eƛi-s.
DEM-ERG [LD.REFL.1SG forty   year be.PRES=QUOT] say-PST
"He$_i$ said that **he$_i$**/*I was 40 years old."

If Tsez indexical shift works like it does in the Turkish variety discussed here, then this pattern is expected. Under this model, indexical shift is essentially pronominal binding. Therefore, when a pronoun is reflexive and so *must* be bound, if its binder is a 3rd-person DP it necessarily receives a shifted interpretation. We also predict that we should see a similar pattern of intervention to that found in Turkish: that is, it should not be possible for an unbound indexical pronoun to intervene between a bound pronoun and the operator which binds it. And, to some extent, Polinsky's data provides some evidence of this. In (17a), we see that a non-reflexive indexical pronoun *dow-ƛ'o-r* 'with you' may happily co-exist in the same clause as a shifted reflexive indexical pronoun *ditow* 'I-self'. In (17b), however, the non-reflexive pronoun *di* 'I' intervenes between the reflexive pronoun *dowƛ'ortow* 'with yourself' and its binder, and the resulting sentence is degraded.

---

[5]Polinsky uses language-internal diagnostics to show that (15a) is an instance of clausal embedding rather than quotation.

(17) a. Irbaħin-ä   Zarema-q-or   [**ditow**   dow-ƛ'o-r   bixzi
          Ibrahim-ERG Zarema-POSS-LAT [LD.REFL.1SG.ABS you-SUPER-LAT angry
          oq-si=ƛin]   eƛi-s.
          become-PST=QUOT] say-PST
          "Ibrahim$_i$ told Zarema$_j$ that **he$_i$** was angry with her$_j$"

     b. ? Irbaħin-ä   Zarema-q-or   [di   **dowƛ'ortow**
          Ibrahim-ERG Zarema-POSS-LAT [1SG.ABS LD.REFL.2SG.SUPER.LAT
          bixzi oq-si=ƛin]   eƛi-s.
          angry become-PST=QUOT] say-PST
          "Ibrahim$_i$ told Zarema$_j$ that he$_i$ was angry with **her$_j$**"

The degraded nature of (17b) can be blamed on the *De Re* Blocking Effect – an unbound pronoun intervenes between the bound (shifted) pronoun and its antecedent. However, this explanation crucially relies on the non-reflexive pronoun *di* 'I' being unbound. If *di* was bound, then it would not function as an intervener, and there would be nothing odd about (17b). Because *di* is both shifted and unbound, it then follows that there must be an alternative route to indexical shift available in Tsez. Given that Polinsky describes Tsez as exhibiting Shift Together, I assume that Tsez also makes use of context-shifting operators to shift indexicals.

5.   Conclusions

We have seen that Shift Together is not a universal fact about indexical shift, and that some speakers treat certain Shift Together violations as grammatical. Crucially, the impossible Shift Together violations are always the ones where an unshifted pronoun intervenes between the shifted pronoun and its antecedent. In this way, they resemble the cases of pronoun binding that are ruled out by the *De Re* Blocking Effect, itself a reflex of a rule of binding locality. As such, I have proposed that speakers who allow violations of Shift Together do in fact have a route to indexical shift that involves pronoun binding.

Note that so far, I have left ambiguous the structural framing of the binding relationship that may be 'intervened'. This is because it does not clearly involve either linear precedence or c-command: if the indirect object in (6) is scrambled over the direct object, as in (18), the original judgments remain in place. That is to say, there is a kind of local scrambling that neither feeds nor bleeds the shiftability of a pronoun:[6]

(18) Tunç Ayşe-'ye [patron **sen-inle ben-i** nere-de   tanış-tır-acak]   de-miş?
      Tunç Ayşe-DAT [boss   you-COM I-ACC where-LOC meet-CAUS-FUT] say-DUB

      "Where did Tunç$_i$ say to Ayşe$_j$ that the boss would introduce   iii.   ?him$_i$ to you?"
                                                                          iv.   *me to her$_j$?"

---

[6]The judgments for (18) are somewhat trickier than for the other examples. This stems from the fact that there is a competing reading of (18) in which *sen-inle ben-i* 'you-COM me-ACC' is interpreted as a coordinated phrase ('you and me'), rather than two separate arguments of the verb. Özyıldız (2012) shows that indexical pronouns in coordinated phrases must shift together, and so reading (iii) is only available under the *non*-coordinated interpretation.

This indicates that the relevant relation is computed over some abstract representation – the exact nature of this representation is a topic for future work.

References

Adesola, Oluseye. 2005. Pronouns and null operators: A-bar dependencies and relations in Yoruba. Doctoral dissertation, Rutgers University.

Anand, Pranav. 2006. De De Se. Doctoral dissertation, MIT.

Anand, Pranav. 2007. Dream report pronouns, local binding, and attitudes de se. In *Proceedings of SALT*, volume 17, 1–18.

Anand, Pranav, & Feng-fan Hsieh. 2005. Long-distance reflexives in perspective. In *Proceedings of WCCFL*, volume 24.

Anand, Pranav, & Andrew Nevins. 2004. Shifty operators in changing contexts. In *Proceedings of SALT*, volume 14, 20–37.

Charnavel, Isabelle. 2015. Let You Be Bound to Me (and Me to You). In *Proceedings of WCCFL*, volume 33.

Dahl, sten. 1973. On so-called sloppy identity. *Synthese* 26:81–112.

Fox, Danny. 2000. *Economy and semantic interpretation*. MIT Press.

Lakoff, George. 1972. Linguistics and natural logic. In *Semantics of natural language*, ed. D Davidson & G Harman. Dordrecht: Reidel.

Maier, Emar. 2011. On the roads to de se. In *Proceedings of SALT*, volume 21, 393–412.

Özyıldız, D. 2012. When I is not me: A preliminary case study of shifted indexicals in Turkish. ENS.

Podobryaev, Alexander. 2014. Persons, imposters, and monsters. Doctoral dissertation, MIT.

Polinsky, Maria. 2015. Embedded finite complements, indexical shift, and binding in Tsez. *Languages of the Caucasus* 1.

Schlenker, Philippe. 2003. A plea for monsters. *Linguistics and philosophy* 26:29–120.

Şener, Nilüfer Gültekin, & Serkan Şener. 2011. Null subjects and indexicality in Turkish and Uyghur. In *Proceedings of the 7th Workshop on Altaic Formal Linguistics*, 269–284.

Shklovsky, Kirill, & Yasutada Sudo. 2014. The syntax of monsters. *Linguistic Inquiry* 45.

von Stechow, Arnim. 2003. Feature deletion under semantic binding: tense, person, and mood under verbal quantifiers. In *Proceedings of NELS*, volume 33.

Tyler, Matthew. 2015. Reflexes of locality and A'-movement in indexical shift. Yale University.

Zushi, Mihoko. 1995. Long-distance dependencies. Doctoral dissertation, McGill University.

**An Ultrasound Study of the Articulatory Correlates of Vowel Anteriority
in three Turkic Languages**[*]

Jonathan North Washington

Indiana University

1.   Introduction

This paper examines the extent to which a difference in tongue body and tongue root
position is correlated with the vowel anteriority ("backness") contrasts in three Turkic
languages—Kazakh, Kyrgyz, and Turkish—using ultrasound imaging of the tongue. The
motivation for this study is the assertion by Vajda (1994) that Kazakh's "front-back" vowel
contrast is actually associated with the position of the tongue root, and not the position of
the tongue body as normally assumed. There appears to be no mention in the literature
on any other Turkic language of a tongue-root contrast in relation to the vowel anteriority
system; instead, all other sources on Turkic vowels (including all other sources on Kazakh)
presume a tongue-body contrast as the correlate of vowel "backness".

    This study briefly overviews the literature on vowel anteriority and claims of tongue-
root vowel systems in languages of Central Eurasia (section 2), explains the methodology
used for data collection (section 3), presents (section 4) and discusses (section 5) the results,
and draws preliminary conclusions (section 6).

2.   Background

Vowel backness is normally associated articulatorily with the front-back position of the
arched part of the tongue body during vowel production: for front vowels, the highest part
of the tongue body is further forward than for that of back vowels.

    Some languages of the world have two types of anteriority contrast in their vowel
systems: both a backness contrast and a tongue root contrast. In these languages, any given
vowel is either front or back, as well as tongue-root advanced or retracted:[1] each vowel
must match one of the four logical combinations, and vowels of all four types are present.

---

[*]This paper contains color graphics that we could not print. They are available in the online version of the
proceedings volume.
   [1]Such vowel systems appear to vary in terms of whether they contrast tongue-root advanced / neutral,
advanced / retracted, or even neutral / retracted.

This pattern is found most notably among quite a few Western African languages belonging to a range of language families (Ladefoged & Maddieson 1996, ch. 9).

A number of Central Eurasian languages have also been described as exhibiting this sort of system (i.e., one involving both a tongue body contrast and a tongue root contrast), including both Mongolic and Tungusic languages. The Mongolic languages include Western Buriat (Kang & Ko 2012), Buriat (Бураев 1959), Tsongol Buriat (Kang & Ko 2012), Halh (Svantesson 1985, Svantesson et al. 2005), Baarin (Svantesson 1985), and Šiliingol (Svantesson 1985). The Tungusic languages of this type include Solon (Svantesson 1985) and Eastern Ewen (Kang & Ko 2012).

Vajda (1994) describes Kazakh not as a typical tongue root language, as the Mongolic and Tungusic languages are described, but as having a vowel system which contrasts *only* tongue root position, and not tongue body position.

The purpose of this study is to investigate whether the anteriority contrast in Kazakh and two related languages—Kyrgyz (a close relative) and Turkish (a well studied but more distant relative)—can be attributed in any way to tongue root position as claimed by Vajda (1994) for Kazakh, or whether a tongue body analysis is more viable.

The above-mentioned Mongolic and Tungusic languages, as well as Kazakh, comprise a set of Central Eurasian languages for which a tongue root system has been described. The map in (1) displays the areas where these languages are primarliy spoken, as well as the additional languages examined in this study (Kyrgyz and Turkish), and the origins of the six participants in this study.

(1)   A map depicting where each Central Eurasian language is spoken for which a tongue root system has been described, as well as the two additional Turkic languages examined in this study. The hometowns of the six participants in this study are also presented.
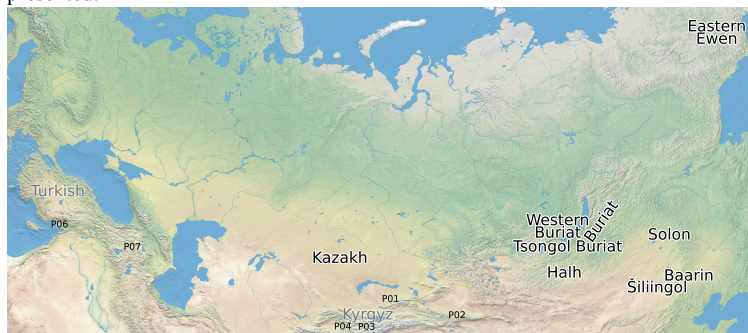
3.   Methodology

Native speakers of Kazakh, Kyrgyz, and Turkish read stimuli embedded in carrier phrases. The stimuli and carrier phrases are presented in §3.1. Audio recordings were made, and the position of the tongue was recorded using ultrasound imaging. How ultrasound imaging of

the tongue works is discussed in §3.2 and how the specific measures for this study were produced is discussed in §3.3.

3.1   Stimuli

The stimuli used in this study were mostly multi-syllabic morphological forms of CVC stems. The vowel of these stems was the target of analysis. The stem consonants were varied within each language so that the short vowel phonemes of each language occurred in a range of contexts. The shapes of the stems for each language (and the total number of stimuli of each shape) and the vowels found in each shape are presented in (2) for Kazakh, (3) for Kyrgyz, and (4) for Turkish.

(2)   Stmiuli shapes for Kazakh: 12 K_sVCCV, 11 t_sVC(CV)

    a.   K_s stems (12): [ə, iə, ʉ, yʉ, æ, ɑ, uʊ, ʊ, ə]

    b.   t_s stems (9): [ə, iə, ʉ, yʉ, ɑ, uʊ, ʊ, ə]

    c.   t_sVC stems (1): [æ]

(3)   Stimlui shapes for Kyrgyz: 8 t_ʃVCCV, 11 K_rVCCV(r), 10 K_lVCCV(r)

    a.   t_ʃ stems (8): [i, y, e, œ, ɑ, ɯ, u]

    b.   K_r stems (11): [i, e, œ, ɑ, o, ɯ, u]

    c.   K_l stems (10): [y, e, œ, ɑ, o, ɯ, u]

(4)   Stimuli shapes for Turkish: 7 b_zCVCV(m), 6 d_SCVCVC, 10 k$^{(j)}$_rCVCVC

    a.   b_z stems (7): [i, y, e, a, o, u]

    b.   d_ʃ stems (6): [i, y, ø, ɯ, u]

    c.   k$^{(j)}$_r stems (10): [i, y, ø, a, o, ɯ, u]

A number of filler stimuli were also recorded, but not analysed. Each target word was recorded once in each of two carrier sentences. The stimulus sentences were randomised and presented in a series of "slides" on a screen; each slide contained 6 sentences (corresponding roughly to the upper limit of the ultrasound system's recording buffer), and there were up to around 150 slides presented, varying some by language.

The carrier phrases positioned the stimuli in similar prosodic environments, and always after a bilabial stop. One carrier sentence each was used containing an anterior vowel and a posterior vowel before the bilabial stop. The sentences were roughly equivalent in the three languages, not only having very similar forms and consisting of cognates, but also having similar semantics.

Examples (5), (6), and (7) present the carrier phrases for Kyrgyz, Kazakh, and Turkish, respectively. Each example presents the anterior- and posterior-vowel carrier phrase, and includes their orthographic forms, their general phonetic realisation, a morphophonological representation, a morpheme-by-morpheme gloss, and a smooth English translation.

(5)   Kyrgyz anterior-vowel and posterior-vowel carrier phrases

    a.   Үйгө кирип, _____ деп айттым.

        yjɣœ ɣiɾíp   _____   deβ ájtt^{h}ɯm

yj–GA kiɾ–(I)p _____ de–(I)p ɑjt–DI–m
house–DAT enter–vADV _____ QUOT–vADV say–RCNTPST_DIR–1SG
'I entered the house and said _____.'

b.   Үйгө барып, _____ деп айттым.
     yjɣœ βɑɾɯ́p _____ deβ ájttʰɯm
     yj–GA bɑɾ–(I)p _____ de–(I)p ɑjt–DI–m
     house–DAT go–vADV _____ QUOT–vADV say–RCNTPST_DIR–1SG.
     'I went home and said _____.'

(6)   Kazakh anterior-vowel and posterior-vowel carrier phrases

a.   Үйге кіріп, _____ деп айттым.
     ʉjɣɪə kʰəɾə́p _____ dɪəβ ájttʰəm
     ʉj–GA kʰəɾ–(I)p _____ dɪə–(I)p ɑjt–DI–m
     house–DAT enter–vADV _____ QUOT–vADV say–RCNTPST_DIR–1SG.
     'I entered the house and said _____.'

b.   Үйге барып, _____ деп айттым.
     ʉjɣɪə bɑɾə́p _____ dɪəβ ájttʰəm
     ʉj–GA bɑɾ–(I)p _____ dɪə–(I)p ɑjt–DI–m
     house–DAT go–vADV _____ QUOT–vADV say–RCNTPST_DIR–1SG.
     'I went home and said _____.'

(7)   Turkish anterior-vowel and posterior-vowel carrier phrases

a.   Eve girip, _____ dedim.
     ɛvɛ giɾíp _____ dedim
     ev–(j)A giɾ–(j)Ip _____ de–DI–m
     house–DAT enter–vADV _____ say–RCNTPST_DIR–1SG.
     'I entered the house and said _____.'

b.   Eve varıp, _____ dedim.
     ɛvɛ vaɾɯ́p _____ dedim
     ev–(j)A vaɾ–(j)Ip _____ de–DI–m
     house–DAT reach–vADV _____ say–RCNTPST_DIR–1SG.
     'I reached the house and said _____.'

3.2   Ultrasound tongue imaging

Ultrasound imaging was used in this study to capture the shape of the tongue during the articulation of vowels. Ultrasound imaging allows for reasonably high-speed recording of the shape of the tongue surface during speech production (Stone 2005). A transducer is held in place under the chin, and standard 2-dimensional ultrasound systems are able to capture an area like that shown in (8).

(8)   An approximation of the area of the vocal tract imaged using ultrasound. Adapted from Cardinal vowel tongue position-front.svg © User:Badseed / CC BY-SA 4.0.
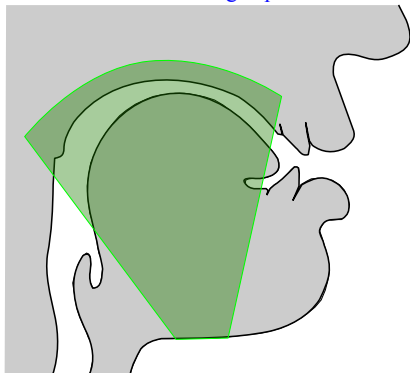
In this study, a Philips EPIQ 7 ultrasound system was used with an X6-1 transducer. An Articulate Instruments Probe Stabilisation Headset (Articulate Instruments Ltd. 2008) was used to stabilise the probe. An image of the author wearing the headset and probe in a similar configuration as participants did is presented in (9).

(9)   The Articulate Instruments Probe Stabilisation Headset and Philips X6-1 ultrasound transducer as worn by the author.

There are several limitations to ultrasound imaging of the tongue. While the area from the tip of the tongue to the tongue root may be captured, the mandible and hyoid bones, respectively, create "shadows" in these areas. Also, because of how ultrasound imaging

works, it is usually not possible to image anything beyond the tongue-air interface; in other words, any parts of the vocal tract above the surface of the tongue are not generally imaged.

A typical output frame from the ultrasound system is presented in (10).

(10)   A raw frame of the ultrasound-imaged tongue. In the left-most of the two subframes, the front of the tongue is to the right, and the back is to the left.
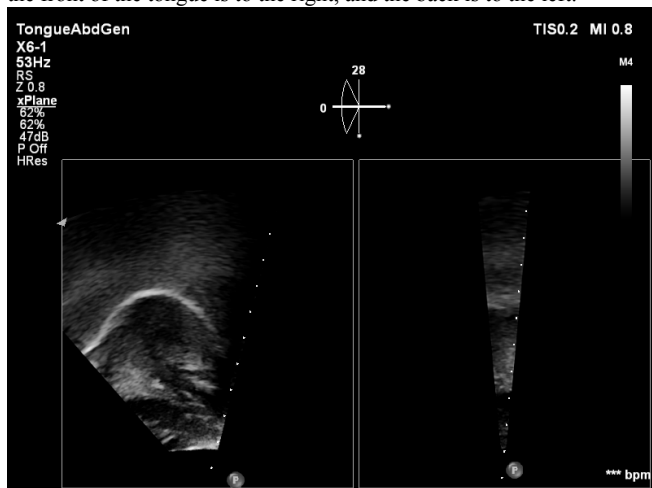
The right side of the midsagittal view of the imaged tongue surface to the left demonstrates the "jaw shadow" caused by the mandible bone. The view to the right is of a plane perpendicular to the midsagittal plane, that was not used for this study.

3.3   Measures

Both acoustic and articulatory data were considered in this study. The data extracted from the acoustic signal included the first formant (F1), the second formant (F2), and the duration (D) of the vowel. For monophthongs, the formants were measured at the half point of the vowel (½), and for diphthongs measurements were taken at the one-third (⅓) and two-thirds (⅔) points. These measures were made using Praat (Boersma 2001), as depicted in (11).

(11)   A Praat TextGrid aligned to the Kazakh vowel [yʉ] in the word [tyʉs], indicating
         the acoustic measures taken in this study.
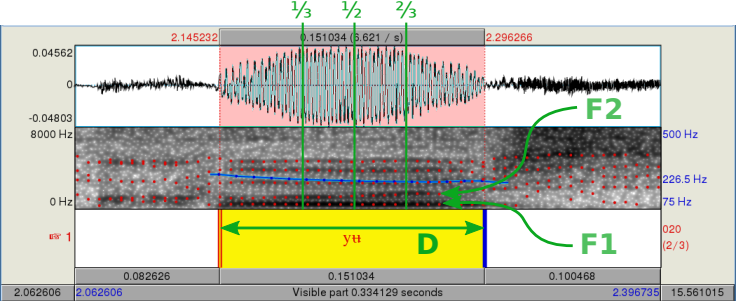
The articulatory data consisted of measurements of the ultrasound images at the time indexes for the midpoint of monophthongs and the one-third and two-thirds points of diphthongs. Measurements were taken using a Python script written by the author. Two points were marked, from which three measurements were extracted. The intersection of a line placed consistently across the bottom of frames with the imaged tongue surface was marked as an approximation of the location of the tongue root, and the distance of the line to that point was measured (TR). In addition, the "highest" point of the tongue (or point furthest from the transducer) was marked, and the vertical distance (TBy) and horizontal distance (TBx) of this point were measured from the edge of the subframe. These measurements are summarised in (12).

(12)   An image depicting the three articulatory measures used in this study: the relative
         backness of the tongue body (TBx), the relative height of the tongue body (TBy),
         and the relative position of the tongue root (TR).
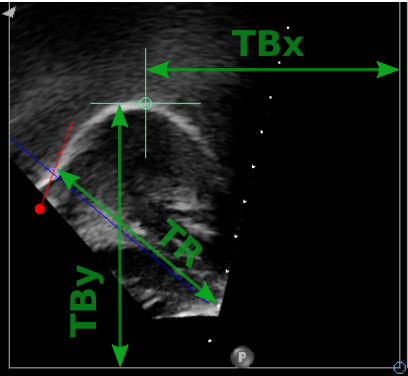
4.   Results

This section presents the results of the various measures for the two speakers of each of the three languages.

4.1   Formant measures

In the graphs of formant measures, anterior vowels are encoded in blue, and posterior vowels are encoded in red. Each vowel token measured is plotted, with diphthong components connected using dark grey lines.

The measurements of the formants for the Kazakh speakers are presented in (13).

(13)   The formant measurements for the Kazakh speakers, P01 and P02.
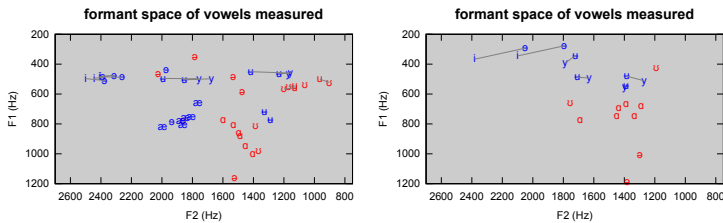
Some of the vowels for the Kazakh speakers exhibit some spurious measurements, due to voiceless articulation and measuring errors. Also, there were not tokens available for all vowel categories in P02's data. These facts, together with the fact that the diphthongs have monophthong-like components, account for some of the overlap seen with many of the vowels. Overall, the posterior vowels appear both lower (i.e., have a higher F1) as well as backer (i.e., have a lower F2) than the anterior vowels.

(14)   The formant measurements for the Kyrgyz speakers, P03 and P04.
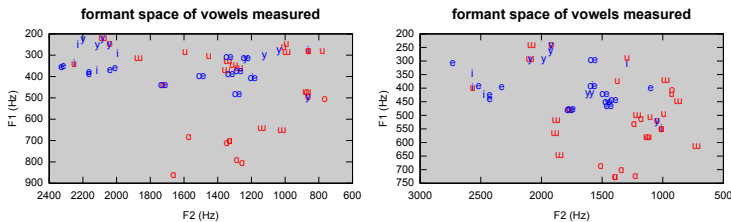
As with the Kazakh vowels, there are a few errors in the measurement of the Kyrgyz vowels. Despite this, it is clear that the posterior vowels in Kyrgyz are lower (have a higher F1) and backer (have a lower F2) than the anterior vowels. Anterior from posterior vowels are not separated by F2 as might be predicted, nor by any combination of F1 and F2.

(15)   The formant measurements for the Turkish speakers, P06 and P07.
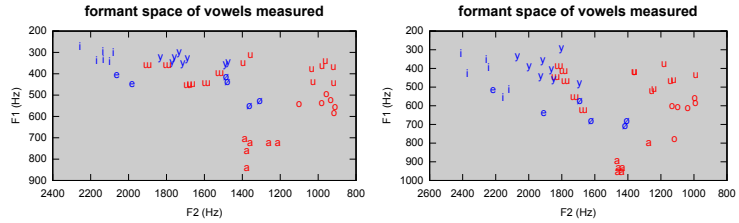
The Turkish vowels have a somewhat more standard distribution than the Kazakh and Kyrgyz vowels. The front rounded and back unrounded vowels overlap to a certain degree in F2 measurements, as might be expected, and [ɑ] appears to be the only low vowel. Otherwise, there is only a small degree of difference in F1 between anterior and posterior vowels, and F2 distinguishes the two sets well on its own.

4.2   Duration measures

In the graphs of duration measures, anterior vowels are encoded in blue, and posterior vowels are encoded in red. Non-high vowels are in the middle of the plots and high vowels are at the edges. Box plots represent first to third quartiles, with a median line in the middle, and whiskers represent the full range of measurements.

(16)   The duration measurements for the Kazakh speakers, P01 and P02.
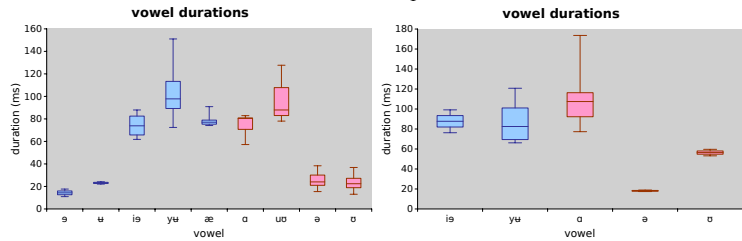
While there were not measurable tokens for all of the vowels in P02's recordings, and high vowels were sparse in P01's recordings as well, it is clear that the traditionally defined "high" vowels ([ɵ ʉ ə ʊ]) are much shorter than non-high vowels in Kazakh.

(17)   The duration measurements for the Kyrgyz speakers, P03 and P04.
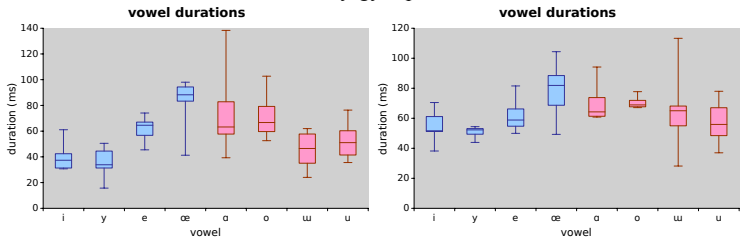
There is more overlap in the duration of high and non-high vowels among the Kyrgyz speakers than among the Kazakh speakers, but even for the Kyrgyz speakers, the tendency for non-high vowels to be longer is present in Kyrgyz as well. There may be one or two measurement errors affecting the length of [ɯ] tokens for P04.

(18)   The duration measurements for the Turkish speakers, P06 and P07.
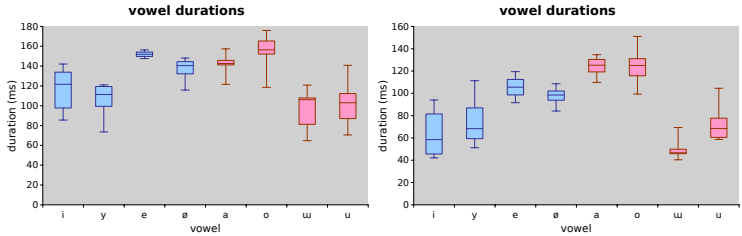
The durations of vowels in the recordings of Turkish speakers were in general longer than those of Kyrgyz and Kazakh speakers, but as in Kyrgyz and Kazakh, the non-high vowels appear to be significantly longer than the high vowels.
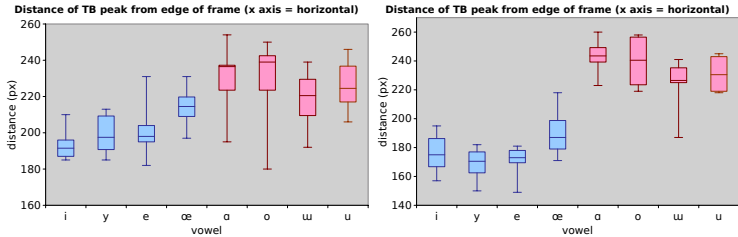
4.3   TR measures

In the graphs of tongue-root distance measures (TR), anterior vowels are encoded in blue, and posterior vowels are encoded in red. Distances measured are the number of pixels to the intersection with the imaged tongue surface of a line overlaid onto the ultrasound frame beginning from an arbitrary point (kept consistent for each speaker) extending roughly parallel to the edge of the graph. Box plots represent first to third quartiles, with a median line in the middle, and whiskers represent the full range of measurements.

(19)   The TR measurements for the Kazakh speakers, P01 and P02.
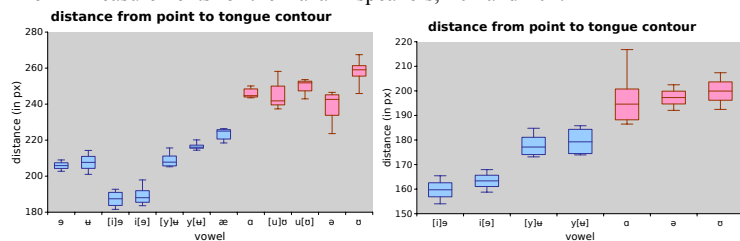
As seen from the plots in (19), the anterior and posterior vowels of Kazakh can be distinguished from one another almost entirely by the TR measure for these speakers. The [æ] vowel, however, appears to be somewhat intermediate between anterior and posterior vowels. This correlates to the vowel's occasionally ambiguous behaviour in the phonology of Kazakh.[2]

(20)   The TR measurements for the Kyrgyz speakers, P03 and P04.
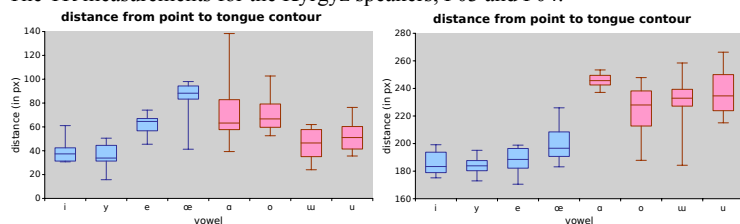
The plots in (20) appear to reflect two different behaviours. P04's data is similar to the Kazakh pattern, in that the TR measure corresponds well to the anteriority contrast of Kyrgyz vowels. For P03, however, this measure appear to correspond more to vowel height than to vowel anteriority.

(21)   The TR measurements for the Turkish speakers, P06 and P07.

---
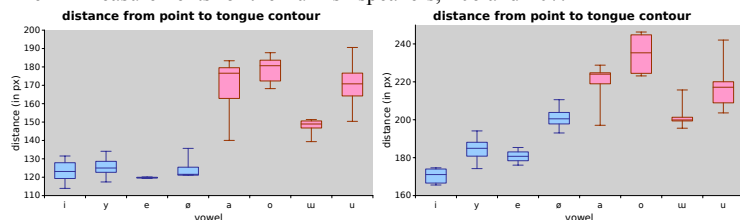
[2]Normally vowels harmonise to /æ/ as if it were a front vowel, e.g. /æn-GA/ [æŋɡiə] 'song-DAT'; however, stems ending in /æ/ behave as if they end in a back vowel, e.g. /kʉnæ-GA/ [kʉnæʁɑ] 'sin-DAT'.

The TR measures for the Turkish speakers, shown in (21), appear slightly different from one another. For P06, the TR measure entirely divides the anteriority classes from one another. For P07, this is true except for [ø], which overlaps with posterior vowel measures. The vowel [ɯ] for both speakers appears to have an intermediate quality.

4.4   TBx measures

In the graphs of tongue-body "backness" measures (TBx), anterior vowels are encoded in blue, and posterior vowels are encoded in red. Distances measured are the number of pixels from the edge of the ultrasound frame to the highest point on the tongue horizontally. Box plots represent first to third quartiles, with a median line in the middle, and whiskers represent the full range of measurements.

(22)	The TBx measurements for the Kazakh speakers, P01 and P02.
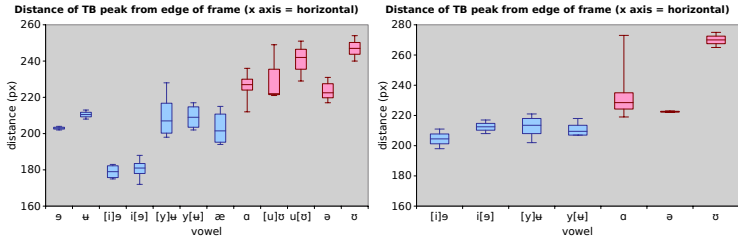
The TBx measures shown in (22) appear to mostly divide the anterior and posterior vowels for Kazakh speakers. However, the ability of this measure to separate the two anteriority classes is not as strong as that of the TR measure.

(23)	The TBx measurements for the Kyrgyz speakers, P03 and P04.

For the Kyrgyz speakers, P03 and P04, the TBx measures shown in (23) do a better job at dividing the anteriority classes than the TR measures in (20).

(24) The TBx measurements for the Turkish speakers, P06 and P07.
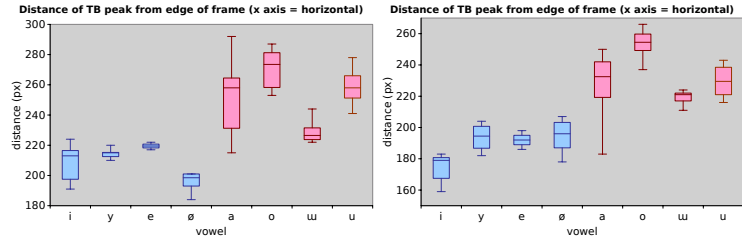
The TBx measures for Turkish speakers in (24) show a good separation of anteriority classes. For P06, it is not as good as the TR measure, but for P07 it is somewhat better.

4.5 TBy measures

In the graphs of tongue-body "height" measures (TBy), anterior vowels are encoded in blue, and posterior vowels are encoded in red. Distances measured are the number of pixels from the edge of the ultrasound frame to the highest point on the tongue vertically. Box plots represent first to third quartiles, with a median line in the middle, and whiskers represent the full range of measurements.

(25) The TBy measurements for the Kazakh speakers, P01 and P02.
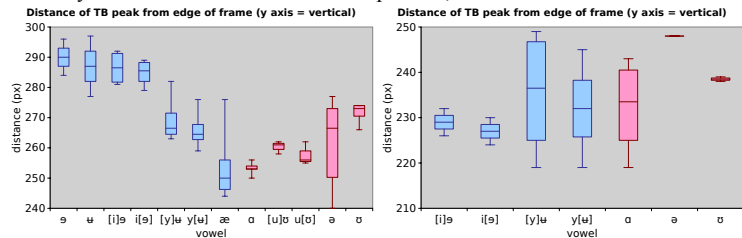
The Kazakh TBy measures in (25) show more or less a correspondence to height (as would be expected) for both speakers. There were very few "high" vowel tokens measured for P02. For P01, the unrounded diphthongs pattern with high vowels and the rounded diphthongs appear to be more like mid vowels. For P02, the rounded diphthong appears to have a wider range and a much "higher" extent than the unrounded diphthong. It appears that all three diphthongs are falling diphthongs for both speakers.

(26)   The TBy measurements for the Kyrgyz speakers, P03 and P04.
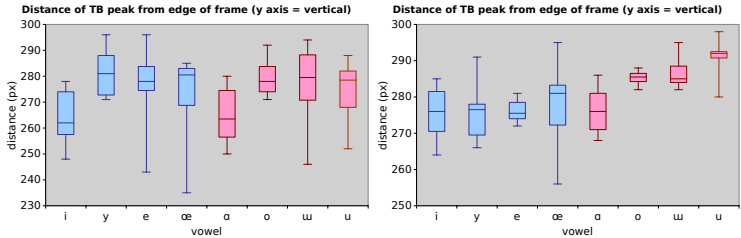
No consistent correlation with height (or backness) is clear in the TBy measurements for Kyrgyz speakers (26).

(27)   The TBy measurements for the Turkish speakers, P06 and P07.
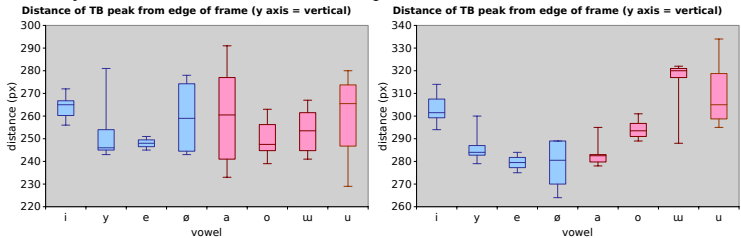
While [ɑ], [u], and [ø] have a wider range than might be expected in the TBy measures for P06, the TBy measurements of P07's vowels show a nice correlation with vowel height.

4.6   TBx by TBy

The TBx by TBy graphs plot the highest point of the tongue during the production of each vowel in a two-dimensional space. Anterior vowels are shown in blue and posterior vowels are shown in red. A grey line links the first and second data points for diphthongs.

(28)   The TBx by TBy plots for the Kazakh speakers, P01 and P02.
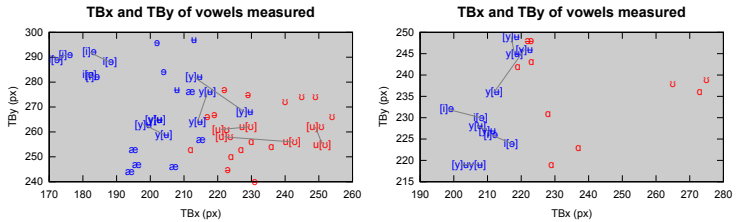
While not as neat or well organised as a vowel formant plot might be expected to be, the TBx by TBy plots for Kazakh speakers in (28) are quite reminiscent of vowel formant plots. The TBx measure (as seen before) separates the vowels by anteriority.

(29)   The TBx by TBy plots for the Kyrgyz speakers, P03 and P04.
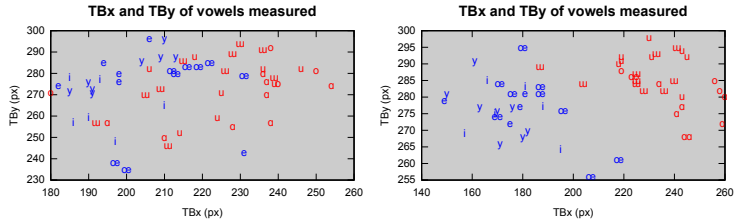
Like for Kazakh, the Kyrgyz TBx by TBy plots in (29) are reminiscent of vowel formant plots, with the TBx dimension dividing vowels by anteriority. The data for P03, however, is quite messy, and may reflect incorrectly extracted ultrasound frames.[3]

(30)   The TBx by TBy plots for the Turkish speakers, P06 and P07.
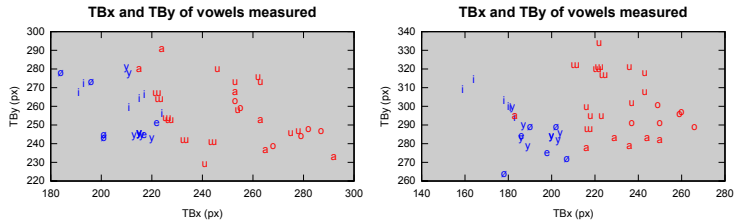
The TBx by TBy plots for Turkish (30), like for those of Kazakh and Kyrgyz, are reminiscent of vowel formant plots. There appears to be an incorrectly measured [a] vowel for speaker P07, but otherwise TBx divides the anteriority classes for P07 well. For P06, a combination of TBx and TBy can be used to divide the anteriority classes.

4.7   TR by TBx

The plots in this section show the correlation between TR and TBx, the two anteriority measures in this study. Anterior vowels are shown in blue and posterior vowels are shown in red. A grey line links the first and second data points for diphthongs. A best fit line is presented in each plot in purple.

As seen in the plots for Kazakh (31), Kyrgyz (32), and Turkish (33), there is a reasonable level of correlation between the TR measures and the TBx measures.

---

[3]Ultrasound frame extraction was automated, though the particular algorithm appears to have had trouble with some of the Kyrgyz data. This will be investigated in future work.
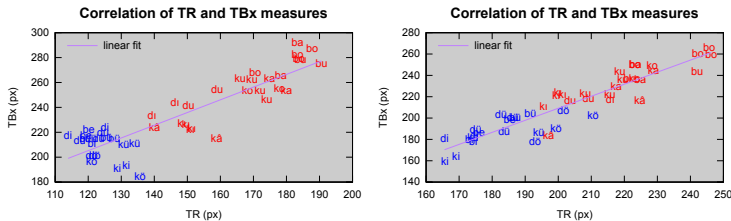
(31)   The TRTBx measurements for the Kazakh speakers, P01 and P02.
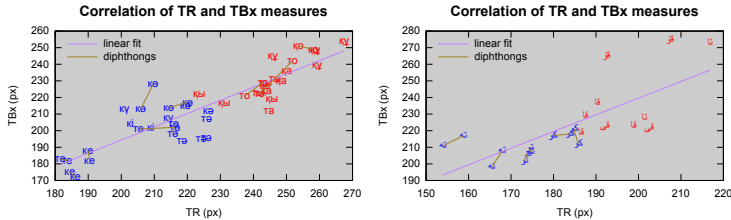
(32)   The TRTBx measurements for the Kyrgyz speakers, P03 and P04.
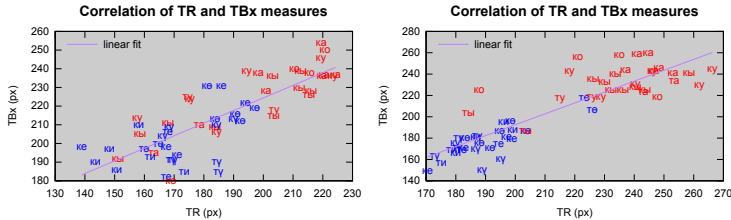
(33)   The TRTBx measurements for the Turkish speakers, P06 and P07.

The correlation is most "clean" in the Turkish data. The overlap of measurements in both dimensions for P03 may be due to measurement errors or improper processing of the data.

5.   Discussion of Results

While not the primary target of this study, some interesting generalisations appear in the duration measures.

    For the Kazakh speakers, the "high" vowels, [ə, ʉ, ə, ʊ], are very short, and are often devoiced or deleted in the first syllable of words. Hence very few tokens were recorded or able to be measured, and those that were measured appear to have suffered from errors in formant analysis. This makes the results of this study not fully interpretable for Kazakh.

For the Kyrgyz and Turkish speakers, the high vowels also appear to be—overall—shorter than the low vowels. However, there are apparent errors in the measurements made for some of the data, especially concerning P03's data.

In general, the formant measurements are very messy, and a simple F2 value is not able to separate anterior and posterior vowels from one another in any of the languages. Stated differently, anterior and posterior vowels have overlapping values of F2 for the phonologically short vowels as uttered by Kazakh, Kyrgyz, and Turkish speakers in the specific environments examined by this study.

The TR measure examined in this study (roughly, tongue root position), though, does appear to be a good predictor of vowel anteriority in the languages studied, especially for Kazakh speakers. It is not as strong a predictor for Turkish speakers, and is only predictive for one Kyrgyz participant—the data for the other Kyrgyz speaker (P03) appears to exhibit some measurement errors.

The TBx measure (roughly, tongue body backness) appears to be a robust predictor of vowel anteriority in all three languages. For the Turkish speakers, TBx is about as good a predictor of vowel anteriority as TR; for the Kyrgyz speakers, it is a stronger predictor; and for the Kazakh speakers, it is less strong a predictor.

These results for TR and TBx measurements taken together show that Kazakh vowel anteriority is best associated with tongue root position, Kyrgyz vowel anteriority is best associated with tongue body position, and Turkish vowel anteriority is associated equally well with the position of both. However, tongue root and tongue body position are well correlated in all three languages. This may simply simply reflect the fact that the position of the tongue root and tongue body are intrinsically linked, but it also may show that the results are not yet fully interpretable.

6.   Preliminary Conclusions

The preliminary conclusions of this study are that Kazakh speakers differentiate anterior and posterior vowels primarily through the position of the tongue root, Kyrgyz speakers differentiate the two vowel classes primarily through the position of the tongue body, and Turkish speakers differentiate them using both equally.

Future work will investigate these preliminary conclusions in more depth. Besides examining more speakers of each language, investigating a wider range of languages (related and not), and reevaluating questionable measurements, more robust measures will be used. Among new, more powerful measures, the entire shape of the tongue should be considered as a whole. This will require fully traced tongue contours.

References

Articulate Instruments Ltd. 2008. *Ultrasound stabilisation headset users manual: Revision 1.4.* Articulate Instruments Ltd.

Boersma, Paul. 2001. Praat, a system for doing phonetics by computer. *Glot International* 5:341–345.

Kang, Hijo, & Seongyeon Ko. 2012. In search of the acoustic correlates of tongue root contrast in three Altaic languages: Western Buriat, Tsongol Buriat, and Ewen. *Altai Hakpo* 22:179–203.

Ladefoged, Peter, & Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford: Blackwell.

Stone, Maureen. 2005. A guide to analysing tongue motion from ultrasound images. *Clinical Linguistics & Phonetics* 19:455–501.

Svantesson, Jan-Olof. 1985. Vowel harmony shift in mongolian. *Lingua* 67:283–327.

Svantesson, Jan-Olof, Anna Tsendina, Anastasia Karlsson, & Vivan Franzén. 2005. *The Phonology of Mongolian*. The Phonology of the World's Languages. New York: Oxford.

Vajda, Edward J. 1994. Kazakh phonology. In *Studies on East Asia*, ed. Edward H. Kaplan & Donald W. Whisenhunt, volume 19: "Opuscula Altaica: Essays Presented in Honor of Henry Schwarz", 603–650. Western Washington University.

Бураев, Игнатий Дмитриевич. 1959. *Звуковой состав Бурятского языка*. Улан-Удэ: Академия Наук СССР, Сибирское Отделение.

Jonathan North Washington
jonwashi@indiana.edu

Aspect and Evidentiality in Azeri[*]

Gita Zareikar

University of Ottawa

1. Introduction

Azeri is a Turkic language spoken in north-west Iran. In Azeri the verbal morpheme *-miş*
has an aspectual or evidential reading in certain contexts. In this paper I present the double
reading of *miş* and whether this overlap is due to polysemy or to morphological haplology.
In order to answer these questions I suggest two hypotheses; first: there is only one mor-
pheme and the evidential reading is just another reading of the perfect, and hence it is an
epistemic modal (Izvorski 1997). This hypothesis predicts that it is the syntactic position
of the morpheme that yields a distinct reading in each case. Second: there are two distinct
morphemes, one that marks aspect and the other that marks evidentiality. In conclusion I
claim that the morphological facts support the second hypothsis.

2. *Miş* morpheme as an aspect marker

The aspect sense of *miş* is seen in (1) and (2). (1) and (2) are perfects of result or present
relevance. In both cases the result state of the telic event, namely "my being home" or
"juice being available" hold at the time of utterance.

(1)   Gəl-**miş**-əm.   (2)   Şərbət tut-**muş**-am.
      come-PARTP-1SG   drink   make-PARTP-1SG
      'I have come home.'   'I have made some juice'
      (And I am home now.)   (And it is still there.)

    Iatridou et al. (2001) propose that perfect locates an eventuality relative to some refer-
ence point. In all the examples above the reference time overlaps with the utterance time
and hence the eventuality of the event extends to the utterance time. McCoard (1978) dis-

---

tinguishes four types of perfect. These four types are introduced as universal perfect, experiential perfect, perfect of results and perfect of recent past.

In the universal perfect the underlying eventuality continues at the utterance time and thus the universal perfect behaves in part like present (Brugger 1978).

The examples in (1) and (2) above are suggestive of the universal perfect in Azeri. Not all four types of perfect are found in all languages. Perfect in Azeri seems to align with some types suggested by McCoard (1978). Since the language does not distinguish distance in past, the perfect of recent past will be developed by means of an adverbial, as in (3)[1].

(3)   **Təzə**   gəl-**ip**.
       shortly come-PARTP.3SG
       'He has come a short while ago.'

The combination of *-miş* and past tense marker, i.e. *-di* yields an anteriority reading, as in (5). The aorist reading is derived by *-di* in (4) and the presence of *miş* before the tense marker gives rise to pluperfect reading of the sentence.

(4)   Al-**dı**-m.   (5)   Əri   öl-**müş-dü**.
       buy-PAST-1SG   husband die-PARTP-PAST.3SG
       'I bought.'   'Her husband had died.'

3.   Miş morpheme as an evidential marker

There are cases in which *-miş* may be used in a structure that derives an evidential reading, more specifically the one indicating that eventuality is indirect. In (6) similar in form to (1), an evidential meaning exists in the statement that the father arrived is marked as being deduced from available evidence rather than directly observed.

(6)   Baba-n   gəl-**ip**.
       father-GEN come-PARTP
       'Apparently, your dad has come home.' (The lights are on.)

Furthermore, a similar structure could be used to talk about an event that the speaker heard about and has no direct evidence for.

This is the hearsay interpretation of *miş* as in (7). In such case the speaker does not commit themselves to the truth of the statement, in (6) and (7).

(7)   Saray adam-ı   gör-**müş**.
       Saray person-ACC see-PARTP
       'Reportedly, Saray has seen the person.'

---

[1]The reader is invited to assume that *-ip* and *-miş* are allomorphs in Azeri. *Miş* has an extensive use and it is taking over the use of *-ip* in first and second person. The process of this shift is beyond the scope of this paper and is subject to a future study.

In the above example, *-miş* is attached to the root. In this case *miş* refers to some time prior to the reference time but in addition this form indicates that the source of the speaker's evidence is not direct. The event time precedes the reference time, E<R. This reading is distinct from the aspectual reading discussed in the previous section. This happens because of the evidential component; however, it is not always aspectually different from the non-evidential readings observed in (1) and (2).

The examples below show that *miş* in the present perfect form could yield ambiguous readings. (8) gives only a perfect aspect, (9) recognizes an evidential reading and the verb maintains its perfect aspect.

(8) Uç   saat-dir gəl-**miş**-əm.
     three hour-be come-PARTP-1SG
     'It has been three hours since I have come.'

(9) Çörəh-lər-ı   Aida yi-y-**ip**.
     bread-PL-ACC Aida eat-EP-PARTP
     'Reportedly, Aida has eaten the bread.'

It is undeniable that the perfect of evidentiality morphology cannot entirely preclude the temporal reading of the perfect. The perfect meaning can still be present on evidentials, as in (10).

(10)   Sənəm ömr-ü-ndə   iki dəfə gızıl tap-**ip**.
       Sanam life-POSS-DAT two times gold find-PARTP.3SG

       a.   'Sanam has discovered gold twice in her life.'
            (I have witnessed that/present perfect.)
       b.   'Reportedly, Sanam has discovered gold twice in her life.'
            (I have an external source of evidence/present perfect.)

4.   Correlation of *miş* with other aspectual morphemes

Examples (11) and (12) are derived in a context where the speaker is telling a story that she has heard from someone else in the past. The reference time in this conversation is set in the past. The forms used in these examples contain two occurrences of *miş*. The resulted reading is the one with a reportative component. In both cases the speaker claims that she has received indirect information about the event.

(11)   Əri   öl-**müş-ü-müş**.
       husband die-PARTP-PAST-PARTP.3SG
       'Reportedly, her husband had been dead.'

(12)   Saray xalan ol-**muş-u-muş**.
       Sara aunt born-PARTP-PAST-PARTP.3SG
       'Reportedly, aunt Sara had been born.'

Miş with future presents its own peculiarities. *Miş* follows future *-acak* rather than preceding it, as in (13). This shows that *-acak* is a prospective marker. In the following example the speaker expresses his lack of awareness about doing the writing event. This event was expected to be completed in the past but never got completed because the agent was not aware that he had to do that.

(13)   Yaz-acağ-ı-**mış**-am.
        write-FUT-EP-PARTP-1SG
        'Apparently, I had to write it.'

The next section discusses the occurrence of the *miş* morpheme as an aspect and evidential marker in various contexts and provides an explanation for the double occurrence of this morpheme, as shown in (11) and (12) above.

5.   Discussion and analysis

Various readings of *-miş* in Turkish have been examined in much previous work. To begin with, we will examine the analysis proposed by Izvorski (1997). Izvorski claims that there are languages with perfect of evidentiality in which the perfect morphology serves to indicate report or inference. Hence, it is common for languages to distinguish between direct and indirect evidence and it is normally the indirect evidentiality that is morphologically marked. Sentences without the perfect of evidentiality morphology should be used for direct evidence and cannot have the perfect of evidentiality reading either.

Izvorski argues that "the morphology of the present perfect or a form historically derived from the present perfect, expresses a particular evidential category, one that indicates the availability of indirect evidence for the truth of a proposition". In Izvorski's account, evidentiality is marked on perfects and an evidential reading is not available in forms other than perfect. In (14), the suffix *miş* is the present perfect inflection and it has an indirect evidential interpretation.

(14)   Gel-**miş**-im.   Turkish
        come-PERF-1SG
        'I have come.' (present perfect) and/or 'I apparently came.'   (Izvorski 1997, 222)

To support the idea that the ambiguity between the present perfect and an evidential reading is not accidental, Izvorski gives examples in which *miş* does not give rise to an evidential reading when it forms NP-modifying participles, as in (15). In other words, it is only the perfect form of the morpheme that yields an evidential reading.

(15)   öl-**müş** adam   Turkish
        dead   man
        'dead man' ≠ 'a man who is apparently dead'   (Izvorski 1997, 223)

Additionally, when *miş* surfaces with the past or future perfect or even when it occurs in non-finite clauses, the evidential interpretation is absent, as illustrated in (16) and (17)

(16)   **Gel-miş**-tim.   Turkish
       come-PERF-1SG.PAST
       'I had come.'≠ 'I apparently came/had come.'

(17)   **Gel-miş**   ol-acak-ım.   Turkish
       come-PERF be-FUT-1SG
       'I will have come.'≠ 'I will apparently come.'   (Izvorski 1997, 223)

     Examples in (18) from Azeri illustrate the *-miş* morpheme giving rise to both perfect of result readings (18a), (18c) and evidential readings (18b), (18d).

(18)   a.   Bir ev   al-**mış**-am.
              a   house buy-PARTP-1SG
              'I have bought a house.'
              (I am telling a friend about the house I have bought.)

       b.   Gör-dü-m   bir ev   al-**mış**-am.
              see-PAST-1SG a   house buy-PARTP-1SG
              'I had a dream that I have bought a house.'
              (I am talking about my dream in which I was seeing myself buy a house.)

       c.   Bir ev   al-**ıp**.
              a   house buy-PARTP-3SG
              'S/He has bought a house.'
              (I know that and I was there.)

       d.   Sanam Paris-i   gəz-**ip**.
              Sanam Paris-ACC visit-PARTP.3SG
              'Reportedly, Sanam has visited Paris.'
              (I have heard about it.)

Examples above provide evidence for an ambiguity between present perfect and indirect evidentials. On the other hand, it is undeniable that an evidential interpretation of *miş* entirely preclude the temporal-aspectual reading of the perfect for the indirect evidence with a perfect aspect, in (18d). Accordingly, it is hard to strongly argue for the present perfect morpheme to be only the perfect of evidentiality morphology since evidentiality seems to be the secondary reading of the perfect, as illustrated in (18a) and (18c) above. In Azeri it is impossible to argue that the perfect of evidentiality morphology does not play an aspectual role and it is the source of modality instead. This has been studied for other languages as well. According to Matthewson et al. (2007) the elements which encode information source may or may not fall into the category of epistemic modals, contrary to Izvorski (1997). In (19), the sentence preserves its temporal reading and the present perfect morpheme is

clearly marking an existential perfect interpretation in addition to being an epistemic modal in Izvorski's terms. The expected aorist reading (perfective) is not available in this case.

(19) Sənəm ömr-ü-ndə   iki dəfə gızıl tap-**ıp**.
      Sanam life-POSS-DAT two times gold find-PARTP.3SG

    a.   'Sanam has discovered gold twice in her life.'
        (I have witnessed that/present perfect/direct.)
    b.   'Reportedly, Sanam has discovered gold twice in her life.'
        (I have an external source of evidence/present perfect/indirect.)

Another difference between Azeri and Izvorski's Turkish data is that Azeri reveals an evidential reading in the future. In (20) (future indirect) the speaker is giving a statement without making any commitment to the truth of the information. (20) instantiates a morphological possibility that is absent in Turkish, in (17) (Izvorski 1997).

(20) a.   Bügün gəl-**əcək-i-miş**-əm.
        today   come-FUT-PAST-PARTP-1SG
        'Apparently, I should have come today'

    b.   *Bügün gəl-**miş-əcək**-əm.
        today   comePARTP-FUT-1SG
        '(Intended:) Apparently, I should have come today.'

It is notable that in the example above *miş* is not able to precede the future morpheme, in (20b). The fact that the future morpheme cannot follow *miş* is an evidence that *acak* and *miş* are order sensitive. Further morphological evidence supports the idea that *acak* is occupying the aspect position, rather than tense in the language, as in (21).

(21) *Gəl-**ir**-əcək.
    come-PROG-FUT
    '(Intended:) S/he will be coming.'

The outcome of the discussion above reveals that in positions where *acak* precedes *miş*, the latter must scope over *acak* and hence cannot precede it. This leaves only one place for *miş* and that is somewhere higher than Aspect. The semantics of the example in (20a) suggests that *miş* is giving an evidential interpretation to the structure. This observation yields that *miş* with an evidential interpretation can occur in non-perfect structures, contrary to what is claimed by Izvorski for Turkish.

Similar cooccurrence facts are provided by the progressive and habitual aspect morphemes. These morphemes, similar to future, precede *miş*, as in (22).

(22) a.   Yaz-ır-**mış**.   (*Yaz-**mış**-ır.)
        write-DUR-PARTP.3SG
        'Reportedly, he was writing.'

b.   Yaz-ar-**mış**.   (*Yaz-**mış**-ar.)
     write-HAB-PARTP.3SG
     'Reportedly, he used to write.'

The facts are summarized as follows: in conjunction with other aspectual morphemes *miş* has an exclusively evidential interpretation; the morphological reflection of this is the inability of *miş* to occur inside the other aspectual morphemes.

So far the morphology and semantics of the above examples lead us to accept that there are two distinct slots for the *miş* morpheme, one for the aspect and the other for evidentiality but does not explain the existance of two distinct morphemes. In providing the examples I did not separate the occurrence of *miş* and *-ip* and assumed that they should alternate equally in all the examples, but a closer look reveals that the alternation of *-miş* and *-ip* is not possible when *-miş* follows an aspect marker, as in (23). This is extendable for all the other aspect morphemes. However, the question remains unanswered why in structures such as (23) *-miş* derives a grammatical indirect reading, (22), but *-ip* is unable to attain such a reading. The answer for the alternation of *-miş* and *-ip* is a topic for future study.

(23)   a.   *Yaz-ır-**ıp**.
            write-PROG-PARTP.3SG
            (Intended:) 'Reportedly, he has been writing.'

       b.   *Yaz-acak-**ıp**.
            write-FUT-PARTP.3SG
            (Intended:) 'Reportedly, he was going to write.'

Now that we have seen various correlations with *miş* of aspectual morhemes, we are ready to include the evidential plulerfect. The impossibility of the alternation of *-miş* and *-ip* becomes clearer by looking at the pluperfect. In addition to what was shown above we need to be able to explain the double occurrence of the morpheme in the evidential form of the pluperfect, as illustrated in (24).

(24)   O   zaman gəl-**miş**-i-**miş**-lər.
       that time   come-PARTP-PAST-PARTP.3PL
       'Reportedly, they had come at that time.'

Examples like (24) is a motivation to defend the second hypothesis which proposes there are two distinct *-miş* morphemes in the language. Hence, the morphological and semantic evidence is in favour of the existance of two distinct morphemes but we still need to explain how to distinguish the occurrence of two separate morphemes in the present perfect and that the second *-miş* is evidential.

Further independent evidence to show that the present perfect morphology is not the same as evidential morpheme in all cases comes from the examples below. In (25a) *-di* in combination with the *-ir* morpheme refers to an ongoing event in the past. It is expected for the progressive morpheme *-ir* to mark aspect and be closer to the stem than the tense

morpheme *-di*. In the case of (25a) it is not possible to insert an additional *-miş* into the structure; the result will be ungrammatical, as in (25b). Considering the latter example it is clear that the position of the aspect is already filled by the progressive morphology, hence, there is no position for a second aspect marker, admitting that *miş* and the other aspect morphemes are in complementary distribution. This is an indication that the *-miş* morpheme when preceding *-di* marks aspect.

Now looking at the next example from this set, in (25c) the combination of the imperfective aspect and *-miş* should exemplify a similar instance as we observed in (25b). The presence of the imperfective aspect marker is expected to immediately reject the occurrence of *-miş* as an aspect marker here.[2] Though the two morphemes look morphologically the same, they seem to occupy two different syntactic positions. Furthermore their semantic interpretation is also distinct. While (25a) expresses a direct evidence, (25c) is about a reported event. The conclusion drawn here is that the *-miş* morpheme in (25c) is not the same morpheme as the one in the present perfect that gives an aspectual reading. This is another support for the claim that there are two types of *-miş* in the language.

(25)   a.   Baba-m   işl-ir-**di**.
              father-GEN work-PROG-PAST.3SG
              'My dad was working.'

        b.   *Baba-m   işl-ir-**miş**-**di**.
              father-GEN work-PROG-PARTP-PAST.3SG
              '(Intended:) My dad has been working.'

        c.   Baba-m   işl-ir-**miş**.
              father-GEN work-PROG-PARTP.3SG
              (i)   'Apparently/reportedly, my dad is/was working.'
              (ii)   'My dad is/was working.'

The table below summarizes the co-occurrence of the morphemes discussed in this paper. It considers allomorphs for tense and evidentiality. The examples in (27a) illustrate the occurrences presented in the table.

(26)   *Table 1. morphological co-occurrence*

ASP	TENSE	EV
unmarked→ ø	-di	-miş/-ip→IND
perfect→ -miş/-ip	-i	ø→DIR
future→ -acak	ø	
progressive→-ir		
habitual→ -ər		

---

[2]One might be tempted to consider the existence of double aspect markers but this is rejected as soon as the tense morpheme *-di* is inserted in (25b) which should not if there were two aspect markers.

(27) a. O   zaman gǝl-**miş**-i-**miş**-lǝr.   [IND→miş-i-miş]
          that time   come-PARTP-PAST-PARTP.3PL
          'Reportedly, they had come at that time.'

      b. Uç   saat-dir gǝl-**miş**-ǝm.   [DIR→miş-ø-ø]
          three hour-be come-PARTP.1SG
          'It has been three hours since I have arrived.'

      c. Baba-n   gǝl-**ip**.   [IND→ø-ø-ip]
          father-GEN come-PARTP.3SG
          'Apparently, your dad has arrived.' (the lights are on)

When there is no *-di* the co-occurrence of two *miş* is possible in an environment of *-i*. In non-past cases the [V-miş-ø- ø] means present perfect [DIR] as the aspect is marked but in unmarked aspect cases [V-ø-ø-miş] there is an [IND] reading. So unmarked aspect + *-di* would mean aorist and unmarked aspect + ø + *miş* means present perfect evidential. The co-occurrence of the *-miş* is also explained in this model as the first one marks aspect and the second one conveys evidentiality. This table is able to account for the co-occurrence of tense, aspect and evidentiality morohemes but nonetheless has some shortcomings. There are certain occurrences that this model cannot account for. For example it will predict the occurrence of [*acak-di-miş*]. We also need to explain the cases in which the morpheme *-i* is selected rather than *-di* as in [*miş*-i-miş] and the impossiblity of having [*miş*-miş] occurrences. The explanation of the latter is possible under the account of haplology. Further discussion on the haplology is a subject for future study. Haplology is not observable in simple past since the simple past takes the direct evidential marker above Tense which is null. Haplology happens in cases where the evidentiality morpheme is present.

6.   Conclusion

The morphological facts indicate that the *miş* morpheme precedes the tense marker wherever it marks aspect. This is an indication of syntactic variation and is an evidence that supports the existence of the evidential morpheme above Tense in the derivation in the same way that Tense scopes over Aspect. Following the discussion above we can conclude that Izvorski's argument about the perfect of evidential morphology and its marking the reports and inferences instead of playing a temporal-aspectual role is not explanatory enough for all the cases in Azeri. Further morphological evidence from imperfective and future also supported this claim.

References

Aikhenvald, Alexandra Y. 2004. *Evidentiality*. Oxford University Press.
Brugger, Gerhard. 1978. Eventive time properties. In *Proceedings of the 21st Annual Penn Linguistics Colloquium*, 51–63.

Comrie, Bernard. 1976. *Aspect: An introduction to the study of verbal aspect and related problems*, volume 2. Cambridge university press.

Comrie, Bernard. 1985. *Tense*, volume 17. Cambridge University Press.

Guillaume, G. 1929. *Temps et verbe*. Champion.

Hankamer, Jorge, J. 2012. The exponence of Turkish aorist. *University of California* .

Iatridou, Sabine, Elena Anagnostopoulou, & Roumyana Izvorski. 2001. Observations about the form and meaning of the perfect. *Current Studies in Linguistics Series* 36:189–238.

Izvorski, Roumyana. 1997. The present perfect as an epistemic modal. *Semantics and linguistic Theory* 7:222–239.

Klein, W. 1994. *Time in language*. Routledge.

Komfilt, Iaklin. 1996. On copular clitic forms in Turkish. *ZAS Papers in Linguis-tics* 6:96–114.

Kratzer, Angelika. 1998. More structural analogies between pronouns and tenses. In *Proceedings of SALT*, volume 8, 92–110.

Matthewson, Lisa, Henry Davis, & Hotze Rullmann. 2007. Evidentials as epistemic modals evidence from st'át'imcets. *Linguistic variation yearbook* 7:201–254.

McCoard, Robert W. 1978. *The English perfect: tense-choice and pragmatic inferences*. North Holland Publishing Company.

Simeonova, Vesela, & Gita Zareikar. 2015. The syntax of evidentials in azeri, bulgarian and persian. In *Proceedings of Annual Conference of Canaidan Linguistic Association*.

Slobin, I. Dan, & A. Ayhan Aksu. 1982. Tense, aspect and modality in the use of the Turkish evidential. *Typological studies in Language* 1:185–200.

Gita Zareikar
g.zareikar@uottawa.ca