

# UC San Diego

## San Diego Linguistic Papers, Issue 7

### Title

Systematic overabundance exhibits systematic differentiation: On person marking, perfect marking, and evidentiality in Azerbaijani

### Permalink

<https://escholarship.org/uc/item/11t996k7>

### Author

Zaslansky, Matthew

### Publication Date

2019-09-16

## Systematic overabundance exhibits systematic differentiation: On person marking, perfect marking, and evidentiality in Azerbaijani

Matthew Zaslansky\*

**Abstract.** This article investigates the synchronic interaction between person-marking and the Standard Azerbaijani perfect paradigms. The second and third persons exhibit a phenomenon known as overabundance (Thornton 2011, 2012). Unlike many previous examples of overabundance in the literature, the variation of the present perfect in Standard Azerbaijani applies to all the relevant verb lexemes in the language, exhibiting systematic overabundance (Bonami & Stump 2016). I argue that systematic overabundance can yield systematic differentiation. The distribution of perfect markers along different persons is understood, at least in part, as indicating a contrast with the evidential clitic. ‘Differentiation by Person’ (Dmitriyev 1927, Əfəndiyeva 2005) seems to be a general structural property of the Azerbaijani verb paradigm, at least in certain TAM categories. This also provides a potential language-internal motivation for the (apparent) restriction of evidential readings to the third person.

### 1. Introduction

Like most Turkic languages, Azerbaijani (Western Oghuz) tends to be characterized by regular agglutinative morphology, i.e., juxtaposing synthetic mappings of form to meaning, with numerous bound forms in a word, and with each morpheme having few and phonologically predictable allomorphs (Johanson 1998a). The perfect aspect in Azerbaijani is a non-predictable violation of this otherwise characteristic transparency of form and meaning. Most traditional descriptions of Azerbaijani identify two synonymous perfect suffixes,  $-(y)Ib^1$  and  $-mIş$ , and claim that  $-(y)Ib$  can mark second or third persons, freely varying with  $-mIş$  (Şirəliyev and Sevortyan 1971: 125, Hüseynzadə 2007: 151, Fəxrəddinqızı 2010: 73-74, Kazımov 2014: 175). The focus of this article is on the competition between the rival perfect suffixes  $-mIş$  and  $-(y)Ib$  in Standard Azerbaijani. Both  $-mIş$  and  $-(y)Ib$  each have their own predictable sets of allomorphs, largely due to vowel harmony. For example,  $-mIş$  occurs with a high front vowel /i/ when the stem ends in a front unrounded vowel, as in *getmişəm* ‘I have gone’, *əmişdirmişəm* ‘I have suckled’; with a front high rounded vowel /y/ when the stem ends in a front rounded vowel, e.g. *söndürmü(ş)sən<sup>2</sup>*

---

\* I am grateful to Lamiyə Adışova, Lalə Mahmudova, Aydan Mammadova, Günel Salayeva, and 10 other anonymous Azerbaijanis who provided me with their judgements on the data in this presentation. I thank Farrell Ackerman, the UCSD BrownBag community for their useful feedback on an earlier draft of this paper. All mistakes are my own. Author: Matthew Zaslansky, University of California, San Diego ([mzaslans@ucsd.edu](mailto:mzaslans@ucsd.edu)).

<sup>1</sup> I employ Standard Azerbaijani Latin orthography throughout this paper. The orthography tends to correspond to IPA equivalents in broad transcription, except  $j = /ʒ/$ ,  $ş = /ʃ/$ ,  $ç = /tʃ/$ ,  $c = /dʒ/$ ,  $k = /c~k/$ ,  $g = /ʒ/$ ,  $q = /g/$  (often spirantized as [x] in codas),  $ğ = /ɣ/$ ,  $y = /j/$ ,  $ə = /æ/$ ,  $ö = /œ/$ ,  $ü = /y/$ ,  $ı = /w/$ . Vowel and consonant harmony spreads from stems to suffixes. Following turkological convention, archiphonemes which undergo harmony are written with capital letters: *I* (unifies /i/, /y/, /u/, and /w/), *A* (unifies /æ/ and /ɑ/) and *Q* (unifies /c/ and /g/).

<sup>2</sup> Depending on the speaker’s idiolect, the  $ş /ʃ/$  in  $-mIş$  may be deleted/assimilated when followed by either the second person singular  $-sAn$ , the second person plural  $-sInLz$ , or the conditional suffix  $-sA$ . While such deletion does seem to have been a phonetically motivated process of sibilant hapology at some point in the language, it appears that—for those speakers who do allow for hapology here—the process is not necessarily phonetic or even phonological, since it only happens with  $-mIş$ . Consider the forms from one speaker, e.g. *yaz-mi-san* ‘you have writ-

‘you have extinguished’, *hürkmüşük* ‘we have been startled’, etc.<sup>3</sup> Table 1 shows the typical paradigm for the perfect forms, demonstrated with the verbs *al-* ‘take’ and *get-* ‘go’.

	<i>al-</i> ‘take’	<i>get-</i> ‘go’
1SG	<i>almışam</i>	<i>getmişəm</i>
2SG	<i>almi(ş)san~alıbsan</i>	<i>getmi(ş)sən~gedıbsən</i>
3SG	<i>almışdır~alıb(dır)</i>	<i>getmişdir~gedib(dır)</i>
1P	<i>almışığ</i>	<i>getmişik</i>
2P	<i>almi(ş)sınız~alıbsınız</i>	<i>getmi(ş)siniz~gedıbsiniz</i>
3P	<i>almışdırlar~alıb(dır)lar</i>	<i>getmişdirlər~gedib(dır)lər</i>

Table 1: Perfect paradigms for the verbs *al-* ‘take’ and *get-* ‘go’

While the existence of competing morphologically complex forms is not characteristic of Turkic inflection generally, it is perhaps not uncommon across languages. Rival forms of this sort have often been referred to as doublets (see e.g. Kroch 1994, Fehring 2004), exemplified in English by certain classes of verbs which exhibit past tense pairs such as *dived/dove* and *leaped/leapt*, among others (Haber 1976). The phenomenon—when linguists have cared to address it—has sometimes been referred to as doubletism (Lečić 2017). Thornton (2011, 2012) has observed that some languages, such as Latin and Italian, sometimes allow for more than two competing forms for certain cells in inflectional paradigms. Traditional ‘doublets’ are therefore instances of a more general sort of morphological variation called OVERABUNDANCE, which Thornton defines as the existence of two more forms (referred to as CELL-MATES) for the realization of a single cell in a paradigm.

The present paper is part of the first detailed study of the overabundance of the perfect in Standard Azerbaijani. By Standard Azerbaijani (henceforth just Azerbaijani), I mean the standard language written and spoken in the Republic of Azerbaijan. Non-standard varieties are spoken throughout the republic, as well as in Iran, Georgia, Russia (Daghestan), Eastern Turkey, and Iraq. Descriptions of the competing forms present perfect forms like *getmi(ş)sən/gedıbsən* ‘you have gone’ have typically described them as just-so synonymous inflections, typically without further comment, except to note that *-(y)lb* does not mark the first person. As Şirəliyev (1958) has demonstrated, the standard variety contains features primarily from the dialects of Şirvan and Bakı-Şamaxı, but also from other dialects. However, the synchronic overabundance of the present perfect is not obviously traceable to dialect mixing. I have discussed the diachrony of the relevant markers elsewhere in more detail (Zaslansky, submitted), but will only briefly touch on that here. Much of the data in this paper is taken from a pre-publication version of the submitted paper.

Unlike most previous examples of overabundance discussed in the literature, the situation in Azerbaijani is not a property of individual lexemes or groups of lexemes, e.g. declension or conjugation classes à la English (see above), Latin (Thornton 2011), Italian (*ibid.*, Cappellaro 2018), or Croatian (Lečić 2017). The variation of the present perfect in Azerbaijani applies to all the

---

ten’, but *soruş-san* ask-COND.2SG ‘if you ask’ (but not *\*soru-san*); *günəş-sən* sun-COP.2SG ‘you are (a/the) sun’ (but not *\*günə-sən*). For many speakers, the exceptional hapology in 2<sup>nd</sup>-person *-mİş* is optional, and so may be considered yet another form of overabundance. Pursuing this analysis is beyond the scope of the present paper.

<sup>3</sup> /w/ when the stem ends in a back unrounded vowel, e.g. *qalmışdır*, etc. *-(y)lb* shows the exact same patterns, e.g. *gedıbsən* ‘you have gone’, *söndürübsən* ‘you have extinguished’, with the additional caveat that *-(y)lb* predictably occurs with a glide when following a vowel, e.g. *oxuyubdur* ‘(s)he has read’, cf. *oxumuşdur* ‘(s)he has read’.

relevant verb lexemes in the language and shows no indication of developing verb classes. As we will see, Azerbaijani rather exhibits what Bonami and Stump (2016: 16-17) call SYSTEMATIC OVERABUNDANCE in that each verb in the language has several sets forms in the present perfect. The synchronic facts of the present perfect and evidential markers are discussed in §2. I discuss the likely historical pathway for the current organization of the present perfect in §3. I then discuss the composite inflection known as the evidential of the perfect in §4. In §5 I invoke Tserunian's 'Differentiation by Person' (Dmitriyev 1927) as a general reason for why systematic overabundance exhibits systematic differentiation in Azerbaijani, and give some quantitative distributional evidence for this systematic differentiation by person in §6.

## 2. The category of the present perfect in Azerbaijani

The data in this section reflect the acceptability judgements of 14 native speakers of Azerbaijani, elicited in Baku. All participants were educated in Azerbaijani. Their judgements represent the spoken standard of the literary (=standard) language. I carried out elicitations in August-September 2014, July 2017, as well as over Skype between September and October 2017. In order to properly set the stage for our description of the Azerbaijani perfect markers, let us begin with two important observations:

(i) Just as in in other Oghuz languages—prototypically Turkish—verbs marked by *-miş* in Azerbaijani are reported to carry both temporal (perfect) or evidential (indirective, meaning 'evidently,' 'apparently,' 'reportedly,' etc.) readings (Hüseynzadə 2007: 169; among many others). However, such verbs in Azerbaijani may be ambiguous between the temporal and evidential readings, tending strongly towards a strictly temporal perfect reading rather than an evidential reading, unlike e.g. Turkish. Evidentiality is thus the weaker reading.

(ii) Unlike *-miş*, the *-(y)ib* perfect marker has no secondary reading. It is always unambiguously temporal.

While the default assumption following Johanson (1998b, 2002: 147, and elsewhere) has been that the secondary nature of the evidential reading for verbs marked by *-miş* is due to Persian influence, I propose the hypothesis (which is not necessarily mutually exclusive with Johanson's) that there also reasons internal to the Azerbaijani verb paradigm for the weaker evidential readings of *-miş*. Namely, I propose that person marking asymmetries in the present perfect paradigm created a morphologically marked distinction between the perfect and evidential senses of *-miş*, which is not marked in other persons. Hypothetically, this facilitated a retreat of evidentiality in the first and second persons. Our focus is therefore on person-marking.

### 2.1. The puzzle of person marking in the Azerbaijani perfect

As shown above in Table 1, the first person suffix *-Am* cannot co-occur with the *-(y)ib* perfect. The judgements in (1) reflect the traditional descriptions of asymmetrical person marking of suffixes *-miş* and *-(y)ib*: the two suffixes can be used to paraphrase each other, except in the first person.

- (1) a.   Könül    hər gün                   dolma                   ye-**miş**-dir / ye-**yib**(-dir).  
          Könül    every day               dolma.ACC             eat-PRF.MIŞ-3SG / eat-PRF.IB-3SG  
          'Könül has eaten dolma every day.'

- b. Mən hər gün dolma ye-**miş**-əm / \*ye-**yib**-əm.  
 Könül every day dolma.ACC eat-PRF.MIŞ-1SG / eat-PRF.IB-1SG  
 ‘I have eaten dolma every day.’

While most traditional descriptions only report that *-(y)Ib* is restricted to the second and third persons, many of my younger consultants gave degraded judgements even for the second person forms, as in (2).

- (2) a. Sən hər gün dolma ye-**mi(ş)**-sən / ?ye-**yib**-sən.  
 2SG every day dolma.ACC eat-PRF.MIŞ-2SG / eat-PRF.IB-2SG  
 ‘You have eaten dolma every day.’  
 b. Siz hər gün dolma ye-**mi(ş)**-sınız / ?ye-**yib**-sınız.  
 2P every day dolma.ACC eat-PRF.MIŞ-2P / eat-PRF.IB-2P  
 ‘You guys have eaten dolma every day.’

Speakers still generally accept the *-(y)Ib* forms in (2), but note that they probably would not use them frequently.

Only the third person singular suffix *-dIr* is reported to be optional, in contrast to the first and second persons. The exact nature of this optionality has not been clear in previous reports. Consider the following proposals regarding the distribution of the third person marker. Table 2 shows a present perfect paradigm adapted from the description given in Şirəliyev & Sevortjan, who claim that *-dIr* is always optional in the third person singular and plural of the perfect. On this view, that *-dIr* is a straightforward marker of the third person, separable from the third person plural marker *-lAr*.

	Singular	Plural
1	<i>yaz-miş-am</i>	<i>yaz-miş-ıq</i>
2	<i>yaz-mi(ş)-san~yaz-ıb-san</i>	<i>yaz-mi(ş)-sınız~yaz-ıb-sınız</i>
3	<i>yaz-<b>miş(-dIr)</b>~yaz-ıb(-dIr)</i>	<i>yaz-<b>miş(-dIr)</b>-lar~yaz-ıb(-dIr)-lar</i>

Table 2: Present perfect paradigm of *yaz-* ‘write’ (Şirəliyev and Sevortjan 1971: 125)

Table 3, on the other hand, shows a paradigm adapted from Öztopçu, who shows the same verb, but indicates that the *-dIr* suffix does not occur in the plural, and does not co-occur with *-(y)Ib*.

	Singular	Plural
1	<i>yaz-miş-am</i>	<i>yaz-miş-ıq</i>
2	<i>yaz-mi(ş)-san~yaz-ıb-san</i>	<i>yaz-mi(ş)-sınız~yaz-ıb-sınız</i>
3	<i>yaz-<b>miş-dır</b>~yaz-ıb</i>	<i>yaz-<b>miş-lar</b>~yaz-ıb-lar</i>

Table 3: Present perfect paradigm of *yaz-* ‘write’ (Öztopçu 2003: 331)

On this view, *-dIr* is a marker of the third person singular perfect in its *-mİş* form, while the third person singular in its *-(y)Ib* form is zero-marked. *-lAr* is then still a marker of the third person plural.

It should not be understated how important it is to properly characterize the facts of perfect marking here. The grammar by Şirəliyev and Sevortjan 1971 represents the standard academic reference on Azerbaijani for Russophone linguists and turkologists. While Öztopçu 2003 is a pedagogical text rather than an academic reference, it nonetheless represents the most comprehensive English-language reference for standard Azerbaijani. There is the earlier pedagogical grammar by Householder & Lofti (1965), but this is a grammar of the Tabrizi dialect rather than of what I have referred to as Standard Azerbaijani. In any case, we are left in a posi-

tion of having different information about the same variety in Russian and English. It is worth considering what information we might find in publications in the vernacular itself. Türkan Əfəndiyeva’s (2005) book-length *Felin keçmiş zaman formaları* [Forms of the Past Tense of the Verb] is a particular insightful overview of the category, including the perfect aspect, which is typically categorized as one of the ‘past-tenses’ in the Azerbaijani turkological tradition. The book appears to be return to the topic from the author’s earlier unpublished dissertation work (Hacıyeva [Əfəndiyeva] 1958).

Table 4 reflects Əfəndiyeva’s description of the facts.

	Singular	Plural
1	<i>yaz-mış-am</i>	<i>yaz-mış-ıq</i>
2	<i>yaz-mı(ş)-san~yaz-ıb-san</i>	<i>yaz-mı(ş)-sınız~yaz-ıb-sınız</i>
3	<i>yaz-mış-dır~yaz-ıb(-dır)</i>	<i>yaz-mış-dır-lar~yaz-ıb(-dır)-lar</i>

Table 4: Present perfect paradigm of *yaz-* ‘write’ (based on Əfəndiyeva 2005)

On this view, *-dır* is an optional marker of the third person singular and is separable from the plural marker *-lar* (as in Table 2), but it is not optional in the marking the third person perfect in its *-miş* forms. The three descriptions of third person marking given above present conflicting views of the facts. It is not possible for all of them to be correct. The descriptions either reflect different varieties (regional dialects, ethnolects, sociolects, idiolects) of Azerbaijani, or they reflect a confusion of the facts. Consider briefly that descriptions all agree that *-dır* displays a similar sort of variation elsewhere in the language, as seen in Table 5.

	Present	Aorist	Future/Prospective	Necessitative	Optative
3SG	<i>yaz-ır</i>	<i>yaz-ar</i>	<i>yaz-acaq(-dır)</i>	<i>yaz-malı(-dır)</i>	<i>yaz-a</i>
3P	<i>yaz-ır-lar</i>	<i>yaz-ar-lar</i>	<i>yaz-ar(-dır)-lar</i>	<i>yaz-malı(-dır)-lar</i>	<i>yaz-a-lar</i>

Table 5: Various partial paradigms of *yaz-* ‘write’

All descriptions agree that *-dır* is an optional marker of the third person in the future/prospective and in the necessitative, and not, for example, in the present (e.g. *yazır* ‘(s)he is writing’ but *\*yazırdır*), the aorist (e.g. *yazar* ‘(s)he would write’ but *\*yazardır*), or optative (e.g. *yaza* ‘that (s)he write’ but *\*yazardır*). The cause for the proliferation of conflicting descriptions cannot be due to a simple failure to characterize optionality, since previous authors have all agreed that *-dır* is optional in other paradigms. The paradigms given in Tables 3 and 4 even have *-dır* as an obligatory marker in some cells. This would run contrary to the other patterns seen in Table 5, where *-dır* is either optional or *cannot mark* the verb. I propose that the reason for this proliferation lies in a failure to properly characterize the relationship between the perfect and the evidential in Azerbaijani as distinct but related categories.

## 2.2. Perfect *-miş* and evidential *=(i)mış*

Johanson (2000: 80) analyzes the evidential form as an enclitic copular evidential particle *=(i)mış*, which has harmonizing and non-harmonizing variants, the former being formally identical to the perfect suffix *-miş*. According to Johanson, the evidential *=(i)mış* also differs from the perfect *-miş* in that the former does not carry pitch accent. This would seem to suggest that it is distinguishable from the perfect *-miş* in principle. These observations do not seem to be true in a straightforward way when there is only one *mış* on a verb stem, though it is the case that *-miş* carries stress rather than *=(i)mış* when both occur, as we will later see. (3) shows an example of potentially ambiguous *-miş/=(i)mış*.

- (3) The evidential  $=(i)mI\dot{s}$  in Azerbaijani (Şirəliyev & Sevortjan: 127; Kazimov 2010: 244).
- a. yaz.**mıŝ**-am  
write.M1ŝ-1SG  
SENSE 1 (default sense): ‘I have written.’  
SENSE 2 (alternative): ‘I wrote/write/have written, apparently (they say, evidently).’
  - b. yaxşı=**ymıŝ**(\*-dir) / yaxşı=**imıŝ**(\*-dir)  
good=EVD / good=EVD  
<sup>OK</sup>‘This is/was/has been good, apparently (they say, evidently).’  
\*‘This has been good.’

While  $-mI\dot{s}$  is possibly (for some speakers, at least) ambiguous when attached to a verb stem as in (3a), it is never ambiguous when attached to adjectives.

A careful evaluation of the morphological evidence seems to suggest that the third person suffix that  $-dir$  displays split behavior. As seen in (3b),  $-dir$  is categorically ungrammatical when an adjectival stem is marked by the evidential. The judgements in (4) mirror the pattern seen in (3), which suggests that the distribution of  $-dir$  in Table 4 is correct.

- (4) Evidential  $=(i)mI\dot{s}$  vs. perfect  $-mI\dot{s}\sim(y)Ib$
- a. ?/\*oxu.**mıŝ**  
read.M1ŝ  
INTENDED 1: ‘(s/he) has read.’ (=perfect) ☒  
INTENDED 2: ‘I wrote/write/have written, apparently.’ (=evidential) ☒
  - b. oxu.**mıŝ**-dur  
read.M1ŝ-3SG  
INTENDED 1: perfect ☒  
INTENDED 2: evidential
  - c. oxu-**yıb**  
read-1B  
INTENDED 1: perfect ☒  
INTENDED 2: evidential
  - d. oxu-**yıb**-dur  
read-1B-3SG  
INTENDED 1: perfect ☒  
INTENDED 2: evidential

Verb+ $=(i)mI\dot{s}$  constructions in the third person—without  $-dir$ , like (4a)—are not acceptable for all speakers; especially bad for those who perceive it as being ‘Turkish’. But for those who do accept it, it uniquely has a stronger evidential reading (as well as a perfect reading, much like Turkish), something *not* found in the other third person perfects (4b-d). Note that this is only a property of the third person. This confirms Johanson’s observation that the Azerbaijani system tends towards pure perfect readings, unlike Turkish. As Johanson (1971: 64, 2000: 80) has pointed out,  $=(i)mI\dot{s}$  is temporally indifferent across Turkic languages, where as  $-mI\dot{s}$  tends to have a temporal sense.

Even if  $-mI\dot{s}$  as an exponent of the perfect is distinguishable from the evidential  $=(i)mI\dot{s}$ , we have not yet considered the meaning(s) of  $-mI\dot{s}$  relative to  $-(y)Ib$ . The Turkish cognate suffix  $-mI\dot{s}$  is also sometimes described as a perfect suffix, but as Bowler and Ozkan (2018) have argued, Turkish  $-mI\dot{s}$  does not actually contribute English-type perfect readings. It has been observed that the English present perfect imposes a ‘lifetime effect’, such that the individuals in the

utterance must be alive at utterance time (Leech 1969; Chomsky 1972: 111-3; Anderson 1973). Bowler and Ozkan (2018: 2, 5) report that the same effect does not apply to Turkish *-mİş*.

(5) Lifetime effects in English and Turkish (Bowler and Ozkan 2018)

a. **English**

*Context:* Einstein is dead, but you have just seen his signature in the physics department guestbook at Princeton.

??Einstein has visited Princeton.

b. **Turkish -mİş**

*Context:* Einstein is dead, but you have just seen his signature in the physics department guestbook at Princeton.

Einstein Princeton-1 ziyaret et-miş

Einstein Princeton-ACC visit do-MİŞ

‘(I have indirect evidence that) Einstein visited Princeton.’

Lifetime effects hold in English, but not in Turkish. Just as in (4), the Azerbaijani perfect exponents *-mİş* and *-(y)İb* pattern together in (6). In this case, the facts support the traditional analyses, which treat both suffixes as synonymous markers of perfect aspect.

(6) Lifetime effects in Azerbaijani

*Context:* Charlie Chaplin is dead, but you have just read that he had been in Japan in 1932.

- |    |           |             |         |                    |
|----|-----------|-------------|---------|--------------------|
| a. | */?Çaplin | Yaponiya-nı | ziyarət | et- <b>miş</b>     |
|    | Chaplin   | Japan-ACC   | visit   | do-MİŞ             |
| b. | ?#Çaplin  | Yaponiya-nı | ziyarət | et- <b>miş-dir</b> |
|    | Chaplin   | Japan-ACC   | visit   | do-MİŞ-3SG         |
| c. | ?#Çaplin  | Yaponiya-nı | ziyarət | ed- <b>ib</b>      |
|    | Chaplin   | Japan-ACC   | visit   | do-MİŞ-3SG         |
| d. | ?#Çaplin  | Yaponiya-nı | ziyarət | ed- <b>ib-dir</b>  |
|    | Chaplin   | Japan-ACC   | visit   | do-MİŞ-3SG         |
- ‘Chaplin has visited Japan.’

(6b-d) are infelicitous for the majority of my consultants, though they were accepted by four of my consultants. (6a) was highly unacceptable for most consultants, with the exception a couple people who noted that—just as in (4)—this form is sometimes acceptable only due to the influence of Turkish. We might consider the general unacceptability of (6) to be the result of a violation of the implicature that the individuals in the utterance exist, similar to the stronger implicature in English, but not similar to Turkish. Those speakers who do accept (6a) do not necessarily accept (6b-d). This is why the present perfect—unlike the other TAM paradigms in Table 5—is exceptionally obligatorily marked by *-dİr* when it is realized by *-mİş*, but not when it is realized by *-(y)İb*.

### 3. Reorganization in the Azerbaijani present perfect

We are now in a better position to describe the asymmetrical distribution of perfect aspect cell-mates along the category of person in Azerbaijani:

**First person (singular and plural):** *-mİş*

**Second person (singular and plural):** *-mİş~(y)İb*

**Third person (singular and plural):** *-mİşdİr~(y)İb~(y)İbdİr*



Unlike the other persons, the third person singular and plural share a segmentable suffix in the realization of their person exponence in the perfect. Also unlike other persons, this suffix is optional with *-(y)Ib*, but obligatory with *-mIş*. Zero marked third person *-mIş* forms do exist for some speakers, but they always necessarily carry evidential readings, rather than perfect readings. By contrast, the first and second persons in the present perfect do not have any reliable strategy to signal a difference in evidential vs. perfect readings of *-mIş*, and they default to temporal readings in verbs. I have suggested that this asymmetry might have facilitated the restricted distribution of evidentiality by allowing for ambiguity in the first and second persons, but not in the third person. What about *-(y)Ib*? Perhaps this suffix never developed a full paradigm for all persons. After all, it has been in competition with *-mIş* so there is no functional need for a full paradigm. A closer look at historical materials provides evidence against this hypothesis.

There is a rather well-established grammaticalization pathway which led to existence of competing forms in the perfect paradigm. I assume, following Doerfer (1977), that modern Azerbaijani is a continuation of the Seljuk language used in Anatolia prior to the 15<sup>th</sup> century, which then diverged into Old Ottoman and what we may call Old Azerbaijani. The suffix *-(y)Ib-dIr* was also present in Ottoman and can be traced to the periphrastic construction *\*X-b tur-ur* in both Ottoman and Azerbaijani. The third person copular suffix *-dIr* comes from the lexical verb *tur-* ‘stand’, which was originally used to periphrastically express the perfect with a converbial construction. The semantic change from changed ‘stand’>‘dwell’>‘be’ (Johanson 2000) likely coincided with the phonological reduction of *durur* to *-dur*, which had become voice-initial and then became a third person suffix with regular vowel harmony on analogy with other person suffixes (Mansuroğlu 1953: 349). The *X-b* is a verb stem plus the non-finite converb suffix *-(I)b*. The grammaticalization of *-(y)Ib* as separable exponent of the perfect in its own right in Azerbaijani can be dated to the 15<sup>th</sup> century (Tanrıverdi 2017: 301). Both the converb suffix *-(y)Ib* (homophonous and cognate with the perfect marker) and the lexical verb *-dur-* ‘stand’ exist in the modern language, but their combined use as a perfect is non-standard, and will not be considered here. *-mIş* is of older origin, and may have expanded from perfect to indirective readings once the *-DI* past tense ceased to mark the witnessed/non-witnessed distinction in (Proto-)Oghuz (Tenishev 2002: 194).

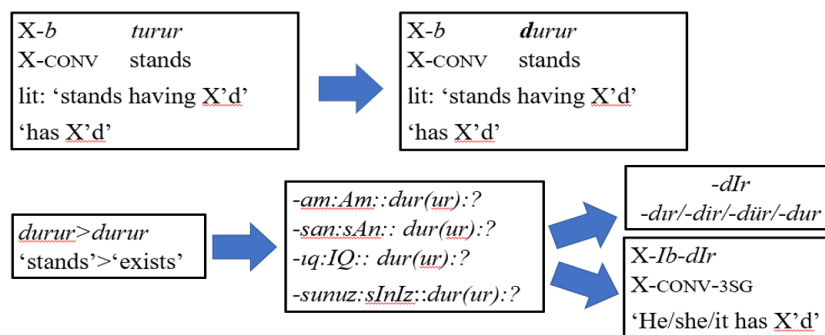


Figure 1. The grammaticalization of *-(y)Ib-dIr* to *-(y)Ib* and *-dIr*.

Given the observations that (i) *-(y)Ib* does not take first person forms; (ii) the second person forms are less acceptable for some speakers; and (iii) the use of *-(y)Ib* as a perfect suffix was historically grammaticalized from a construction involving *durur*, which itself grammaticalized

as a way to mark the third person, perhaps-(y)Ib never developed a full paradigm for all persons, as suggested at the end of the previous section. But consider (7).

- (7) First person -(y)Ib (Füzuli's Bəngü Badə, XVI cent.)  
 نه گناه ائيلهديم كي خوار اولوبام؟  
 Nə günah elə-ydim ki, xar ol-**ub**-am?  
 What sin do-PST.1SG that, shame be-PRF-1SG  
 'What sin have I committed, that I have brought shame upon myself?'

Füzuli is a towering figure in the early Azerbaijani literary canon. First person -(y)Ib forms are found in his writings, but more generally also in the writing of other authors before the 17<sup>th</sup> century. Moreover, Kazımov (2010) reports that -(y)Ib can be marked for first and second persons, specifically in the 'evidential of the perfect' forms, to which we now turn our attention.

#### 4. Person marking in 'composite' perfects

So far I have only considered the present perfect. One complication in the morphology is seen in the evidential of the perfect, as seen in (8).

- (8) The evidential of the perfect (Əfəndiyeva 2005: 49)  
 ...sonra yad-ım-a düş-dü ki, bu əsər-lər  
 ...then memory-1SG.POSS-DAT fall-PST that, DEM work-P  
 haqqında haradasa oxu-**muş**=**muş**-am.  
 about somewhere read-PRF=EVD-1SG  
 '...then I remembered that I **had evidently** read about these works somewhere.'

Verbforms which morphologically express multiple TAM categories in Azerbaijani are traditionally referred to as 'composite tenses' or 'complex tenses.' E.g. for the complex perfects, Şirəliyev & Sevortyan (1971: 127) call them 'complex/composite forms of the past tense tense' [сложные/составные формы прошедшего времени]; Kazımov (2010: 173-175) groups them with 'complex verbs' [Mürəkkəb fellər]; Əfəndiyeva calls them 'complex forms of the past tense' [Keçmiş zaman kateqoriyasının mürəkkəb formaları], etc.

In the case of the evidential of the perfect in (8), Əfəndiyeva points out that it is the inner suffix which is the perfect and identifies the outer suffix as =(i)mIş. The evidential of the -mIş perfect has two possible suffix orders (ibid.: 241). One ordering is adjacent, as seen in Table 6. In the non-adjacent ordering in Table 7, the -mIş perfect is separated from =(i)mIş by the person and number markers.

	Singular	Plural
1	yaz <b>mış</b> muşam~yaz <b>mış</b> imişəm	yaz <b>mış</b> muşuq~yaz <b>mış</b> imişik
2	yaz <b>mış</b> mi(ş)san~yaz <b>mış</b> imi(ş)sən	yaz <b>mış</b> mi(ş)sınız~yaz <b>mış</b> imi(ş)siniz
3	yaz <b>mış</b> muş~yaz <b>mış</b> imiş	yaz <b>mış</b> muşlar~yaz <b>mış</b> imişlər

Table 6: Adjacent ordering of the evidential of the of -mIş perfect of yaz- 'write' (Pitch accent added)

	Singular	Plural
1	yaz <b>mış</b> ammuş~yaz <b>mış</b> am imiş	yaz <b>mış</b> ıqmuş~yaz <b>mış</b> ıq imiş
2	yaz <b>mı</b> (ş)sanmuş~yaz <b>mı</b> (ş)san imiş	yaz <b>mı</b> (ş)sınızmuş~yaz <b>mı</b> (ş)sınız imiş
3	yaz <b>mış</b> muş~yaz <b>mış</b> imiş (same as table 6)	yaz <b>mış</b> larmuş~yaz <b>mış</b> lar imiş

Table 7: Non-adjacent ordering of the evidential of the of -mIş perfect of yaz- 'write' (Pitch accent added)

Notably, the third person forms here are marked by *-dlr*. As we have already seen, the third person suffix *-dlr* has strictly temporal, non-evidential readings. The evidential of the perfect reportedly has a strongly evidential reading, which explains why it is never marked by *-dlr*. The harmonizing and non-harmonizing variants of the adjacent and non-adjacent orderings shown in Tables 6 and 7 are all synonymous with each other and with the evidential of the *-(y)Ib* perfect, seen in Table 8. Unlike present perfect *-(y)Ib* perfect forms, the first person is permissible in the evidential of the perfect (ibid.: 242).

	Singular	Plural
1	<i>yazıbmışam~yazıb imişəm</i>	<i>yazıbmışıq~yazıb imişik</i>
2	<i>yazıbmı(ş)san~yazıb imi(ş)sən</i>	<i>yazıbmı(ş)sınız~yazıb imi(ş)siniz</i>
3	<i>yazıbmış~yazıb imiş</i>	<i>yazıbmışlar~yazıb imişlər</i>

Table 8: The evidential of the *-(y)Ib* perfect of *yaz-* ‘write’ (Pitch accent added)

The six cell-mates of the evidential of the perfect are a striking example of overabundance across all persons. To say ‘I have apparently written’, one could say any of *yazmışmışam / yazmış imişəm / yazmışam / yazmışam imiş / yazıbmışam / yazıb imişəm*. That is not to say that all these forms are universally used. Şirəliyev (2008: 270) notes that the forms in Table 6 are more common in the Western dialects of dialects Karabakh, Qazakh, and Ağdam. That being said, it is not clear what would allow speakers to choose between the forms in Table 7 and Table 8; my consultants accept all forms in Tables 6-8, and these forms are sometimes listed in standard grammars.

Presumably, one could also add another table of forms in which person marking follows *-(y)Ib* and precedes *=(i)mIş*, but I have not seen this possibility mentioned in the literature. The forms are, however, attested. Consider the following examples, pulled from the Sketch Engine Turkic Web corpus of Azerbaijani (Baisa and Suchomel 2015), a large (~115,000,000 tokens) text corpus of Azerbaijani government and news websites, as well as some blogs and other .az domain websites.

- (9) Bax-ıb gör-dü-k ki, müqavilə-ni bir ay  
 look-CONV see-PST.1P that, contract-ACC one month  
 müddət-i-ndə bağla-yıb-lar=**mış**  
 period-EZ-LOC close-PRF-3P=**EVD**  
 ‘We looked and saw that they have/had evidently/apparently conducted/closed the contract in a month.’
- (10) Valideyn-lər-i o-na həтта İslam-ı da  
 parent-3P-POSS.3 3SG-DAT even Islam-ACC also  
 qəbul elə-t-dir-ıb-lər=**mış**  
 accept do-CAUS-CAUS-PRF-3P=**EVD**  
 ‘His/her parents have/had evidentially even made him/her accept Islam.’

We clearly see in (9) and (10) that the person/number marker *-lAr* intervenes between the perfect marker *-(y)Ib* and the evidential *=(i)mIş*, so this order must be possible, though unfortunately no other person markers besides *-lAr* occur in the SketchEngine corpus. We do, however, find other person/number markers in this position in, e.g., the historical writings of some authors. For example, in the 19<sup>th</sup> century playwright Axundov’s *Sərgüzəşti-mərdi-xəsis: Hacı Qara* [The Adventure of the Miser: Haji Gara] the *naçalnik* character says the sentence in (11) to the titular Haji Gara:

- (11) Bəli, çox böyük qulluq-lar ed-**ib-sən=miş!**  
 yes, very great service-3P do-**PRF-2SG=EVD**  
 ‘Yes, you have/had evidentially done a great service!’

Speakers seem to still accept the form in (11). They are less sure about the first-person forms when they are adjacent to *-(y)Ib*, as indicated by the question marks in Table 9, but the judgments here are not as degraded as the first-person marked *-(y)Ib* simple perfect forms.

	Singular	Plural
1	? <i>yazıbammiş~yazıbam imiş</i>	? <i>yazıbıqmış~yazıbıq imiş</i>
2	<i>yazıbsanmış ~yazıbsan imiş</i>	<i>yazıbsınızmış ~yazıbsınız imiş</i>
3	<i>yazıbmış~yazıb imiş</i>	<i>yazıblarmış~yazıblar imiş</i>

Table 9: Non-adjacent ordering of the evidential of the *-(y)Ib* perfect of *yaz-* ‘write’ (Pitch accent added)

As the historical evidence in (7) demonstrates, the *-(y)Ib* marker used to mark the first person even in the first person, but it has become specialized for non-first persons, and—as I suggest in Zaslansky (submitted)—is moving towards becoming a marker of the third person, i.e. it is declining in relative frequency as a marker in the second person present perfect. The composite ‘evidential of the perfect’ forms have not been subject to this apparent shift in person marking. How can we account for the facts of first person marking in this section and in the previous section?

First, it is worth considering that the evidentials of the perfect are rarer than the simple present perfect. Thornton (2012) has suggested that overabundance in paradigms is better preserved in low frequency cells than in high-frequency cells. Cappellaro (2018) reports similar findings. One could, for example, look across a large enough modern corpus—and across the writings of several historical authors—and scarcely encounter any of the forms in Tables 6-9. Consider the results in Table 10.

Author	Nəmətullah Kışvəri	Xətai	Füzuli	Vaqif	Axundov	Şirvani	--
Suff. / Period	(1490-1502)	(1487-1524)	(1494-1556)	(1717-1797)	(1812-1878)	(1835-1888)	Modern
<i>mİş=mİş-X</i>	0	0	0	0	6	0	0.06
<i>mİş-X=mİş</i>	0	0	9.39	0	102.07	6.12	0.4
<i>Ib=mİş-X</i>	26.94	0	0	0	0	6.12	30.33
<i>Ib-X=mİş</i>	0	0	0	0	17.01	24.49	3.12

Table 10: Normalized relative frequencies (per million) for the evidential of the perfect

The frequencies in Table 10 come from corpora of drastically different sizes and so are normalized by dividing the observed frequencies by the total number of tokens in the corpus, then multiplied by 1,000,000 for consistent comparison. This normalization is admittedly arbitrarily done, but is easily interpretable (see e.g. Kilgarriff 2009 for discussion of normalization). The first row of results in Table 10 represents the frequencies for the forms in Table 6, the second row for forms in Table 7, the third row for forms in Table 8, and the fourth row for forms in Table 9. These results are across all lexemes, not just *yaz-* ‘write,’ and include all six possible person and number combinations. It goes without saying that these forms are exceedingly rare. For context, compare the frequency of the simple past tense marker in Table 11.

Author	Nəmətullah Kışvəri	Xətai	Füzuli	Vaqif	Axundov	Şirvani	--
Suff. / Period	(1490-1502)	(1487-1524)	(1494-1556)	(1717-1797)	(1812-1878)	(1835-1888)	Modern
<i>dI</i>	36209.82	45123.27	50789.51	32930.42	30690.19	63122.64	20116.48

Table 11: Normalized relative frequencies (per million) for the simple past

This is consistent with the findings of Thronton (2012) and Cappellaro (2018) that low frequency forms preserve variation in the paradigm. Interestingly, the variation here is *throughout* categories in the paradigm (i.e. first-person evidential of perfect, second-person evidential of perfect) rather than at the level of the cell (e.g. first-person singular evidential of the perfect). The distribution of markers observed so far is summarized in Table 12.

	Present perfect	Evidential of perfect
First person (singular and plural):	<i>-mİş</i>	<i>-mİş~(y)İb</i>
Second person (singular and plural):	<i>-mİş~(y)İb</i>	<i>-mİş~(y)İb</i>
Third person (singular and plural):	<i>-mİşdlr~(y)İb~(y)İbdlr</i>	<i>-mİş~(y)İb</i>

Table 12: Distribution of perfect exponence by person

### 5. Differentiation by person in the (present) perfect and evidential

Given that speakers tend to give more degraded judgements when the second person present perfect is marked with *-(y)İb* and that *-mİşdlr* and *-(y)İb* seem to be the preferred pattern for the third person present perfect, it seems that the perfect markers are moving towards differentiation by person in the present perfect but not in the evidential of the perfect. What would the cause behind this apparent distribution be?

I believe that that the bare evidential marker is relevant here. The evidential marker by itself belongs to entirely different syntagmatic distribution which just happens to be homophonous with the perfect marker in certain contexts when it is marked on the verb. As we saw in (4) and (6), forms marked by *mİş*—if they are ever truly ambiguous between evidential and perfect readings—tend towards the temporal perfect reading. This corroborates previous claims by Johanson (1998b, 2002: 147). But my consultants really only seem to have strong evidential readings in the third person (when they do not reject it for sounding ‘Turkish’), and only rarely in the first and second persons, as indicated by the question marks in Table 13.

	Base evidential	Evidential of perfect
First person (singular and plural):	?=(y)mİş~imiş	=(y)mİş~imiş
Second person (singular and plural):	?=(y)mİş~imiş	=(y)mİş~imiş
Third person (singular and plural):	=(y)mİş~imiş	=(y)mİş~imiş

Table 13: Distribution of evidential exponence by person

Interestingly, just as with the *-(y)İb* perfect marker, the evidential *=(i)mİş* is more acceptable as a marker of evidentiality in the first and second persons in the composite evidential of perfect category! It is not obvious that there should be something pragmatically odd about the base evidential first or second person as opposed to the composite evidential of perfect first or second person. One hypothesis is there is a cross-paradigmatic contrast here.

The *-dlr* marker—which is typically optional where it does occur (as in Table 5)—is not optional in the perfect when marked by *-mİş* because it functions as an additional marker of contrast with the evidential *=(y)mİş*. The third person, at least for some speakers, thus clearly allowed for differentiation of perfect and evidential readings in a way that the first and second persons did not. I cannot work out the details here, but I speculate that this contrast only in the

third person has led to—at least for some speakers—a more restricted distribution of evidential readings in the first and second persons, where there was not typically any morphological contrast between evidential and perfect readings of *mİş*. This creates a contrast in person in the perfect category, which—speculatively—seems to have led to the differentiation of the *-(y)İb* marker by person as well. As Əfəndiyeva reminds us, this concept of sensitivity to personhood goes back to the Ottoman scholar Tserunian’s ‘Differentiation by Person’, something he had identified in Ottoman Turkish, in which only in the third person was there a contrast between evidential and perfect forms of *-mİş*, e.g. Ottoman *al-mİş-tur* ‘(S)he has bought’ contrasted with *al-mİş* ‘(S)he has bought’ (Dmitriyev 1927).

Differentiation by person seems to be a reliable trend in Azerbaijani. What would the alternative be?

### 6. Concluding remarks: From systematic overabundance to systematic differentiation

It is in principle possible that rival affixes would become associated with different lexemes (Kroch 1994; Thornton 2011; Aronoff & Lindsay 2015). For example, English verbs may be grouped into inflectional classes based on their patterns of exponence in the past and past participle (following Carstairs-McCarthy 1994). Thus some verbs like *heal* express the past and past participle with the voiced *-(e)d* suffix, and other verbs—like *feel*—express the same categories via the voiceless *-t* and ablaut.

Categories	‘heal’ Class	‘feel’ Class	‘give’ Class	‘kneel’ Class	‘burn’ Class
Past	heal(e)d	felt	gave	knelt / kneeled	burned/burnt
Past participle	heal(e)d	felt	giv(e)n	?knelt	burned/burnt
Imperative	heal (Ø)	feel (Ø)	give (Ø)	kneel (Ø)	burn (Ø)

Table 14: English verb classes (Carstairs-McCarthy 1994; Levin 2009)

Other verbs may be marked by both patterns in the past (preterite) but perhaps not in the past participle—e.g. *kneel*—and other verbs may be marked by both markers patterns in both the past preterite and the past participle forms, e.g. *burn* (Levin 2009). While verbs like *kneel* and *burn* may exhibit variation, there seem to be clearly distinct *heal* and *feel* classes which take one pattern or the other.

The same sort of class formation is also true of the German nouns reported in Fehringer (2004), which exhibit different patterns of genitive singular formation.

Categories	-s Class	-es Class	Variable Class	Feminine Class
NomSg	Deutsch	Busch	Vaterland	Hand
GenSg	Deutchs	Busches	Vaterlands/Vaterlandes	Hand (Ø)
Translation	German	bush	fatherland	hand

Table 15: English verb classes (Carstairs-McCarthy 1994; Levin 2009)

Feminine nouns like *Hand* do not overtly mark the genitive singular, but strong masculine and neuter nouns mark the genitive singular with *-s* or *-es*, depending on which class the noun falls into. For example, nouns denoting a language, like *Deutsch*, fall into the *-s* class, whereas most nouns ending in <sch>, like *Busch*, fall into the *-es* class. Fehringer describes a large class of nouns like *Vaterland*, which may be marked by either suffix. While the *Vaterland* class is quite large, it is nonetheless the case that there are several nouns which fall clearly into either the *-s* class or the *-es* class, so the competing markers in the *Vaterland* class may be described as belonging to different classes, much like the doublets *burned* and *burnt* may be described as belonging to different classes.

We do not find any indication of sort of class formation in the Azerbaijani perfect. Consider, for example, that the third person displays the most variability in the present perfect, being marked by one of *-(y)Ib*, *-(y)IbdIr*, or *-mİşdir*. The most common verbs marked by these suffixes in the third person largely overlap in the SketchEngine corpus, as seen in Table 16.

Rank	<i>-(y)Ib</i>			<i>-(y)IbdIr</i>			<i>-mİşdir</i>		
	Verb	Token frequency	Per million	Verb	Token frequency	Per million	Verb	Token frequency	Per million
1.	<i>et-</i> 'do'	113,204	981.99	<i>ol-</i> 'be'	579	5.02	<i>et-</i> 'do'	28,681	248.79
2.	<i>ol-</i> 'be'	79,554	690.09	<i>et-</i> 'do'	452	3.92	<i>ol-</i> 'be'	25,966	225.24
3.	<i>bildir-</i> 'inform'	52,691	457.07	<i>edil-</i> 'do' (PASS)	335	2.91	<i>edil-</i> 'do' (PASS)	20,278	175.90
4.	<i>ver-</i> 'give'	43,234	375.03	<i>olun-</i> 'be' (REFL)	333	2.89	<i>olun-</i> 'be' (REFL)	15,365	133.28
5.	<i>olun-</i> 'be' (REFL)	41,337	358.58	<i>veril-</i> 'give' (PASS)	188	1.63	<i>keçiril-</i> 'pass' (PASS)	11,348	98.44
6.	<i>edil-</i> 'do' (PASS)	38,255	331.84	<i>yaradıl-</i> 'create' (PASS)	167	1.45	<i>veril-</i> 'give' (PASS)	8,032	69.67
7.	<i>de-</i> 'say'	29,970	259.97	<i>ver-</i> 'give'	154	1.34	<i>veril-</i> 'give' (PASS)	7,562	65.60
8.	<i>keçiril-</i> 'pass' (PASS)	20,552	178.28	<i>çevril-</i> 'turn' (intr.)	146	1.27	<i>bildir-</i> 'inform'	4,820	41.81
9.	<i>veril-</i> 'give' (PASS)	19,269	167.15	<i>art-</i> 'increase'	127	1.10	<i>yaradıl-</i> 'create' (PASS)	4,510	39.12
10.	<i>başla-</i> 'start'	16,345	141.78	<i>çat-</i> 'reach'	112	0.97	<i>başla-</i> 'start'	3,513	30.47
11.	<i>vurğula-</i> 'stress'	13,613	118.09	<i>görül-</i> 'see' (PASS)	107	0.93	<i>al-</i> 'take'	3,302	28.64
12.	<i>al-</i> 'take'	13,239	114.84	<i>tikil-</i> 'build' (PASS)	99	0.86	<i>görül-</i> 'see' (PASS)	2,969	25.75
13.	<i>gəl-</i> 'come'	10,556	91.57	<i>yanan-</i> 'create' (intr.)	83	0.72	<i>aparıl-</i> 'carry' (PASS)	2,885	25.03
14.	<i>tutul-</i> 'hold' (PASS)	9,468	82.13	<i>al-</i> 'take'	81	0.70	<i>tutul-</i> 'hold' (PASS)	2,658	23.06
15.	<i>alın-</i> 'take' (PASS)	9,347	81.08	<i>keçiril-</i> 'pass' (PASS)	80	0.69	<i>yarat-</i> 'create' (tr.)	2,652	23.00

Table 16: Most frequent third-person present perfect verbs by suffix

Since the perfect *-(y)Ib* is homophonous with the converb *-(y)Ib*, all of the results in Table 16 only reflect verbs followed directly by a period (i.e. the final words in declarative sentences), ruling out the possibility of including any converbs.<sup>4</sup> Generally, the exact same verbs are the most frequent ones across the three suffixes. Some verbs which only appear in the top 15 for one suffix nonetheless occur with the other suffixes as well, as seen in Table 17.

<sup>4</sup> I have also excluded common non-verbs which happen to end in <ub>, <ib>, <ıb>, and <üb>, which did appear in the search results, e.g. *sahib* 'owner', *vacib* 'important', *münasib* 'reasonable', *ayıb* 'shame', etc.

Verb	-(y)Ib	-(y)IbdIr	-mIşdIr	Ratio (rounded)
<i>de-</i> 'say'	29,970	25	992	<b>1200:1:40</b>
<i>vurğula-</i> 'stress'	13,613	34	2,366	<b>2720:7:480</b>
<i>gəl-</i> 'come'	10,556	38	1,790	<b>265:1:45</b>
<i>çevril-</i> 'turn' (intr.)	5,189	146	2,489	<b>52:1:25</b>
<i>art-</i> 'increase'	6,319	127	2,303	<b>63:1:23</b>
<i>çat-</i> 'reach'	7,853	112	2,336	<b>79:1:23</b>
<i>tikil-</i> 'build' (PASS)	1,426	99	827	<b>14:1:8</b>
<i>yarat-</i> 'create' (tr.)	4,392	56	2,652	<b>44:1:27</b>

Table 17: Relative frequencies for verb stems occurring in the top 15 only once in Table 16

As Table 17 demonstrates, the relative frequencies represented by the ratios (rounded to the nearest hundreds) are more or less constant across all of the verb stems which occur in only one of the top 15 rankings in Table 16: *-(y)Ib* is the most frequent marker, followed by *-mIşdIr*, followed by the relatively rare *-(y)IbdIr*. There is no class formation here. Instead we find systematic differentiation by person.

What we do have is differentiation within the paradigm along the category of person, as demonstrated by the frequency measures in Table 18 (again, using only the final words in declarative sentences).

	-(y)Ib	-mIş		-(y)Ib	-mIş
1SG (-Am)	1	<b>30,617</b>	1P (-IQ)	0	<b>18,746</b>
2SG (-sAn)	93	<b>1,628</b>	2P (-sInIz)	134	<b>3,451</b>
3SG (-dIr)	<b>1,020,837</b>	277,080	3P (-dIr)lAr)	<b>111,376</b>	35,059

Table 18: Raw token frequencies for by sentence-final present perfect forms by person

The SketchEngine-derived raw token frequencies in Table 18 reflect—once again—the final verbs in declarative sentences. The third person forms with *-(y)Ib* include forms both with and without the overt *-dIr* marker, as do all of the third person plural results for both markers, since the *-lAr* marker may occur both with and without *-dIr*. I manually removed any non-verbs found in the 100 most frequent types for each person and number combination in an effort to ensure that the measures above faithfully reflect the distribution of present perfect markers by person. The single first person occurrence with *-(y)Ib* is actually a quote from Füzuli, who—as we saw in (7)—lived at a time when *-(y)Ib* was still used in the first person. What we have here is a clear paradigmatic contrast in person.

Paradigmatic contrasts of the sort are found elsewhere in the language. An admittedly speculative but sensible hypothesis is that when there is variation in inflectional markers, there is a systematic tendency in Azerbaijani for these markers to become differentiated by person. For



example, the suffixes of the negative aorist category are differentiated by person, exhibiting a contrast between the first person and non-first persons (Fəxrəddinqızı 2010: 119).

	Affirmative aorist	Negative aorist
First person (singular and plural):	-Ar	-mAr
Second person (singular and plural):	-Ar	-mAz
Third person (singular and plural):	-Ar	-mAz

Table 19: Distribution of aorist suffixes by person

The situation in Old Anatolian Turkish reportedly remarked all negative aorist forms with *-mAr* (Fəxrəddinqızı 2010: 121), but after a (later?) period of competition between *-mAr* and *-mAz*, the latter came to be associated with the first person. Similarly, several of the Azerbaijani dialects in Şirəliyev's (2008) survey seem to have come to associate *-(y)Ib* or *-mIş* with various different persons in the perfect.

Systematic overabundance (Bonami & Stump 2016) therefore exhibits systematic differentiation rather than class-based differentiation, at least in Azerbaijani.

### Abbreviations

- 1 = First person
- 2 = Second person
- 3 = Third person
- ACC = Accusative
- CAUS = Causative
- CONV = Converb
- DAT = Dative
- EVD = Evidential
- EZ = Ezafe
- GEN = Genitive
- intr. = Intransitive
- LOC = Locative
- NOM = Nominative
- P = Plural
- PASS = Passive
- POSS = Possessee
- PRF = Perfect
- PST = Past
- REFL = Reflexive
- SG = Singular
- TAM = Tense, Aspect, Mood
- tr. = Transitive

### References

- Əfəndiyeva, Türkan. 2005. *Felin keçmiş zaman formaları* [Forms of the Past Tense of the Verb]. Baku: Nurlan. Anderson,
- Anderson, John. 1973. On Existence and the Perfect. *Foundations of Language* 10(2). 333-337.
- Aronoff, Mark & Mark Lindsay. 2015. Partial organization in languages: la langue est un système où la plupart se tient. In Sandra Augendre, Graziella Couasnon-Torlois, Déborah

- Lebon, Clément Michard, Gilles Boyé & Fabio Montermini (eds.), *Proceedings of the 8th Décembrettes*, 1–14. Toulouse: CNRS & Université Toulouse – Jean Jaurès.
- Axundov/Axundzadə, Mirzə Fətəli. 1852. *Sərgüzaştı-mərdi-xəsis: Hacı Qara* [The Adventure of the Miser: Haji Qara]. Available via [Wikisource.org](https://www.wikisource.org).
- Baisa, Vít & Vít Suchomel. 2015. Turkish language support in Sketch Engine. In *Proceedings of the 3rd International Conference on Computer Processing in Turkic Languages (TurkLang 2015)*, 214–223.
- Bonami, Olivier & Gregory T. Stump. 2016. Paradigm Function Morphology. In Andrew Hippisley & Gregory T. Stump (eds.) *The Cambridge Handbook of Morphology*, 449-481. Cambridge: Cambridge University Press.
- Bowler, Margit & Sozen Ozkan. 2018. Turkish –*mİş* does not contribute (English-type) present perfect. In Céleste Guillemot, Tomoyuki Yoshida & Seunghun J. Lee (eds.) *Proceedings of WAFL 13*.
- Carstairs-McCarthy, Andrew. 1994. Inflection classes, gender, and the principle of contrast. *Language* 70.737–88.
- Cappellaro, Chiara. 2018. Genesis and diachronic persistence of overabundance. In Antje Dammel, Matthias Eitelmann, & Mirjam Schmuck (eds.) *Reorganising Grammatical Variation: Diachronic studies in the retention, redistribution and refunctionalisation of linguistic variants*. Amsterdam: John Benjamins Publishing Company. 119-148. <https://doi.org/10.1075/slcs.203.05cap>.
- Chomsky, Noam. 1972. *Deep Structure, Surface Structure and Semantic Interpretation*. The Hague: Mouton.
- Dmitriyev, Nikolay. 1927. K voprosu o znaceniі osmanskoj glagol’noj formy na “mysh” [On the question of the meaning of the Ottoman verbal form “mış”]. *Zapiski Kollegii Vostokovedov*, 2 [Notes from the College of Orientalists, 2]. Leningrad: Izdatel’stvo Akademii Nauk SSSR.
- Doerfer, Gerhard. 1977. Das Chorasantürkische. *Türk Dili Araştırmaları Yıllığı-Belleten*. 127-204.
- Fəxrəddinqızı, Gülnarə. 2010. *Türk Dillərində Zaman Kateqoriyası* [The Category of Time in the Turkic Languages]. Baku: Nurlan.
- Fehringer, Carol. 2004. How stable are morphological doublets? A case study of /ə/~/Ø/ variants in Dutch and German. *Journal of Germanic Linguistics* 16(4). 285–329.
- Füzuli, Məhəmməd. 1524/2005. *Bəngü-Badə* [Opium and Wine]. In Həmid Araslı & Teymur Kərimli (eds.) *Məhəmməd Füzuli Əsərləri: İkinci Cild* [The words of Məhəmməd Füzuli: Second Volume]. Baku: Şərq-Qərb.
- Haber, Lyn R. Leaped and Leapt: A Theoretical Account of Linguistic Variation. *Foundations of Language* 14(2). 211-238.
- Hacıyeva [Əfəndiyeva], Türkan. 1958. *Sistema form proshedshego vremeni glagola v sovremennom azerbajdzanskom jazyke*. [The system of of the past-tense forms of the verb in the Modern Azerbaijani Language]. Ph.D dissertation. Moscow.
- Hüseynzadə, Muxtar. 2007. *Müasir Azərbaycan Dili, Üçüncü Hissə: Morfologiya* [Modern Azerbaijani, Part Three: Morphology]. Baku: Şərq-Qərb.
- Householder, Fred, & Mansour Lotfi. 1965. *Basic Course in Azerbaijani*. Bloomington: Indiana University Publications, Uralic and Altaic Series, 4.
- Johanson, Lars. 1971. Aspekt im Türkischen: Vorstudien zu einer Beschreibung des türkeitürkischen Aspektsystems. *Studia Turcica Upsaliensia*. Uppsala: Almqvist & Wiksell.

- . 1998a. The Structure of Turkic. In Lars Johanson & Éva Agnes Csató (eds.) *The Turkic Languages*, 30-66. London & New York: Routledge.
- . 1998b. Code-copying in Irano-Turkic. *Language Sciences* 20. 325-338.
- . 2000. Turkic Indirectives. In Johanson & Utas (eds.) *Evidentials: Turkic, Iranian, and Neighboring Languages*, 61-88. Berlin: Mouton de Gruyter.
- . 2002. *Structural Factors in Turkic Language Contacts*. Richmond & Surrey: Curzon.
- Levin, Magnus. 2009. The formation of the preterite and the past participle. In Günter Rohdenburg and Julia Schlüter (eds.), *One Language, Two Grammars? Differences between British and American English*, 60–85. Cambridge: Cambridge University Press.
- Kazımov, Qəzənfər. 2010. *Müasir Azərbaycan Dili: Morfologiya* [Modern Azerbaijani: Morphology]. Baku: Elm və Təhsil.
- . 2014. *Seçilmiş əsərləri, 9cu cild* [Selected works, Vol 9]. Baku: Elm və Təhsil.
- Kilgarriff, Adam. 2009. Simple maths for keywords. Paper presented at *Corpus Linguistics 2009*, University of Liverpool.
- Kroch, Anthony. 1994. Morphosyntactic variation. In Katharine Beals (ed.) *Papers from the 30th Regional Meeting of the Chicago Linguistics Society: Parasession on Variation and Linguistic Theory*, 180–201. Chicago: Chicago Linguistic Society.
- Lečić, Dario. 2017. *Morphological Doublets in Croatian: A multi-methodological analysis*. PhD Dissertation. University of Sheffield.
- Leech, Geoffrey N. 1969. *Towards a Semantic Description of English*. London: Longmans.
- Mansuroğlu, Mecdud. 1953. Türkçede -miş ekinin fonksiyonları [The functions of the -miş suffix in Turkish]. In *60 doğum yılı münasebetiyle Fuad Köprülü armağanı*, 345-351. İstanbul: Ankara Üniversitesi Dil ve Tarih-Coğrafya Fakültesi.
- Öztopçu, Kurtuluş. 2003. *Elementary Azerbaijani*. Santa Monica, California.
- Şirəliyev, Məmmədəğa Şirəli oğlu. 1958. O dialektnoj osnove azerbajdzanskogo natsional'nogo literaturnogo jazyka [On the dialectal basis of the national literary Azerbaijani language]. *Voprosy jazykoznanija* 1958 1. 78-85.
- . 2008. Dialekty i govory azerbajdzanskogo jazyka [Dialects and subdialects of the Azerbaijani language]. In Anna Dybo *Dialekty tyurkskix jazykov: Ocerki* [Dialects of the Turkic languages: Notes], 36-82. Moskow: Vostocnaja literatura RAN.
- & Ervand Sevortjan. 1971. *Grammatika Azerbajdzanskogo Jazyka: Fonemika, Morfologija i Sintaksis* [Grammar of the Azerbaijani Language: Phonetics, Morphology and Syntax]. Baku: Akademija Nauk Azerbajdzanskoj SSR.
- Tanrıverdi, Əzizxan. 2017. *Azərbaycan Dilinin Tarixi Qrammatikası* [Historical Grammar of Azerbaijani]. Baku: Elm və təhsil.
- Tenishev, Egdem R. 2002. *Sravnitel'no-istoriceskaya grammatika tjurkskix jazykov: Regional'nye rekonstruksii* [Comparative-Historical Grammar of Turkic Languages: Regional Reconstruction]. Mosow: Nauka.
- Thornton, Anna M. 2011. Overabundance (multiple forms realizing the same cell): a non-canonical phenomenon in Italian verb morphology. In Martin Maiden, John Charles Smith, Maria Goldbach & Marc-Olivier Hinzelin (eds.), *Morphological Autonomy: Perspectives from Romance Inflectional Morphology*, 358-381. Oxford: Oxford University Press. [doi.org/10.1093/acprof:oso/9780199589982.003.0017](https://doi.org/10.1093/acprof:oso/9780199589982.003.0017).
- . 2012. Reduction and maintenance of overabundance. A case study on Italian verb paradigms. *Word Structure* 5(2). 183-207. [doi.org/10.3366/word.2012.0026](https://doi.org/10.3366/word.2012.0026).

Zaslansky, Matthew. Submitted to *the Proceedings of the 4th workshop on Turkic and languages in contact with Turkic (Tu+4)*.