

## Mismatched suppletion in Azeri as morphology/phonology competition

Colin Davis\*

**Abstract.** This paper focuses on a puzzle about suppletion of the perfect/evidential suffix in north Azeri/Azerbaijani. This morpheme’s default form is *-mİf*, but it can also be realized as *-(j)İb* when the subject is 2nd or 3rd person, with one exception: in certain contexts where this suffix is used twice, the first instance is realized as *-(j)İb*, even if the subject is not 2nd/3rd person. I hypothesize that this over-application of suppletion is driven by a requirement to avoid strings with adjacent identical morphemes. If this analysis is correct, this phonologically-forced morphological mismatch provides evidence that lexical insertion interacts and competes with phonological constraints (Wolf 2008, 2009, Pertsova 2015, a.o.). I go on to consider some remaining puzzles about the distribution of *-(j)İb*.

**Keywords.** morphology, suppletion, morphology-phonology interaction, Turkic

**1. Introduction.** In this paper, I use original fieldwork data to investigate a puzzle about suppletion in the northern dialect of Azeri/Azerbaijani (Turkic). In particular, I examine suppletion of the perfect/evidential suffix, whose default form is *-mİf*. This research has found, in agreement with Şirəliyev & Sevortjan (1971) and Öztopçu (2003), that this morpheme has an optional allomorph *-(j)İb* that is typically only used with 2nd and 3rd person subjects.<sup>1</sup> Since there is no predictable phonological process that is capable of obviously deriving *-(j)İb* from *-mİf*, I take this to be an instance of morpho-syntactically conditioned suppletion:

### (1) Typical distribution of *-(j)İb* in 2nd/3rd person<sup>2</sup>

- a. biz d̪ɜ̌ɜl-miʃ/\*ib-ij  
 1PL come-PRF-1PL  
 ‘We have come’
- b. mən d̪ɜ̌ɜl-miʃ/\*ib-əm  
 1SG come-PRF-1SG  
 ‘I have come’
- c. siz d̪ɜ̌ɜl-ip-siniz  
 2PL come-PRF-2PL  
 ‘You (pl) have come’
- d. o gatʃ-miʃ/ib-ø  
 3SG run-PRF-3SG  
 ‘He/she/it has run’

\*All Azeri data reported here was gathered in 2013-2014 at the University of Minnesota, with the help of native speaker Samir Karimov from Baku, Azerbaijan. Thanks to comments from David Pesetsky, Norvin Richards, and Matthew Zalansky, as well as audiences at MIT and the Tu+4 workshop at NYU. In the end, this work raises more questions than it answers, but I hope this short paper clarifies some directions for future work. Author: Colin Davis, Massachusetts Institute of Technology ([colind@mit.edu](mailto:colind@mit.edu)).

<sup>1</sup>Though Zalansky (2019) argues that *-(j)İb* is retreating towards marking only 3rd person.

<sup>2</sup>Since the data presented here was originally gathered in a fieldwork context, it is transliterated in a modified version of IPA rather than the original Azerbaijani orthography. The modifications to IPA used here are as follows: [ä] = low front vowel, [a] = low back vowel, [ö] = mid front rounded vowel, [i] = back unrounded vowel, [r] = rhotic tap. Glossing conventions: ACC = accusative, COP = copula, DAT = dative, EVI = evidential, GEN = genitive, IMPV = imperfective, PL = plural, PRES = present, PRF = perfect, PROS = prospective, PST = past, SG = singular.

- e. pfi ät je-j**ib**-di-ø  
 cat meat eat-PRF-PST-3SG  
 ‘The cat had eaten meat’
- f. sän dʒäl-**ip**-sän  
 2SG come-PRF-2SG  
 ‘You (sg) have come’

This suppletion process appears to simply be optional, but this fact doesn’t impede an analysis of the contexts where it can and cannot occur. While we’ve seen that *-(j)Ib* is usually unavailable in contexts with 1st person subjects like (1a/b) above, interestingly, this is not always the case. There is one unique circumstance where *-(j)Ib* arises even with a first person subject—contexts that involve stacking the evidential on top of the perfect like (2) below:

(2) **Perfect + evidential stacking contexts use *-(j)Ib* in first person**<sup>3</sup>

- a. biz gatʃ-**ib**-miʃ-iʏ  
 1PL run-PRF-EVID-1PL  
 ‘Apparently we had run’
- b. biz comjunizm-i jaʃa-**j**ib****-miʃ-iʏ  
 1PL communism-ACC live-PRF-EVID-1PL  
 ‘Apparently we have lived through communism’
- c. bu sähär män halva-ni je-**j**ib****-miʃ-äm  
 this morning 1SG halva-ACC eat-PRF-EVID-1SG  
 ‘This morning I (apparently) have eaten the halva’
- d. män jat-**ib** gal-**ib**-miʃ-am  
 1SG sleep-CNVB stay-PRF-EVID-1SG  
 ‘Apparently, I have been asleep’

Since *-(j)Ib* suppletion typically requires a 2nd/3rd person subject, we would have expected the 1st person subjects in (2) to make *-(j)Ib* unavailable. Thus (2) presents a puzzle.

I hypothesize that the over-application of *-(j)Ib* suppletion in examples like (2) is forced by a highly-ranked phonological constraint against forms with adjacent identical morphemes (Stemberger 1981, Menn & MacWhinney 1984, Yip 1998, Plag 1999). That is, *-(j)Ib* is used outside of its typical licensing context here because, if it did not, an illegal /-mIʃ-mIʃ/ string would be generated. If this explanation is on the right track, and the required constraint is indeed a phonological one on the shape of adjacent morphemes, then this analysis provides evidence that the determination of morphological form must interact, and compete, with phonological constraints (Wolf 2008, 2009, Pertsova 2015).

1.1 ROADMAP. Section 2 provides more background on the morphemes in question. Section 3 overviews the theory I use to resolve the puzzling over-application of *-(j)Ib* suppletion. Section 4 demonstrates the implementation of the solution. Section 5 considers a remaining puzzle about the distribution of *-(j)Ib*, and the conclusion in section 6 notes a few issues for further study.

<sup>3</sup>In (2d) we also see a use of the converb *-(j)Ib*, which forms certain complex predicates, such as *be asleep* in this example. This morpheme is diachronically related to the *-(j)Ib* variant of *-mIʃ*, but I won’t discuss this here. See Zalansky (2019) for a recent overview of this morpheme’s diachronic history.

**2. Background.** Since this paper is mainly concerned with a word-internal pattern, I will say little about word-external syntax. It is enough to know that Azeri, like other Turkic languages, is strictly head-final with default SOV word order, and agglutinating suffixal morphology.

As mentioned, *-mİf ~ -(j)İb* has both perfect and evidential uses. The closely related language Turkish has a morpheme *-mİf* with similar semantics (Kornfilt 1997, Göksel & Kerslake 2005). However, while the Turkish *-mİf* usually corresponds to an evidential, with perfect uses more restricted, for the Azeri *-mİf ~ -(j)İb* this state of affairs is reversed—a perfect reading is typical, with the evidential reading less salient, and perhaps becoming obsolete (Authier 2010).<sup>4 5</sup>

The evidential interpretation of *-mİf ~ -(j)İb* does clearly emerge in some contexts, however. The first such context is the stacked examples introduced in (2), where we get both perfect and evidential semantics occurring together. The second is configurations where *-mİf ~ -(j)İb* is attached to the copular stem *i ~ -j*, as in (3):

(3) **Evidential interpretation**

- a. mən xəstə-j/i-mİf-äm  
1SG sick-be-EVID-1SG  
'Apparently I was sick (but didn't realize it until now)'
- b. at xəstə i-mİf-ø  
horse sick COP-EVID-3SG  
'The horse really is sick!'<sup>6</sup>

This contrasts with similar examples like (4) which use the copular form *ol* instead, and have a perfect interpretation.<sup>7</sup>

(4) **Perfect interpretation**

- a. mən xəstə ol-muʃ-äm  
1SG sick be-PRF-1SG  
'I have been sick'
- b. daş at ol-ub-ø  
stone horse become-PRF-3SG  
'The stone has become a horse'<sup>8</sup>

In Davis (2017), I argued that while copular V has a default form *ol*, a copula that raises to T suppletes depending on T's features, for instance, to the form *i ~ -j* in (3).<sup>9</sup> Thus the structure of (3) is as shown in (5) below, where copular V moves to a evidential T, expressed as *-mİf*:

<sup>4</sup>Öztopçu's (2003) Azeri grammar refers to *-mİf ~ -(j)İb* only as a perfect, with no mention whatsoever of an evidential use, which is suggestive that this reading is being lost.

<sup>5</sup>The evidential reading is often translated by my consultant as including past tense, even when there is no overt past tense morphology (*-dİ*). I suspect that this is because the evidential reading of this morpheme is essentially an inferential one, the inference typically being made based on events that occurred before the time of speaking.

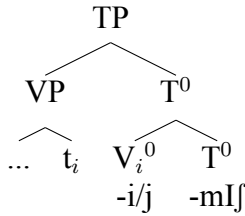
<sup>6</sup>Context: The vet lied to me!

<sup>7</sup>Though (4b) here has a "become" reading rather than a plain predicational reading. *ol* may be generally ambiguous between "be" and "become", though in Davis (2017) I suggest that we can regard *ol* as a copula in either case.

<sup>8</sup>Context: Because a sculptor carved it.

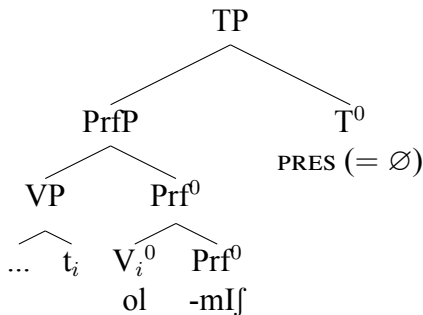
<sup>9</sup>The form *i ~ -j* is used in, roughly speaking, finite environments. When a copula is merged to a non-finite T, as in certain embedded contexts, the copula defaults to *ol*.

(5) **V<sup>0</sup> movement to evidential T<sup>0</sup> -*mIf* feeds copular allomorphy**



In contrast, in Davis (2017) I argued that there is a set of clause-medial functional heads, including various aspects as well as the perfect, which block V to T movement. In such cases we have a structure like in (6), corresponding to (4) above, where V movement is interrupted by the perfect *-mIf*. Since the copula does not raise to T in this circumstance, but is trapped within the lower perfect head, it cannot supplete sensitive to features of T, and thus takes on the default form *ol*.

(6) **V<sup>0</sup> to T<sup>0</sup> movement blocked by low perfect -*mIf*, bleeding copular allomorphy**



In short, the differing copular forms in (3-4) can be attributed to differences in the underlying syntactic position of *-mIf*: the evidential *-mIf* is high, perhaps in T, while the perfect *-mIf* is lower in the functional spine.<sup>10</sup> If this is correct, given the head-finality of this language, we expect the evidential to have been linearized right of the perfect in the stacked examples introduced in (2).

(7) **The evidential is structurally above, and so linearly right of, the perfect**

biz gatʃ-ib-mif-iʏ  
 1PL run-PRE-EVID-1PL

‘Apparently we had run’

The inner of these two morphemes, the perfect, is therefore apparently the one undergoing unexpected *-(j)Ib* suppletion in (2).<sup>11</sup> With this established, in the next section, I’ll overview a theory that can make sense of how this unexpected suppletion occurs.

<sup>10</sup>Similarly, Kelepir (1996) argues for the closely related Turkish that the evidential *-mIf* occupies T, while the perfect *-mIf* sits in a lower position.

<sup>11</sup>We might ask why the evidential doesn’t undergo suppletion instead. This is in fact predicted to be grammatical by the account I’ll offer in section 4, though none of the data I currently have show this, and I did not have the chance to directly test it. If it turns out that the evidential cannot supplete in these circumstances, it may have something to do with the phenomenon considered in section 5.

**3. Integrated morpho-phonology.** I argue for a theory like that of Wolf (2008, 2009), who proposes a particular integration of morphology and phonology. Wolf takes as his starting point the Distributed Morphology framework (Halle & Marantz 1993, Harley & Noyer 1999, a.o.), which hypothesizes that phonological information is assigned post-syntax. That is, the derivation eventually passes a structure off to the PF side of the derivation, which assigns form to the leaves of the tree, depending on their syntactic features. Within Distributed Morphology, this assignment of form is termed *Vocabulary Insertion*, involving rules like the following:

(8) **Some sample Vocabulary Insertion rules for English**

- a. [+Plural]  $\leftrightarrow$  -z
- b. [+Plural]  $\leftrightarrow$  -en / N\_, where  $N \in \{ox, child\}$
- c. [+Present, 3sg]  $\leftrightarrow$  -s
- d. [+Past]  $\leftrightarrow$  -d

In this system, the *subset principle* requires that the morpheme assigned by Vocabulary Insertion at a given syntactic node is the one that matches as many features as possible in the context of insertion, while also not corresponding to any features not present in that context. In reality, this principle is not always strictly obeyed, and Wolf argues that one reason for this is that phonological requirements sometimes out-compete morphological ones. In particular, Wolf argues that Vocabulary Insertion is integrated with the phonological component of the grammar, and in particular, that both phonology and morphology are governed by Optimality Theoretic constraints (Prince & Smolensky 1993) and thus can interact.

The morphological constraint relevant to the Azeri puzzle is stated in (9), which penalizes inserting morphemes that correspond to features missing from the context of insertion:

- (9) **Match(F)**: For every instance of the morphosyntactic feature F which a morpheme is specified as spelling out, assign a violation-mark if that morpheme has been inserted onto a feature-bundle which does not contain F. (Wolf 2009, ex. 9)

I argue that (9) is outranked by a constraint against strings with adjacent phonologically identical morphemes, stated in (10). This formulation will do the job for this project, but numerous works have explored the ways that requirements of this sort ought to be stated/explained (Stemberger 1981, Menn & MacWhinney 1984, Yip 1998, Plag 1999, a.o).

- (10) **No Repetition (NoRep)**: Assign a violation for each sequence of adjacent phonologically identical morphemes.<sup>12</sup>

In the languages of the world, such constraints often trigger processes like haplology or dissimilation. Given that (10) is sensitive to the phonological form of morphemes, it must be evaluated at a stage of the derivation that overlaps with the phonological part of the grammar.

I argue that the unexpected suppletion that this paper is concerned with is, in essence, a case

<sup>12</sup>Matthew Zalansky (p.c.) informs me that there is independent evidence for a constraint of this sort from the facts about certain constructions with double causative or passive marking. In these contexts, the same allomorph of the causative or passive cannot be stacked: it is necessary to dissimilate by using an alternative form. An anonymous reviewer points out that the similar dissimilation in causative stacking is common in Turkic more generally, including Turkish and Kazakh. Facts about the impossibility of multiple genitive marked phrases in Kyrgyz (Laszakovits 2019) could have a related explanation.

of dissimilation: the ranking of NoRep over Match(F) allows suppletion to occur outside of its typical licensing context, to avoid an illegal *-mIf-mIf* string. If NoRep really is a phonological constraint, this morphological dissimilation provides evidence that phonology interacts with the mechanisms that determine morphological exponence. This is because if Vocabulary Insertion rules did not interact with the phonological part of the grammar, there is no way that a constraint like (10) could influence morpheme choice.

**4. Implementation: Forcing a suppletion mismatch.** To review the puzzle, we've seen that the morpheme *-mIf* has a variant *-(j)Ib*, which cannot occur with 1st person subjects (11a). However, *-(j)Ib* exceptionally occurs in certain stacked perfect + evidential contexts, even when the subject is 1st person (11b):


- (11) a. **Typical ban on *-(j)Ib* in 1st person**  
**biz** dʒäl-miʃ/\***ib**-ij  
 1PL COME-PRF-1PL  
 'We have come'
- b. **Stacked contexts allow *-(j)Ib* in 1st person**  
 biz comjunizm-i jaʃa-j**ib**-miʃ-iy  
 1PL COMMUNISM-ACC LIVE-PRF-EVID-1PL  
 'Apparently we have lived through communism'

Since I've claimed that in examples like (11b) unfaithful suppletion applies to the perfect version of *-mIf* ~ *-(j)Ib*, the perfect is what I'll refer to in the following implementation.

First, morphemes expressing the perfect should be specified as corresponding to a feature like [Prf], and thus, their insertion will be governed by the constraint Match([Prf]). Since the *-(j)Ib* variant can only occur with non-1st person subjects, its insertion must be governed by an additional constraint like Match([-Speaker]).<sup>13</sup> I assume that T agrees with the subject, and that the phi-features gathered by T via agree serve as the context for potential *-(j)Ib* suppletion.

Abstracting away from the form of agreement marking on T, the table in (12) shows a circumstance where *-(j)Ib* suppletion fails. Here the input to the morpho-phonological evaluation contains [+Speaker] (1st person) features on T, so insertion of *-(j)Ib* here would incur a violation of Match([-Speaker]), as in (12a). Insertion of *-mIf* is governed only by Match([Prf]), however, and so it can be inserted without violation, as in (12b):

(12) **Typical lack of *-(j)Ib* with a 1st person subject**

V-PRF-T[+Speaker]	NoRep	Match(-Speaker)	Match(Prf)
a. V-(j)Ib-T		*	✓
 b. V-mIf-T		✓	✓

In (12), the highly-ranked constraint NoRep is inactive. This constraint will become active when

<sup>13</sup>The influential feature geometry proposed in Harley & Ritter (2002) does not contain a position that can single out the addressee (2P) and 3rd person, but exclude the speaker, since both 1st and 2nd person are dominated by one [Participant] node. While a hypothetical feature [-Speaker] is indeed the simplest way to refer to the conditions permitting *-(j)Ib*, a way of avoiding this tension would be to posit instead that *-(j)Ib* is independently triggered by 2nd person and 3rd person. In this case, the fact that *-(j)Ib* is triggered by the complement set of 1st person features would simply be coincidence, rather than a reflection of a [-Speaker] feature in the morpho-syntax.

we consider the stacking contexts, modeled in (13) below. Here the (a) form with double *-mIf* doesn't violate Match([-Speaker]), but violates the highly ranked NoRep. Thus the winner is the (b) form, which violates Match([-Speaker]), but satisfies NoRep by using the *-(j)Ib* allomorph:<sup>14</sup>

(13) **NoRep forces *-(j)Ib* even with 1st person subject**

V-PRF-EVID-T[+Speaker]	NoRep	Match(-Speaker)	Match(Prf)
a. V- <b>mIf</b> - <b>mIf</b> -T	*	✓	✓
☞ b. V- <b>(j)Ib</b> - <b>mIf</b> -T	✓	*	✓

In this way, the dominance of NoRep permits an otherwise impossible morphological mismatch.<sup>15</sup> As expected, a double *-mIf* configuration is not grammatical, as we see in (14).<sup>16</sup>

(14) **No double *-mIf* within one word**

\*/?? mən gatʃ-miʃ-miʃ-am  
 1SG run-PRF-EVID-1SG

'Apparently I had run'

But the morpho-phonological evaluation evidently only applies maximally at the word level, since in periphrastic auxiliary constructions with two instances of *-mIf* that are not adjacent, there is no violation, as (15) shows.

(15) **Double *-mIf* possible when separated by a word boundary**

✓ gatʃ-miʃ ol-muʃ-am  
 run-PRF be-EVID-1SG

'I realize that I have run'

To conclude this section, I'll note a detail of this implementation that I have so far overlooked. Notice that the schema for Match constraints in (9) above refers to the mapping

<sup>14</sup>This account predicts double *-(j)Ib* to be impossible. While I have not seen such examples, this should be tested.

<sup>15</sup>This discussion has assumed that lexical insertion for all the morphemes in the word occurs in parallel, since this is the most straightforward way to implement having the form of one morpheme influence the allomorph chosen for another. This is contrary to the widely adopted view that lexical insertion occurs strictly bottom-up, though Deal & Wolf (2016) argue that while spellout does proceed from the lowest domain upward, spellout within a given domain need not be strictly bottom up: domain-internally, spellout may also proceed top-down, or in parallel.

<sup>16</sup>Matthew Zalansky (p.c.) informs me that while such strings are indeed bad for some speakers, they are fine for others. The analysis of this paper thus only applies to speakers who pattern like my own consultant, whose reaction was that (14) sounds like Turkish, not Azeri. Indeed, Turkish does permit examples with double *-mIf*. I hypothesize that such strings are grammatical in Turkish because there is simply no possible form that would be morpho-phonologically superior, given that Turkish does not have the *-(j)Ib* variant of *-mIf* to take advantage of.

- i. **Double *-mIf* in Turkish** (Adapted from Goksel & Kerslake 2005, section 23.2.2, ex. 27.)  
 Meğer Hasan çoktan tıbbi bitirmişmiş  
 'It turns out that Hasan completed his medical studies a long time ago'

Variance of this sort within Azeri could of course be accommodated by positing different constraint rankings for different speakers, though on this I currently can only speculate.

between a morpheme, and the syntactic feature bundle that the morpheme is being inserted onto. I have used the intuition of Match to model the distribution *-(j)Ib*, but in doing so, I have expanded the set of situations to which Match applies. This is because, though *-(j)Ib* suppletion is sensitive to the phi-features of the subject (mediated by T), those phi-features are not directly present on the [Prf] node where *-(j)Ib* is actually being inserted. Those phi-features are local, but they are not in fact present on the same node. Thus Match is, strictly speaking, not applicable here, unless we expand Match to be relevant both to the mapping between morphemes and syntactic nodes, and to the relationship between morphemes and the local context that licenses them.

**5. A remaining puzzle: V-adjacency.** Here I'll comment on another puzzle about the distribution of *-(j)Ib*. While I cannot solve this puzzle here, its solution will likely clarify how exactly the implementation shown in the previous section should be stated.

Recall that in all the examples of over-applied *-(j)Ib* suppletion that we've seen, the evidential stacks on the perfect, and the perfect suppletes to *-(j)Ib*. This avoids producing a *-mIf-mIf* string which would violate NoRep. Notice that suppletion of the evidential to *-(j)Ib* would also satisfy NoRep, since *-mIf-Ib* would be just as good of a form as *-Ib-mIf* in this regard. However, none of the data I currently have shows forms with *-mIf-Ib*.<sup>17</sup> This fact may be related to another puzzle about *-(j)Ib* so far unmentioned: in all the data I currently have access to, *-(j)Ib* must always be adjacent to the V stem, as (16) suggests. Here we see that the presence of aspect morphology between V and *-(j)Ib* is not allowed, though such examples are fine if *-mIf* is used instead:

**(16) Evidence for a V-adjacency requirement on *-(j)Ib***

- a. \* at/o gatf-adʒay-ip-∅  
horse/3SG run-PROS-PRF-3SG  
'The horse / he will have run'
- b. \* at gatf-ir-ip-∅  
horse run-PROS-PRF-3SG  
'The horse has been running'
- c. o gatf-ip ol-adʒay-∅  
3SG run-PRF be-PROS-3SG  
'3SG will have run'
- d. o gatf-ib-di-∅  
3SG run-PRF-PST-3SG  
'3SG had run'

Supposing that there is such a V-adjacency requirement on *-(j)Ib*, if the evidential suppleted in the stacked contexts like (2) above, the resulting form would end up with a *-(j)Ib* which illegally has the perfect between it and V (*V-mIf-Ib*). If this hypothesis is right, however, the presence of a V-adjacency requirement on *-(j)Ib* simply raises another puzzle.

In the implementation of the previous section, I assumed that *-(j)Ib* suppletion is conditioned by the phi-features present on T, gathered by T agreeing with the subject. However, there is some evidence that this is likely not correct. For instance, *-(j)Ib* suppletion can occur in embedded structures where T may not be present, like the bracketed constituents in (17) below. In Davis (2017), I argued that these constituents are sub-TP structures that are embedded under an

<sup>17</sup>Matthew Zalansky (p.c.) informs me that *-mIf-Ib* is almost certainly ungrammatical, but I have unfortunately not had the opportunity to test this.



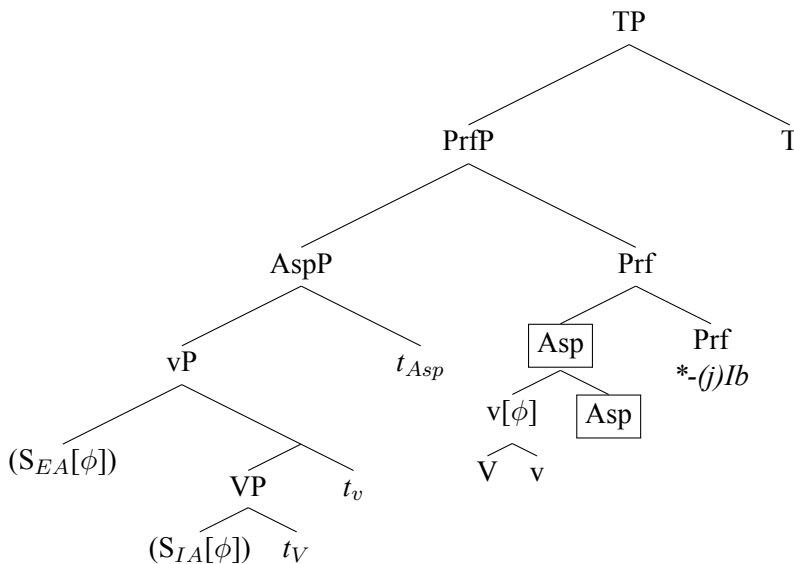
auxiliary copula. If these are indeed sub-TP constituents, *-(j)Ib* suppletion within them evidently succeeds despite the perfect head not being local to T:

(17) **Perfect embedded under auxiliary can still supplete**<sup>18</sup>

- a. o [gatʃ-ip] ol-adʒay-ø  
 3SG run-PRF be-PROS-3SG  
 ‘3SG will have run’
- b. män [at-in gatʃ-mif/ib] ol-diɣ-in-a inan-ir-am  
 1SG horse-GEN run-PRF be-T-3SG-DAT believe-IMPV-1SG  
 ‘I believe that the horse has run’

This fact would be compatible with the suppletion-triggering features not being on T, but rather on v/V. If this is so, the lack of suppletion when aspect intervenes (16a/b) would emerge from a theory where an element undergoing contextual allomorphy must be directly local to the allomorphy trigger (Bobaljik & Harley 2017). For instance, v/V might obtain the subject’s phi-features through either Agree or a selection relationship with the subject. Assume that the agglutinative verbal complex is formed by head movement. If something like an Asp(ect) head is present in the structure, it intervenes between the perfect head and v-V, thus blocking suppletion:

(18) **Aspect intervenes between Prf and v-V, blocking suppletion**<sup>19</sup>



If Asp is absent, head movement will bring v/V local to the perfect head, permitting suppletion.

<sup>18</sup>Such examples could be important for ruling out a separate approach to the facts analyzed in this paper. Notice that in example (1), which initially demonstrated that *-(j)Ib* is ungrammatical with 1st person subjects, we see that both the singular and plural 1st person agreement markers begin with a vowel (*-Am*, *-Ij/Iy*). It turns out that these are the only agreement markers that begin with a vowel in this language. Thus, there remains the possibility that *-(j)Ib* is not actually blocked by the presence of 1st person subjects, but rather for some reason cannot be followed by a vowel-initial suffix. If this hypothesis is right, *-(j)Ib* should be usable with a 1st person subject within the bracketed domains shown in example (17), since these constituents lack agreement marking. In this case, we also predict placing past tense marking between agreement and *-(j)Ib* as in examples like (16d) to permit forms where *-(j)Ib* co-occurs with 1st person agreement, as a reviewer points out. I have not had the chance to test such examples.

<sup>19</sup>In Davis (2017) I argue that head movement goes no further than Asp/Prf (when present), as modeled here.

However, even if what I've just said is roughly correct, it remains unclear why *-(j)Ib* would need to be local to v/V even in circumstances where suppletion over-applies, like (2). Since in such examples the features that typically permit *-(j)Ib* are absent altogether, being adjacent to v/V should be irrelevant. This puzzle will have to remain unresolved for now.<sup>20</sup>

**6. Conclusion.** In this paper, I described a pattern of suppletion in Azeri that can occur outside of its typical licensing context, only when a sequence of phonologically identical elements would be created. I argued that this is a phonologically motivated morphological mismatch, which supports a theory in which phonology and lexical insertion interact and compete. There are many more questions to ask about these facts, however. Many of these questions have been mentioned in footnotes, and I'll end this paper by discussing a couple more.

I have mostly focused on the possible forms of the perfect, assuming that the inner morpheme in the stacked contexts under consideration is indeed the perfect. However, I have not tested in detail whether the evidential *-mIʃ* allows *-(j)Ib* in exactly the same circumstances where the perfect does. The growing obsolete-ness of the evidential in Azeri makes this difficult to test.

Another remaining puzzle is that *-(j)Ib* appears to be unavailable in relative clauses, as we see in (19). If these sentences contain a *wh*-operator co-indexed with the subject, we might expect the 3rd person features of the subject to be able to trigger *-(j)Ib*, contrary to fact. However, it is notable that these are subject relatives, which in Azeri, as in Turkish, lack subject agreement.<sup>21</sup> This lack of agreement could be the cause of the unavailability of *-(j)Ib*, if this morpheme's use is contingent on the appropriate phi-features being present on T post-Agree with the subject, as assumed in the initial part of this paper.

(19) **No *-(j)Ib* in relative clauses**

- a. [gatʃ-miʃ] at jodyn-dyr  
run-PRF horse tired-COP.3SG  
'The horse that has run is tired'
- b. \* [gatʃ-ib] at  
run-PRF horse  
'The horse that has run'
- c. [gatʃ-miʃ] tʃi  
run-PRF man  
'Man who has run'

<sup>20</sup>Furthermore, a reviewer points out that *-(j)Ib* is apparently not blocked by intervening negation, as in an example like *gör-me-yib* (see-NEG-PRF), which raises questions about what exactly the adjacency condition I am hypothesizing here is defined in terms of. I speculate that perhaps negation counts as sufficiently "verbal" for these purposes, but it is difficult to say more on this without systematically studying the interaction of *-(j)Ib* with more potential interveners. This is yet another angle for future work.

<sup>21</sup>While non-subject relatives do show agreement with the subject of the relative clause, *-mIʃ* is actually unavailable in such relatives, which hence makes it unsurprising that *-(j)Ib* is not available in them either:

- ii. \* [it-in itʃ-ib/miʃ-i] su  
dog-GEN drink-PRF-3SG water  
'Water that the dog has drunk'

I thank an anonymous reviewer for pointing out this fact.

- d. \* [gatʃ-ib] tʃi  
 run-PRF man  
 ‘Man who has run’

More data is needed before the cause of this effect can be examined further.

## References

- Authier, Gilles. 2010. Azeri morphology in Kryz (East Caucasian). In Lars, Joanson. (ed.) *Turkic Languages* 14, 14-42.
- Bobaljik, Jonathan and Heidi Harley. 2017. Suppletion is Local: Evidence from Hiaki. Maire Noonan, Glynne Piggot, and Lisa Travis (eds.), *The structure of words at the interfaces*. Oxford: OUP. DOI: [10.1093/oso/9780198778264.003.0007](https://doi.org/10.1093/oso/9780198778264.003.0007)
- Davis, Colin. 2017. Locality and copular allomorphy in North Azeri. In L. Zidani-Eroğlu, M. Ciscel, and El. Koulidobrova (eds.), *Proceedings of the 12th Workshop on Altaic Formal Linguistics* (WAFL12). MITWPL: Cambridge.
- Deal, Amy Rose, and Matthew Wolf. 2016. Outwards-sensitive phonologically conditioned allomorphy in Nez Perce. In Vera Gribova and Stephanie Shih (eds.), *The Morphosyntax Phonology Connection*, pp. 29-60. Oxford: Oxford University Press. DOI: [10.1093/acprof:oso/9780190210304.001.0001](https://doi.org/10.1093/acprof:oso/9780190210304.001.0001)
- Göksel, Asli., & Celia. Kerslake. 2005. *Turkish: A Comprehensive Grammar*. Abingdon, UK: Routledge.
- Halle, Morris & Alec, 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, Heidi and Elizabeth Ritter. 2002. Person and Number in Pronouns: A Feature-Geometric Analysis. *Language* 78. 482-526.
- Harley, Harley & Rolf, Noyer. 1999. Distributed morphology. In *Glott International* 4:4. 3 - 9.
- Kelepir, Meltem. 1996. *Topics in Turkish Syntax: Clausal Structure and Scope*. PhD Dissertation, MIT.
- Kornfilt, Jaklin. 1997. *Turkish*. London: Routledge
- Laszakovits, Sabine. 2019. *Kyrgyz Relative Clauses and Dependent Case Theory*. Talk given at the 4th workshop on Turkic and languages in contact with Turkic (Tu+4).
- Menn, Lise. and Brian, MacWhinney. 1984 The Repeated Morph Constraint. *Language*: 60, 519-541.
- Öztopçu, Kurtuluş. 2003. *Elementary Azerbaijani*. Santa Monica, California.
- Pertsova, Katya. 2015. Interaction of morphological and phonological markedness in Russian genitive plural allomorphy. *Morphology* 25, 229-266.
- Plag, Ingo. 1999. Morphological haplology in a constraint-based morpho-phonology. <https://ling.auf.net/lingbuzz/repo/ROA/article/000481>
- Prince, Alan, and Paul Smolensky (2004 [1993]). *Optimality Theory: Constraint Interaction in Generative Grammar*. Oxford: Blackwell.
- Stemberger, Joseph Paul. 1981. Morphological Haplology. *Language*. 57, 791-817.
- Şirəliyev, Məmmədağa and Ervand, Sevortjan. 1971. *Grammatika Azerbajdzanskogo Yazyka: Fonemika, Morfologiya i Sintaksis* [Grammar of the Azerbaijani Language: Phonetics, Morphology and Syntax]. Baku: Akademija Nauk Azerbajdzanskoj SSR.

- Wolf, Matthew. 2008. Optimal interleaving: Serial phonology-morphology interaction in a *constraint-based model*. PhD Dissertation, University of Massachusetts Amherst.
- Wolf, Matthew. 2009. Lexical insertion occurs in the phonological component. Bernard Tranel (ed.), *Understanding Allomorphy: Perspectives from Optimality Theory*. London: Equinox. DOI: [10.1558/equinox.25223](https://doi.org/10.1558/equinox.25223).
- Yip, Moira (1998): Identity Avoidance in Phonology and Morphology. In: Stephen G. Lapointe, Diane K. Brentari, and Patrick M. Farrell (eds.): *Morphology and its Relation to Phonology and Syntax*. Stanford, CA: CLSI Publications, 216-246.
- Zalansky, Matthew. 2019. *The Overabundance of the Perfect and the Decay of Evidentiality in Standard Azerbaijani*. Talk given at the 4th workshop on Turkic and languages in contact with Turkic (Tu+4).