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)lb1 and -mls, and claim that -(y)lb can mark second or third persons, freely varying with -mls (Şiraliyev and Sevortyan 1971: 125, Hüseynzad 2007: 151, Faxraddinqua 2010: 73-74, Kazım 2014: 175). The focus of this article is on the competition between the rival perfect suffixes -mls and -(y)lb in Standard Azerbaijani. Both -mls and -(y)lb each have their own predictable sets of allomorphs, largely due to vowel harmony. For example, -mls occurs with a high front vowel /i/ when the stem ends in a front unrounded vowel, as in getmişəm ‘I have gone’, omızdirmişəm ‘I have suckled’; with a front high rounded vowel /u/ when the stem ends in a front rounded vowel, e.g. sönərmüş(ə)şən2

* I am grateful to Lamiya Adışova, Lala 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).
1 I employ Standard Azerbaijani Latin orthography throughout this paper. The orthography tends to correspond to IPA equivalents in broad transcription, except j = /j/, s = /ʃ/, ç = /ç/, c = /çk/, k = /k/, g = /g/, q = /q/ (often spirantized as [x] in codas), z̄ = /ʒ/, y = /j/, ə = /æ/, ə = /æ/, u = /u/, i = /i/. Vowel and consonant harmony spreads from stems to suffixes. Following turkological convention, archiphonemes which undergo harmony are written with capital letters: I (unifies /ɨ/, /ɨ/, /u/, and /ɯ/), A (unifies /æ/ and /ʌ/) and Q (unifies /ɛ/ and /ɬ/).
2 Depending on the speaker’s idiolect, the /ʃ/ in -mls may be deleted/assimilated when followed by either the second person singular –sA, the second person plural –slnz, or the conditional suffix –sA. While such deletion does seem to have been a phonetically motivated process of sibilant haplology at some point in the language, it appears that—for those speakers who do allow for haplology here—the process is not necessarily phonetic or even phonological, since it only happens with -mls. Consider the forms from one speaker, e.g. yaz-mi-san ‘you have writ-
‘you have extinguished’, hürkmüşük ‘we have been startled’, etc.\(^3\) Table 1 shows the typical paradigm for the perfect forms, demonstrated with the verbs al- ‘take’ and get- ‘go’.

<table>
<thead>
<tr>
<th></th>
<th>al- ‘take’</th>
<th>get- ‘go’</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>almişşam</td>
<td>getmişşom</td>
</tr>
<tr>
<td>2SG</td>
<td>almuş(s)an~alhsan</td>
<td>getmuş(s)an~gedibşan</td>
</tr>
<tr>
<td>3SG</td>
<td>almuşdir~alb(dir)</td>
<td>getmuşdir~gedib(dir)</td>
</tr>
<tr>
<td>1P</td>
<td>almişşiq</td>
<td>getmişşik</td>
</tr>
<tr>
<td>2P</td>
<td>almuş(s)ızn~alhsızn</td>
<td>getmuş(s)ızn~gedibsızn</td>
</tr>
<tr>
<td>3P</td>
<td>almuşdılır~alb(dir)lar</td>
<td>getmuşdılır~gedib(dir)lar</td>
</tr>
</tbody>
</table>

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, Fehringer 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 (Dagestan), Eastern Turkey, and Iraq. Descriptions of the competing forms present perfect forms like getmiş(s)an/ gedibşan ‘you have gone’ have typically described them as just-so synonymous inflections, typically without further comment, except to note that -(y)Ib does not mark the first person. As Şiraliyev (1958) has demonstrated, the standard variety contains features primarily from the dialects of Şirvan and Baki-Şamaxi, 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 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 in the present perfect in Azerbaijani applies to all the

\(^3\) /uy/ when the stem ends in a back unrounded vowel, e.g. qalımsdır, etc. -(y)lb shows the exact same patterns, e.g. gedibşon ‘you have gone’, səndirlübson ‘you have extinguished’, with the additional caveat that -(y)lb predictably occurs with a glide when following a vowel, e.g. oxumüşdur ‘(s)he has read’, cf. oxumuşdur ‘(s)he has read’.\(^{10}\) But soruş-san ask-COND.2SG ‘if you ask’ (but not *soru-san); günş-san sun-COP.2SG ‘you are (a/the) sun’ (but not *gün-san). For many speakers, the exceptional hapology in 2\(^{nd}\)-person –mIş is optional, and so may be considered yet another form of overabundance. Pursuing this analysis is beyond the scope of the present paper.

---

\(^{10}\) ‘you have extinguished’, hürkmüşük ‘we have been startled’, etc.\(^3\)
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.

   Könül every day dolma.ACC eat-PRF.MİŞ-3SG / eat-PRF.IB-3SG
   ‘Könül has eaten dolma every day.’
b. Man hər gün dolma ye-miş-am / *ye-yib-am.
    Könül every day dolma.ACC eat-PRF.MİŞ-1SG / eat-PRF.IB-1SG
    ‘I have eaten dolma every day.’

While most traditional descriptions only report that -(y)lb 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(s)-sən / ?ye-yib-sən.
    2SG every day dolma.ACC eat-PRF.MİŞ-2SG / eat-PRF.IB-2SG
    ‘You have eaten dolma every day.’

    2P every day dolma.ACC eat-PRF.MİŞ-2P / eat-PRF.IB-2P
    ‘You guys have eaten dolma every day.’

Speakers still generally accept the -(y)lb 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 Şiraliyev & 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.

<table>
<thead>
<tr>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 yez-miş-am</td>
<td>yez-miş-yq</td>
</tr>
<tr>
<td>2 yez-mi(s)-sən-yez-yən-san</td>
<td>yez-mi(s)-siniz-yez-yən-siniz</td>
</tr>
<tr>
<td>3 yez-miş-(dIr)-yez-yən-(dIr)</td>
<td>yez-miş-(dIr)-lar-yez-yən-(dIr)-lar</td>
</tr>
</tbody>
</table>

Table 2: Present perfect paradigm of yaz- ‘write’ (Şiraliyev 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)lb.

<table>
<thead>
<tr>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 yez-miş-am</td>
<td>yez-miş-yq</td>
</tr>
<tr>
<td>2 yez-mi(s)-sən-yez-yən-san</td>
<td>yez-mi(s)-siniz-yez-yən-siniz</td>
</tr>
<tr>
<td>3 yez-miş-dIr-yez-yən</td>
<td>yez-miş-lər-yez-yən-lər</td>
</tr>
</tbody>
</table>

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)lb 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 Şiraliyev 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 Öfendiyeva’s (2005) book-length Felin keçmiş zaman formalari [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 [Öfendiyeva] 1958).

Table 4 reflects Öfendiyeva’s description of the facts.

<table>
<thead>
<tr>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>yaz-miş-am</td>
<td>yaz-miş-tq</td>
</tr>
<tr>
<td>yaz-m(ş)-san–yaz-ib-san</td>
<td>yaz-m(ş)-siniz–yaz-ib-siniz</td>
</tr>
<tr>
<td>yaz-miş-dir–yaz-ib(-dir)</td>
<td>yaz-miş-dir-lar–yaz-ib(-dir)-lar</td>
</tr>
</tbody>
</table>

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

On this view, -dir 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 -ml$ş$ 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 -dir displays a similar sort of variation elsewhere in the language, as seen in Table 5.

<table>
<thead>
<tr>
<th>Present</th>
<th>Aorist</th>
<th>Future/Prospective</th>
<th>Necessitative</th>
<th>Optative</th>
</tr>
</thead>
<tbody>
<tr>
<td>3SG</td>
<td>yaz-ır</td>
<td>yaz-ar</td>
<td>yaz-açaq(-dir)</td>
<td>yaz-mali(-dir)</td>
</tr>
<tr>
<td>3P</td>
<td>yaz-ır-lar</td>
<td>yaz-ar-lar</td>
<td>yaz-ar(-dir)-lar</td>
<td>yaz-mali(-dir)-lar</td>
</tr>
</tbody>
</table>

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

All descriptions agree that -dir is an optional marker of the third person in the future/progressive and in the necessitative, and not, for example, in the present (e.g. yazır ‘(s)he is writing’ but *yazirdır), the aorist (e.g. yazar ‘(s)he would write’ but *yazardır), or optative (e.g. yaza ‘that (s)he write’ but *yazadır). 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 -dir is optional in other paradigms. The paradigms given in Tables 3 and 4 even have -dir as an obligatory marker in some cells. This would run contrary to the other patterns seen in Table 5, where -dir 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 -ml$ş$ and evidential =(i)ml$ş$

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

38
The evidential 

\[=(i)mIş\] in Azerbaijani (Şirəliyev & Sevortjan: 127; Kazımov 2010: 244).

\(a.\) yaz.muş-am
write.miş-1SG

SENSE 1 (default sense): ‘I have written.’
SENSE 2 (alternative): ‘I wrote/write/have written, apparently (they say, evidently).’

\(b.\) yaxshi=müš\(*=-dIr) / yaxşi=imiş\(*=-dIr)

\[\text{good}=\text{EVD} \quad / \quad \text{good}=\text{EVD}\]

Ök: ‘This is/was has been good, apparently (they say, evidently).’

\*‘This has been good.’

While 

\[-mIş\] 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 \n
\[=(i)mIş\] vs. perfect \n
\[-mIş\]~\n
\[-(y)Ib\]

\(a.\) ?/*oxu.miş
read.miş

\[\text{INTENDED} 1: '(s/he) has read.' (=perfect)\]
\[\text{INTENDED} 2: 'I wrote/write/have written, apparently.' (=evidential)\]

\(b.\) oxu.miş-dur
read.miş-3SG

\[\text{INTENDED} 1: \text{perfect}\]
\[\text{INTENDED} 2: \text{evidential}\]

\(c.\) oxu-yib
read-ib

\[\text{INTENDED} 1: \text{perfect}\]
\[\text{INTENDED} 2: \text{evidential}\]

\(d.\) oxu-yib-dur
read-ib-3SG

\[\text{INTENDED} 1: \text{perfect}\]
\[\text{INTENDED} 2: \text{evidential}\]

Verb+\n
\[=(i)mIş\] 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, \n
\[=(i)mIş\] is temporally indifferent across Turkic languages, where as 

\[-mIş\] tends to have a temporal sense.

Even if 

\[-mIş\] as an exponent of the perfect is distinguishable from the evidential 

\[=(i)mIş\], we have not yet considered the meaning(s) of 

\[-mIş\] relative to 

\[-(y)Ib\]. The Turkish cognate suffix 

\[-mIş\] is also sometimes described as a perfect suffix, but as Bowler and Ozkan (2018) have argued, Turkish 

\[-mIş\] 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 -mIş.

(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 -mIş**
   
   *Context:* Einstein is dead, but you have just seen his signature in the physics department guestbook at Princeton.
   
   Einstein Princeton-i ziyaret et-miş
   Einstein Princeton-ACC visit do-MIŞ
   ‘(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 -mIş and -(y)Ib 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-ni ziyarət et-miş
   Chaplin Japan-ACC visit do-MIŞ

b. ?#Çaplin Yaponiya-ni ziyarət et-miş-dir
   Chaplin Japan-ACC visit do-MIŞ-3SG

c. ?#Çaplin Yaponiya-ni ziyarət ed-ib
   Chaplin Japan-ACC visit do-MIŞ-3SG

d. ?#Çaplin Yaponiya-ni ziyarət ed-ib-dir
   Chaplin Japan-ACC visit do-MIŞ-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 -dIr when it is realized by -mIş, but not when it is realized by -(y)Ib.

3. **Reorganization in the Azerbaijani present perfect**

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

**First person (singular and plural):** -mIş
**Second person (singular and plural):** -mIş- -(y)Ib
**Third person (singular and plural):** -mIşdIr- -(y)Ib- -(y)IbdIr
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)lb, but obligatory with -mlş. Zero marked third person -mlş 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 -mlş, 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)lb? Perhaps this suffix never developed a full paradigm for all persons. After all, it has been in competition with -mlş 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 15th century, which then diverged into Old Ottoman and what we may call Old Azerbaijani. The suffix -(y)lb–dlr 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 –dlr is 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 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 –(l)b. The grammaticalization of -(y)lb as separable exponent of the perfect in its own right in Azerbaijani can be dated to the 15th century (Tanrıverdi 2017: 301). Both the converb suffix –(y)lb (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. -mlş 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).

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

![Figure 1. The grammaticalization of -(y)lb–dlr>-(y)lb and -dlr.](image)
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 Bəngə Badoğlu, XVI cent.)

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 17th 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-im-a düş-dü ki, bu aşər-lar
...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, Şiraliyev & Sevortyan (1971: 127) call them ‘complex/composite forms of the past tense’ [сложные/составные формы прошедшего времени]; Kazımov (2010: 173-175) groups them with ‘complex verbs’ [Mürekəkbələr]; Əfəndiyeva calls them ‘complex forms of the past tense’ [Keçiş zaman kategoriyasının mürekəkbələr], 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 =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 =mIş by the person and number markers.

<table>
<thead>
<tr>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 yazmIşmIşam-yazmIş imIşəm</td>
<td>yazmIşmIşəq-yazmIş imIşik</td>
</tr>
<tr>
<td>2 yazmIşmIş(ş)san-yazmIş imIş(ş)sən</td>
<td>yazmIşmIş(ş)siniz-yazmIş imIş(ş)siniz</td>
</tr>
<tr>
<td>3 yazmIşmIş-yazmIş imIş</td>
<td>yazmIşmIşlar-yazmIş imIşlor</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 yazmIşammIş-yazmIşam imIş</td>
<td>yazmIşıqmIş-yazmIşıq imIş</td>
</tr>
<tr>
<td>2 yazmIş(ş)sanmIş-yazmIş(ş)san imIş</td>
<td>yazmIş(ş)sinızmIş ~yazmIş(ş)sinız imIş</td>
</tr>
<tr>
<td>3 yazmIşmIş-yazmIş imIş (same as table 6)</td>
<td>yazmIşlarmIş-yazmIşlar imIş</td>
</tr>
</tbody>
</table>

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)lb perfect, seen in Table 8. Unlike present perfect -(y)lb perfect forms, the first person is permissible in the evidential of the perfect (ibid.: 242).

<table>
<thead>
<tr>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 yazibmişam–yazib imişom</td>
<td>yazibmişiq–yazib imişik</td>
</tr>
<tr>
<td>2 yazibmiş(s)san–yazib imis(s)son</td>
<td>yazibmiş(s)siniz–yazib imis(s)siniz</td>
</tr>
<tr>
<td>3 yazibmiş–yazib imiş</td>
<td>yazibmişlar–yazib imişlor</td>
</tr>
</tbody>
</table>

Table 8: The evidential of the -(y)lb 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 yazmişmişam / yazmiş imişom / yazmişmişam en yazmiş imiş en yazmışmam / yazib imişom. 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)lb and precedes =(i)mış, 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.


‘We looked and saw that they have/had evidently/apparently conducted/closed the contract in a month.’


‘His/her parents have/had evidently 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)lb and the evidential =(i)mış, 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 19th century playwrite Axundov’s Sərgüləsta–mərd–xas: Hacı Qara [The Adventure of the Miser: Haji Gara] the naçalnik character says the sentence in (11) to the titular Haji Gara:
(11) Boli, çok büyük 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)Iβ, as indicated by the question marks in Table 9, but the judgements here are not as degraded as the first-person marked -(y)Iβ simple perfect forms.

<table>
<thead>
<tr>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 ?yazıbammış-yazıbam imiş</td>
<td>?yazıbiqmiş-yazıbiq imiş</td>
</tr>
<tr>
<td>2 yazıbsanmiş-yazybsan imiş</td>
<td>yazıbsimizmiş-yazybsimiz imiş</td>
</tr>
<tr>
<td>3 yazıbmiş-yazıb imiş</td>
<td>yazıblarmış-yazıblar imiş</td>
</tr>
</tbody>
</table>

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

As the historical evidence in (7) demonstrates, the -(y)Iβ 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.

<table>
<thead>
<tr>
<th>Author</th>
<th>Nəmatullah Kışvari</th>
<th>Xətai</th>
<th>Füzuli</th>
<th>Vaqif</th>
<th>Axundov</th>
<th>Şirvani</th>
<th>--</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suff. / Period</td>
<td>(1490-1502)</td>
<td>(1487-1524)</td>
<td>(1494-1556)</td>
<td>(1717-1797)</td>
<td>(1812-1878)</td>
<td>(1835-1888)</td>
<td>Modern</td>
</tr>
<tr>
<td>mlş=mlş-X</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>0.06</td>
</tr>
<tr>
<td>mlş-X=mlş</td>
<td>0</td>
<td>0</td>
<td>9.39</td>
<td>0</td>
<td>102.07</td>
<td>6.12</td>
<td>0.4</td>
</tr>
<tr>
<td>Iβ=mlş-X</td>
<td>26.94</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6.12</td>
<td>30.33</td>
</tr>
<tr>
<td>Iβ-X=mlş</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>17.01</td>
<td>24.49</td>
<td>3.12</td>
</tr>
</tbody>
</table>

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.
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.

Table 12: Distribution of perfect exponence by person

<table>
<thead>
<tr>
<th>Present perfect</th>
<th>Evidential of perfect</th>
</tr>
</thead>
<tbody>
<tr>
<td>First person (singular and plural):</td>
<td>-mlš</td>
</tr>
<tr>
<td>Second person (singular and plural):</td>
<td>-mlš~(y)lb</td>
</tr>
<tr>
<td>Third person (singular and plural):</td>
<td>-mlšdIr~(y)lb~(y)lbdlr</td>
</tr>
</tbody>
</table>

Table 13: Distribution of evidential exponence by person

<table>
<thead>
<tr>
<th>Base evidential</th>
<th>Evidential of perfect</th>
</tr>
</thead>
<tbody>
<tr>
<td>First person (singular and plural):</td>
<td>=(y)mlš~imiş</td>
</tr>
<tr>
<td>Second person (singular and plural):</td>
<td>=(y)mlš~imiş</td>
</tr>
<tr>
<td>Third person (singular and plural):</td>
<td>=(y)mlš~imiş</td>
</tr>
</tbody>
</table>

Interestingly, just as with the -(y)lb perfect marker, the evidential =(i)mlš 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 -dIr marker—which is typically optional where it does occur (as in Table 5)—is not optional in the perfect when marked by -mlš because it functions as an additional marker of constrast with the evidential =(y)mlš. 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 mlŞ. This creates a contrast in person in the perfect category, which—speculatively—seems to have led to the differentiation of the -(y)lb marker by person as well. As Öndiyeva 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 -mlŞ, e.g. Ottoman al- mış- tü ‘(S)he has bought’ contrasted with al- mış ‘(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.

<table>
<thead>
<tr>
<th>Categories</th>
<th>‘heal’ Class</th>
<th>‘feel’ Class</th>
<th>‘give’ Class</th>
<th>‘kneel’ Class</th>
<th>‘burn’ Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Past</td>
<td>heal(ed)</td>
<td>felt</td>
<td>gave</td>
<td>kneel / kneeled</td>
<td>burned/burnt</td>
</tr>
<tr>
<td>Past participle</td>
<td>heal(ed)</td>
<td>felt</td>
<td>give(n)</td>
<td>?kneel</td>
<td>burned/burnt</td>
</tr>
<tr>
<td>Imperative</td>
<td>heal(Ø)</td>
<td>feel(Ø)</td>
<td>give(Ø)</td>
<td>kneel(Ø)</td>
<td>burn(Ø)</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Categories</th>
<th>-s Class</th>
<th>-es Class</th>
<th>Variable Class</th>
<th>Feminine Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>NomSg</td>
<td>Deutsch</td>
<td>Busch</td>
<td>Vaterland</td>
<td>Hand</td>
</tr>
<tr>
<td>GenSg</td>
<td>Deutschs</td>
<td>Busches</td>
<td>Vaterlands/Vaterlandes</td>
<td>Hand(Ø)</td>
</tr>
<tr>
<td>Translation</td>
<td>German</td>
<td>bush</td>
<td>fatherland</td>
<td>hand</td>
</tr>
</tbody>
</table>

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 -mIşdIr. The most common verbs marked by these suffixes in the third person largely overlap in the SketchEngine corpus, as seen in Table 16.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Verb (y)Ib</th>
<th>Token frequency</th>
<th>Per million</th>
<th>Verb (y)IbdIr</th>
<th>Token frequency</th>
<th>Per million</th>
<th>Verb -mIşdIr</th>
<th>Token frequency</th>
<th>Per million</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>et- ‘do’</td>
<td>113,204</td>
<td>981.99</td>
<td>ol- ‘be’</td>
<td>579</td>
<td>5.02</td>
<td>et- ‘do’</td>
<td>28,681</td>
<td>248.79</td>
</tr>
<tr>
<td>2.</td>
<td>ol- ‘be’</td>
<td>79,554</td>
<td>690.09</td>
<td>et- ‘do’</td>
<td>452</td>
<td>3.92</td>
<td>ol- ‘be’</td>
<td>25,966</td>
<td>225.24</td>
</tr>
<tr>
<td>3.</td>
<td>bildir- ‘inform’</td>
<td>52,691</td>
<td>457.07</td>
<td>edil- ‘do’ (PASS)</td>
<td>335</td>
<td>2.91</td>
<td>edil- ‘do’ (PASS)</td>
<td>20,278</td>
<td>175.90</td>
</tr>
<tr>
<td>4.</td>
<td>ver- ‘give’</td>
<td>43,234</td>
<td>375.03</td>
<td>olun- ‘be’ (REFL)</td>
<td>333</td>
<td>2.89</td>
<td>olun- ‘be’ (REFL)</td>
<td>15,365</td>
<td>133.28</td>
</tr>
<tr>
<td>5.</td>
<td>olun- ‘be’ (REFL)</td>
<td>41,337</td>
<td>358.58</td>
<td>veril- ‘give’ (PASS)</td>
<td>188</td>
<td>1.63</td>
<td>keçiril- ‘pass’ (PASS)</td>
<td>11,348</td>
<td>98.44</td>
</tr>
<tr>
<td>6.</td>
<td>edil- ‘do’ (PASS)</td>
<td>38,255</td>
<td>331.84</td>
<td>yaradil- ‘create’ (PASS)</td>
<td>167</td>
<td>1.45</td>
<td>veril- ‘give’ (PASS)</td>
<td>8,032</td>
<td>69.67</td>
</tr>
<tr>
<td>7.</td>
<td>de- ‘say’</td>
<td>29,970</td>
<td>259.97</td>
<td>ver- ‘give’</td>
<td>154</td>
<td>1.34</td>
<td>veril- ‘give’ (PASS)</td>
<td>7,562</td>
<td>65.60</td>
</tr>
<tr>
<td>8.</td>
<td>keçiril- ‘pass’ (PASS)</td>
<td>20,552</td>
<td>178.28</td>
<td>çevril- ‘turn’ (intr.)</td>
<td>146</td>
<td>1.27</td>
<td>bildir- ‘inform’</td>
<td>4,820</td>
<td>41.81</td>
</tr>
<tr>
<td>9.</td>
<td>veril- ‘give’ (PASS)</td>
<td>19,269</td>
<td>167.15</td>
<td>art- ‘increase’</td>
<td>127</td>
<td>1.10</td>
<td>yaradil- ‘create’ (PASS)</td>
<td>4,510</td>
<td>39.12</td>
</tr>
<tr>
<td>10.</td>
<td>başla- ‘start’</td>
<td>16,345</td>
<td>141.78</td>
<td>çat- ‘reach’</td>
<td>112</td>
<td>0.97</td>
<td>başla- ‘start’</td>
<td>3,513</td>
<td>30.47</td>
</tr>
<tr>
<td>11.</td>
<td>varğula- ‘stress’</td>
<td>13,613</td>
<td>118.09</td>
<td>görül- ‘see’</td>
<td>107</td>
<td>0.93</td>
<td>al- ‘take’</td>
<td>3,302</td>
<td>28.64</td>
</tr>
<tr>
<td>12.</td>
<td>al- ‘take’</td>
<td>13,239</td>
<td>114.84</td>
<td>tikil- ‘build’ (PASS)</td>
<td>99</td>
<td>0.86</td>
<td>görül- ‘see’ (PASS)</td>
<td>2,969</td>
<td>25.75</td>
</tr>
<tr>
<td>13.</td>
<td>gol- ‘come’</td>
<td>10,556</td>
<td>91.57</td>
<td>yaran- ‘create’ (intr.)</td>
<td>83</td>
<td>0.72</td>
<td>aparil- ‘carry’ (PASS)</td>
<td>2,885</td>
<td>25.03</td>
</tr>
<tr>
<td>14.</td>
<td>tutul- ‘hold’ (PASS)</td>
<td>9,468</td>
<td>82.13</td>
<td>al- ‘take’</td>
<td>81</td>
<td>0.70</td>
<td>tutul- ‘hold’ (PASS)</td>
<td>2,658</td>
<td>23.06</td>
</tr>
<tr>
<td>15.</td>
<td>alın- ‘take’</td>
<td>9,347</td>
<td>81.08</td>
<td>keçiril- ‘pass’ (PASS)</td>
<td>80</td>
<td>0.89</td>
<td>yarar- ‘create’ (tr.)</td>
<td>2,652</td>
<td>23.00</td>
</tr>
</tbody>
</table>

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. 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.

---

4 I have also excluded common non-verbs which happen to end in <ub>, <ib>, <ib>, and <üb>, which did appear in the search results, e.g. sahib ‘owner, vacib ‘important’, münasib ‘reasonable’, ayib ‘shame’, etc.
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).

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

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).

<table>
<thead>
<tr>
<th></th>
<th>Affirmative aorist</th>
<th>Negative aorist</th>
</tr>
</thead>
<tbody>
<tr>
<td>First person (singular and plural):</td>
<td>-Ar</td>
<td>-mA r</td>
</tr>
<tr>
<td>Second person (singular and plural):</td>
<td>-Ar</td>
<td>-mA z</td>
</tr>
<tr>
<td>Third person (singular and plural):</td>
<td>-Ar</td>
<td>-mA z</td>
</tr>
</tbody>
</table>

Table 19: Distribution of aorist suffixes by person

The situation in Old Anatolian Turkish reportedly remarked all negative aorist forms with -mA r (Fəxrəddinqızı 2010: 121), but after a (later?) period of competition between -mA r and -mA z, the latter came to be associated with the first person. Similarly, several of the Azerbaijani dialects in Şiraliyev’s (2008) survey seem to have come to associate -(y)ləb or -mI ş 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**

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


https://doi.org/10.1075/scls.203.05cap.


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