

A Paradigm Gap in Turkish

Muhammed İleri & Ömer Demirok*

Abstract. In this paper, we argue that Turkish has a gap in the third person plural cell of the person-number agreement paradigm of desiderative constructions formed with the -AsI suffix. We provide evidence for this claim from a corpus search and an acceptability judgment experiment. The corpus search shows that the third person plural suffix is virtually unattested with -AsI desideratives and the results of the experiment show that the third person plural suffix significantly reduces the acceptability of -AsI desideratives. In order to account for the observation that third person plural desideratives are unacceptable for most speakers, we argue that both negative evidence and competition accounts contribute to the existence and persistence of the gap. We discuss that competition accounts are supported by the presence of two competing forms whereas negative evidence accounts are supported by the anomalous relative frequency distribution in the paradigm of desideratives.

Keywords. morphology; paradigm gap; productivity; corpus; Turkish

1. Introduction. Productivity is regarded as one of the defining properties of human language. Speakers/signers with no language related impairment are able to produce and understand novel words and sentences based on their implicit knowledge about the grammar of their language: they can readily inflect words they have never seen before for various grammatical feature categories such as tense, aspect and agreement. However, there are also cases where language users unexpectedly fail to inflect a word for a given set of grammatical features. The inflected forms of certain lexemes that are expected to exist but are absent are called lexical or paradigm gaps in the literature (Albright 2003, Sims 2006, Baerman et al. 2010, Gorman & Yang 2019).

In this paper, we report – for the first time, to the best of our knowledge – a gap in the 3PL cell of the agreement paradigm of a construction we call *-AsI¹ desideratives* in Turkish. That is, there are two attested forms for 3PL desideratives, as exemplified in (1)²; however, we argue that neither of the forms is acceptable for most speakers of Turkish.

- (1) a. * (onlar-ın) Kahve iç -esi -leri var.
(they-GEN) Coffee drink -DESID -3PL.POS exist
'They feel like drinking coffee.'
- b. * (onlar-ın) Kahve iç -e -leri var.
(they-GEN) Coffee drink -DESID -3PL.POS exist
'They feel like drinking coffee.'

In (1), there is a 3PL subject, *onların*, which agrees in person and number with the main predicate of the embedded desiderative clause, *iç-* 'to drink'. *onların* is optional since pronouns can be dropped in Turkish (Göksel & Kerslake 2005). What differs between the two sentences is the form of the desiderative suffix: the form in (1-a) spells out the desiderative morpheme with -A whereas the one in (1-b) spells it out with -AsI. Every speaker prefers one form or the other, but

* Authors: Muhammed İleri, Boğaziçi University (muhammed.ileri@boun.edu.tr) & Ömer Demirok, Boğaziçi University (omerfaruk.demirok@boun.edu.tr).

¹ The capital letters represent archiphonemes: they may change according to the vowel harmony rules of Turkish.

² Abbreviations: 1 = first person, 2 = second person, 3 = third person, DESID = desiderative, POS = possessive, GEN = genitive, ACC = accusative, EVID = evidential, SG = singular, PL = plural

not both. However, most speakers tend to reject even their preferred form. Thus, there is no acceptable form for 3PL desideratives for most speakers of Turkish.

Moreover, as will be discussed in the following sections, 3PL agreement marker on desiderative verbs reduces acceptability more when the lexeme is more frequent. We will argue that this finding is compatible with negative-evidence (or surprisal) accounts, such as the one proposed by (Daland et al. 2007), which posit that speakers are sensitive to the frequency distribution of the forms in a paradigm. Hence, 3PL desideratives are defective with all lexemes; however, the defectivity is stronger for lexemes that are more frequent.

In what follows, we first provide some background on -AsI desideratives and their agreement paradigm, and discuss the source of the gap in section 2. In section 3, we report the results of a corpus search and an acceptability judgment experiment, which corroborate our intuitions that there is a gap in the third plural agreeing desiderative verbs. In section 4, we discuss the results and their implications for theories of defectivity. Section 5 concludes.

2. -AsI desideratives. Turkish has a desiderative construction formed with the suffix -AsI, which attaches to the main verb of an embedded clause, as exemplified in (2). In (2), the embedded clause inside the brackets is the desiderative clause. It is the only argument of the main predicate of the matrix clause, the existential *var*. The subject of the desiderative clause is marked with the genitive case suffix and its main predicate, *yapasım*, agrees in person and number with this genitive subject.

- (2) [Ben-im tatlı yap-ası-m] var.
 [I-GEN dessert do-DESID-1SG.POS] exist
 Literal: ‘My desire to make a dessert exists.’
 ‘I feel like making a dessert.’

Sentences like (2) report the absence or presence of a desire. The desire meaning is conveyed via the -AsI clause, the *desiderative*, whereas the absence or presence meaning is conveyed via the main predicate of the matrix clause. Thus, desiderative clauses can be taken as complement by only a small set of predicates that have an existential meaning, such as *gel-* ‘come’, *git-* ‘go’, *kaç-* ‘go.away’ and *tut-* ‘hold’ in addition to the existential predicates *var* ‘exist’ and *yok* ‘not exist’.

2.1. POSSESSIVE AGREEMENT PARADIGM(S). Desideratives are like most nominalized clauses in Turkish in terms of their morpho-syntactic properties such as case marking and agreement patterns. Most nominalized clauses have a genitive subject, which agrees in person and number with the main predicate of that clause. Normally, when the subject is in Genitive, we see a suffix from the possessive agreement paradigm on the verb. Typically, when the subject is 3SG, the possessive agreement suffix is -(s)I. This pattern is exemplified with a Genitive-Possessive construction in (3) and with a nominalized clause formed with the suffix -mA in (4).

- (3) Merve’nin tatlı-sı
 Merve-GEN dessert-3SG.POS
 ‘Merve’s dessert’
- (4) [Merve’nin tatlı yap-ma-sı] lazım.
 [Merve-GEN dessert do-NMLZ-3SG.POS] necessary
 ‘Merve needs to make a dessert.’

However, the desiderative constructions come with an exceptional pattern: we do not observe an

additional -(s)I suffix on 3SG agreeing main verbs of desiderative clauses. Instead, the -AsI suffix assumes the function of both the desiderative suffix and third person singular agreement marker, as shown in (5).

- (5) Merve-nin tatlı yap-**asI** var.
 Merve-GEN dessert do-**DESID.3SG.POS** exist
 ‘Merve feels like making a dessert.’

Apart from the irregularity in 3SG desideratives, the agreement markers on -AsI desideratives (Table 2) come from the regular possessive agreement paradigm, which is exemplified with the paradigm of -mA nominalization in Table 1. Moreover, Table 2 illustrates that there are two possible but ungrammatical forms in the 3PL desiderative cell: one of them spells out the desiderative morpheme with -AsI whereas the other spells it out with -A.

	SG	PL
1	X-mA-m	X-mA-mIz
2	X-mA-n	X-mA-nIz
3	X-mA-sI	X-mA-lArI

Table 1. The regular possessive agreement paradigm of -mA nominalization.

	SG	PL
1	X-AsI-m	X-AsI-mIz
2	X-AsI-n	X-AsI-nIz
3	X-AsI- \emptyset	*X-AsI-lArI / *X-A-lArI

Table 2. The possessive agreement paradigm of -AsI desideratives.

2.2. SOURCE OF TWO DIFFERENT 3PL FORMS. We argue that there are two reasons why two possible forms exist for 3PL desideratives and why neither of them is grammatical for most speakers. The first reason is the absence of 3PL desideratives in the input and the second reason is the irregularity in the form of 3SG desideratives. Since speakers cannot store a form they have not heard before, they need to derive 3PL desideratives on the fly. In order to do so, they need to rely on the patterns or rules they generalized based on related forms in the input (Albright & Hayes 2003; Yang 2016; Sims 2015). However, as argued by Albright (2003) in his analysis of 1SG gaps in Spanish, the reliability of such rules are affected by the irregularities in the paradigm. Building on this observation, we argue that the irregularity in 3SG desideratives hinders speakers from extracting a unique reliable rule for deriving 3PL desideratives. In particular, we posit that speakers may come up with at least two different patterns of allomorphy, which lead to two different forms for the 3PL cell. However, since speakers have no positive evidence in the input for the well-formedness of either form, they are hesitant about the acceptability of these forms.

To illustrate this process, a speaker cannot confidently decide based on other forms in the paradigm that all desideratives are formed by combining a verbal stem, the desiderative suffix -AsI and a regular possessive suffix since the 3SG desiderative has one complex suffix, -AsI, to spell out both DESID and 3SG.POS morphemes. In order to make a generalization of this sort, the speaker needs to decide how to decompose the portmanteau -AsI in 3SG desideratives. They may proceed to decompose it in at least two different ways: the first one is that -A is the allomorph of DESID and -(s)I is the regular exponent of 3SG.POS (6-a); the other is that -AsI is the regular exponent of DESID and 3SG.POS has a null allomorph in the environment of DESID (6-b).³

³ A reviewer notes that another possible way of deriving a 3SG desiderative is to first combine the regular suffixes for DESID (-AsI) and 3SG.POS (-(s)I) and then to delete the exponent of 3SG.POS via a haplological process (i.e. $X+AsI+sI \rightarrow X+AsI+\emptyset$). Although this analysis is also possible, if all speakers had only this derivation, the exis-

- (6) a. yap-a-sı
do-DESID-3SG.POS
- b. yap-ası-∅
do-DESID-3SG.POS

There is no definitive evidence suggesting that one decomposition is grammatical and the other is not, but this is not a problem to produce and understand 3SG desideratives: the meaning of 3SG desideratives is straightforward even without decomposing -AsI into smaller components. If the verb does not have an overt agreement marker right next to -AsI, then it is interpreted to agree with a 3SG subject; but if it has an agreement marker, then it is one of the other forms and will be interpreted according to the agreement marker. This way, speakers can avoid making a decision for how to decompose the complex suffix -AsI in 3SG desideratives without jeopardizing the comprehension and production of 3SG desideratives.

However, without decomposing the complex -AsI in 3SG desideratives into smaller components, speakers cannot know how to form 3PL desideratives: if speakers decompose *yapası* as in (6-a), there is no evidence available to them to make sure if DESID has an allomorph only in the environment of 3SG. It might as well be that DESID has the allomorph -A in the environment of third person, which extends the environment of -A to 3PL. To illustrate, let's demonstrate the derivation of the verbal stem *yap* 'do'. If speakers adopt the parsing in (6-b), it means they have decided that the desiderative suffix is -AsI in all cells in the paradigm: the irregularity is in the 3SG.POS, which has a null exponent. In this analysis, speakers would produce the form in (7-b) as the 3PL desiderative form of *yap* by combining the verbal stem *yap*, the regular DESID suffix -AsI, and the regular 3PL.POS suffix -larI. However, if speakers adopt the parsing in (6-a), they can produce both forms in (7). If they decide that DESID has the allomorph -A only in the environment of 3SG while it has the exponent -AsI elsewhere, including in the environment of 3PL, they would produce the form in (7-b) for the 3PL desiderative form of *yap*. However, if speakers hypothesize that DESID has the allomorph -A in the environment of the third person feature (i.e. regardless of the number feature), they would produce the form in (7-a) since DESID would need to be spelled out by its allomorph -A in the environment of 3PL and the 3PL.POS would be spelled out by its regular exponent -larI.

- (7) a. *yap-a-ları
do-DESID-3PL.POS
- b. *yap-ası-ları
do-DESID-3PL.POS

We know that both forms in (7) are possible given that some speakers prefer (7-a) and some speakers prefer (7-b) over the other. We also know that Turkish speakers rarely hear, if any, 3PL desideratives which they can memorize since they are virtually absent in the input according to our corpus search. Thus, speakers need to parse the complex suffix -AsI in 3SG desiderative constructions into the exponents of the DESID and 3SG so as to infer a rule for how to produce 3PL desideratives. Moreover, depending on their initial parsing decision, they would also need to decide if the desiderative suffix allomorphy is restricted to third person singular forms or not. This is too complex a task for a rather infrequent construction in the absence of conclusive evidence.

tence of two alternating forms for 3PL desideratives and the split among the speakers regarding their preferred 3PL desiderative form could not be explained. Hence, this derivation cannot be the only option; it can only be one of the many alternatives which lead to the form $X+AsI+larI$ for 3PL desideratives – $*-sI+sI$ haplology would not apply when deriving the 3PL desiderative form. That is, this analysis bears the same result for the form of 3PL desideratives as (6-b), where DESID is -AsI and 3SG.POS is null.

Although every speaker consistently prefers one of the forms in (7) over the other, most of them are indecisive concerning the well-formedness of their preferred form.⁴

In the next section, we turn to empirical evidence from a corpus search and an acceptability judgment experiment for our hypothesis that there is a gap in the 3PL cell of the desiderative agreement paradigm.

3. Method.

3.1. CORPUS DATA. Our purpose for running a corpus search was to detect if the frequency of 3PL desideratives is anomalous relative to other forms in the desiderative agreement paradigm. However, corpus search can be challenging especially for some specific purposes like ours. For instance, automatically annotated corpora of Turkish err significantly when detecting the suffix -AsI due to the fact that it is often mistaken for the optative suffix -A and the third person singular possessive suffix -(s)I, which happens to be the most common suffix in Turkish (Bilgin 2016). When it comes to hand-annotated corpora, they are too small for our purposes due to the low base frequency of desideratives overall. Thus, even relatively frequent desiderative forms would be rarely found in a small corpus, let alone the 3PL desideratives.

Nevertheless, our main purpose is not to fully analyze the corpus frequencies of desideratives. It is to provide evidence for the atypical behavior of 3PL desideratives. Therefore, we utilized the biggest Turkish corpus online: TrTenTen corpus on *sketchengine.eu* (3.3 billion word tokens). Since TrTenTen is not morphologically annotated for the desiderative, we needed to restrict ourselves to a small number of verbs. We decided to do a search for all 6 possible desiderative forms of 4 verbs that are among the top 10 most frequent verbs in Turkish based on overall frequencies (Aksan et al. 2016). Our reasoning is that if 3PL is allowed in desideratives, observing a high frequency verb in 3PL desiderative form would be more likely than observing a less frequent one. More importantly, this strategy would allow us to see the relative frequency of 3PL forms compared to other desideratives.

We tried to approximate the number of occurrences of the desiderative bearing verbs in the corpus by using the small set of predicates (*gel-*, *git-*, *kaç-*, *tut-*, *var*, *yok*) that can take desideratives as arguments. We could find only 17 instances of 3PL desideratives. However, this does not provide enough evidence on its own that the combination of 3PL and DESID features lead to decreased acceptability and uncertainty. As is well known, the absence of a form in a corpus does not prove that the form is ungrammatical (Divjak 2008, Bader & Häussler 2010). For example, if desideratives have very low baseline frequency and are rarely attested overall, then it is very normal that 3PL forms are also attested rarely. For this reason, instead of raw frequencies, the relative frequencies of the forms in a paradigm should be used to detect an anomalous frequency distribution (Sims 2015).

⁴ A reviewer rightly notes that Turkish speakers tend to judge possessive constructions with a possessum that bears the 3PL possessive agreement marker -IArI as unacceptable when it has an overt 3PL possessor, as shown in (i), and asks if it might be a confound in our study.

(i) ?onlar-in at-ları
they-GEN horse-3PL.POS

Since our experimental items do not include overt 3PL possessors and 3PL agreement is obligatory when the subject/possessor is covert in Turkish (Göksel & Kerslake 2005), this observation does not constitute a problem for our analysis: desiderative verbs that have an understood 3PL subject need to bear the exponent of 3PL.POS, -IArI.

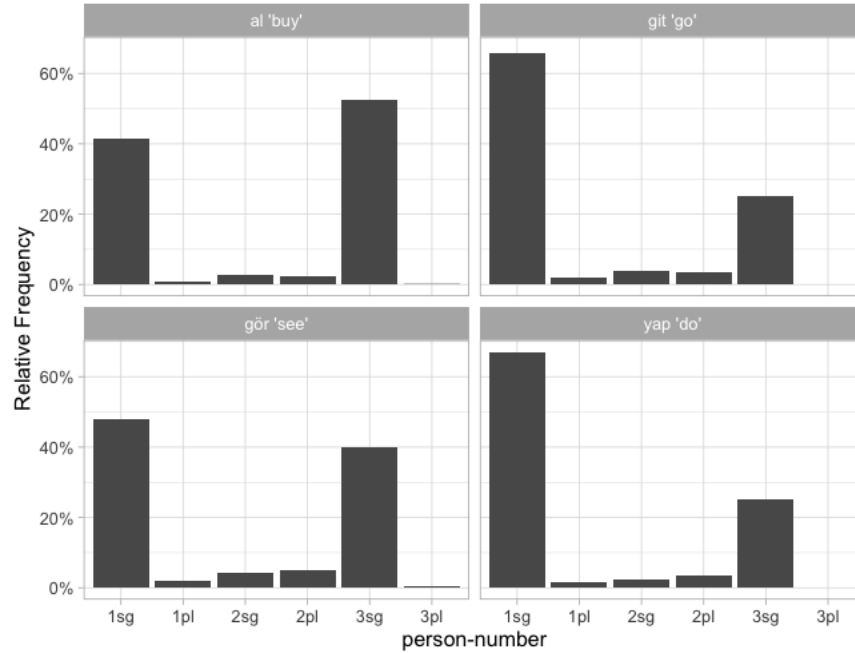


Figure 1. Relative frequencies of the desideratives of 4 frequent verbs with respect to their person-number agreement markers. 1SG and 3SG forms dominate the frequencies while only 2 instances of 3PL forms (*alasıları* and *göresileri*) are attested.

The pattern is similar in all 4 verbs (Figure 1). It can be seen that 1SG and 3SG forms dominate desiderative constructions. There were only 2 attested tokens of 3PL desideratives out of 3764 tokens: *alasıları* and *göresileri*. Even though 1PL forms seem to be rarely attested, too, when we collapse frequencies across the verbs, we see that it is actually 28 times more likely to see a 1PL form than a 3PL, given these verbs.

Moreover, when we check the relative frequency distribution of the forms in the desiderative agreement paradigm (Table 3), it can be observed that the relative frequency of 3PL desideratives are much lower than all the other forms in the paradigm. It is actually so rare that 3PL desideratives have even lower relative frequency than the relative frequency of the non-past 1SG agreeing form of the Russian verb *pobedit'* 'to win', which is a well-established gap in the literature, based on the data in (Sims 2015:226).

	SG	PL
1	59.2%	1.51%
2	3.06%	3.45%
3	32.7%	0.05%

Table 3. Averaged frequencies of person/number suffixes in desideratives given one of the verbs *al-* 'buy', *git-* 'go', *gör-* 'see', *yap-* 'do'.

In conclusion, the corpus data provide evidence that the bare attestation of 3PL desideratives is not simply due to the low base frequency of desideratives. Rather, 3PL has an abnormal relative frequency distribution in its paradigm, which is typical of paradigm gaps reported in the

literature. These data confirm our intuitions that 3PL desideratives pose a problem for speakers of Turkish. Its usage is much lower than expected and it instantiates a paradigm gap.

3.2. EXPERIMENT. This study is approved by the Boğaziçi University Ethics Committee for Master and PhD Theses in Social Sciences and Humanities (number 47879, 13 January 2022). The procedures in this study conformed with the Helsinki Declaration’s ethical principles for research involving human subjects.

The primary aim of the experiment is to test if verbs bearing the desiderative suffix *-AsI* are ungrammatical in Turkish when inflected for third person plural agreement. Our hypothesis is that they are ungrammatical and there is a gap in the third plural cell of the agreement paradigm of desideratives. In addition, we also have a secondary aim which is to detect if the frequency of the verbal lexeme in a third plural agreeing desiderative verb affects its grammaticality. Even though we do not predict a frequency effect on the acceptability of third plural agreeing desiderative verbs based on our native speaker consultants’ (and our) intuitions, several studies on paradigm gaps report that low frequency lexemes reduce the acceptability of an inflected form more than high frequency lexemes (Albright 2003, Sims 2006). Hence, we also test the effect of lexeme frequency on the acceptability of desiderative verbs bearing the 3PL marker.

3.2.1. PARTICIPANTS. 183 students from Boğaziçi University, ranging in age from 18 to 32 years ($M = 21.4$, $SD = 2.0$), were recruited in exchange for extra course credit. Only the responses of native speakers of Turkish are included in the analysis. Hence, non-native speakers of Turkish were excluded from the analysis, resulting in 181 participants in total.

3.2.2. DESIGN. A 7-point Likert scale acceptability judgment experiment was designed where participants were asked to read a sentence provided in the middle of the screen and rate it on the given 7-point scale (1: sounds completely unnatural, 7: sounds completely natural). Our linking hypothesis is that speakers will rate grammatical constructions as natural (closer to 7) and ungrammatical constructions as unnatural (closer to 1). Experimental sentences are divided into two groups with a Latin-square design so that each participant saw each item once and only in one condition.

3.2.3. MATERIALS. 2 factors (frequency and agreement), each with two levels, are manipulated on the main verb of the desiderative clauses in the experiment. The factors, frequency (HIGH, LOW) and agreement (3PL, OTHER), are crossed to derive 4 conditions. To manipulate the lexeme frequency of the desiderative verb, half of the experimental items were built with high frequency verbs and the other half were built with low frequency verbs. The high frequency verbs were selected from Aksan et al. (2016), which is based on a 50-million-word corpus. However, since Aksan et al. (2016) listed only the most frequent verbs in Turkish, we consulted Göz (2020) for the low frequency verbs even though it is not as representative (1 million words) as the former.

We checked the frequencies of all candidate verbs in Turkish National Corpus (TNC) and kept the verbs which were above 500 tokens in a million as high frequency verbs ($Mean = 2568.9$) and the verbs which were below 100 tokens in a million as low frequency verbs ($Mean = 16.0$). 24 sentences were created on 12 high frequency and 12 low frequency verbs. Each sentence had two variants (a minimal pair): one with 3PL desiderative form and one with an OTHER (1SG, 1PL, 2SG, 2PL, or 3SG) desiderative form. In addition, we controlled for the sentence length, word type, main predicate and word order across items. All the desiderative verbs were mono-transitive with only one object. 48 sentences (12 for each condition) were created as in (8).

- (8) a. Haftaya bisiklet-i al-ası-ları var-mış.
Next.week bicycle.ACC buy-DESID-3PL.POS exist-EVID
'They feel like buying a bicycle next week.' 3PL X FREQUENT
- b. Haftaya bisiklet-i al-ası-∅ var-mış.
Next.week bicycle-ACC buy-DESID-3SG.POS exist-EVID
'He/she feels like buying a bicycle next week.' OTHER X FREQUENT
- c. Bugün ufaklığ-ı şımart-ası-ları var-mış.
Today little.one-ACC spoil-DESID-3PL.POS exist-EVID
'They feel like spoiling the little one today.' 3PL X INFREQUENT
- d. Bugün ufaklığ-ı şımart-ası-n var-mış.
Today little.one-ACC spoil-DESID-2SG.POS exist-EVID
'He/she feels like spoiling the little one today.' OTHER X INFREQUENT

3.2.4. PROCEDURE. We divided the experimental sentences into two lists with a Latin-square design in order to prevent participants from seeing 2 sentences that differed only in their desiderative agreement marker. Additionally, we prepared 48 filler sentences (24 grammatical and 24 ungrammatical) which were added to both lists. At the beginning of the experiment, participants were randomly assigned one of the two item lists explained in the stimuli section. We had participants do a training session, where they rated 9 sentences with different levels of acceptability, to enable them to use the whole range of options on the 7-point scale before the trial. In the trial session, each participant saw 72 sentences: 24 experimental items (6 from each condition in (8)) and 48 filler items. The experiment was run on PCIBexFarm (Zehr & Schwarz 2022).

3.2.5. RESULTS. All the analyses on the data are conducted by using the free statistical software R (R Core Team 2021) and RStudio (RStudio Team 2020). In Figure 2, the distribution of responses by experimental condition is plotted. The y-axis represents the number of observations for each response and the x-axis represents the responses, which range from 1 to 7 on the Likert-scale used in the experiment. The top panels illustrate the distribution of responses given for the items that are marked with the 3PL suffix (e.g. *V-ası-ları*). The lower panel represents the distribution of responses for the items that are marked with an agreement suffix other than the 3PL.

All the distributions are skewed towards a 7-response, meaning that the participants tend to accept experimental items regardless of the condition. However, when we compare the plots in the top panel and the plots in the lower panel, we observe that 7-responses are much less frequent for the items marked with 3PL than the items marked with another agreement suffix. Moreover, responses lower than 6 are much more frequent for the 3PL items than the OTHER items. Turning to frequency effect on acceptability, we observe that the plots in the left panel and the ones on the right panel do not differ much, implying that frequency is not a determining factor for the acceptability of desiderative clauses. Hence, at first glance, the distribution of raw responses is in line with our main hypothesis that 3PL agreement lowers the acceptability of desiderative constructions. However, for our secondary manipulation, which we did based on the previous findings in the paradigm gap literature, the plots do not show a detectable difference between the acceptability of desiderative verbs with a high frequency lexeme and the ones with a low frequency lexeme.

In addition to the raw responses in Figure 2, we also calculated within-subject standardized (z-score transformed) responses. Within-subject z-scored responses are obtained by subtracting the mean of each subject's responses from the subject's individual responses and then dividing them by the standard deviation of the subject's responses. This transformation is used for reduc-

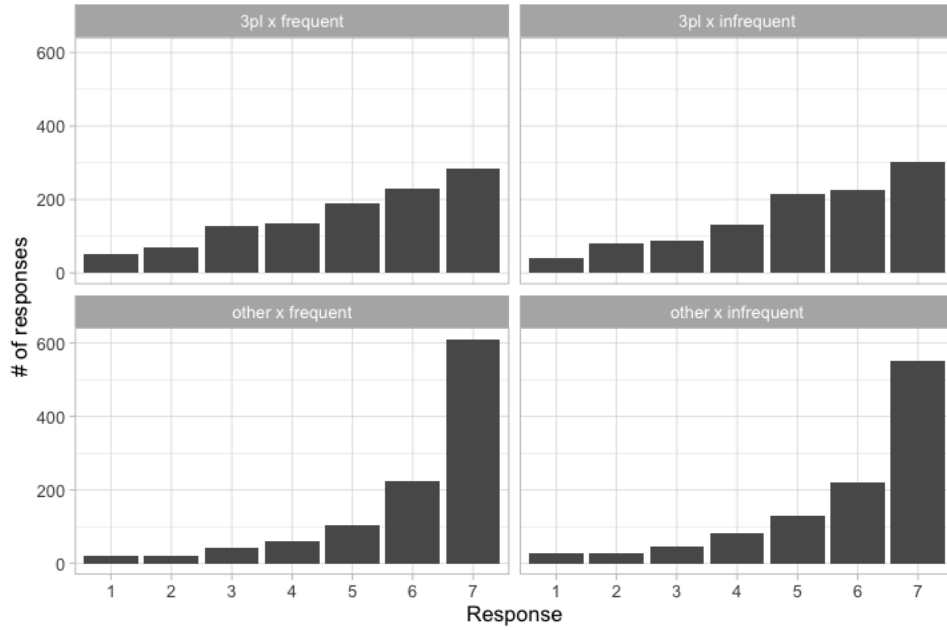


Figure 2. The distribution of the number of Likert-scale (1-7) responses by experimental condition, with 1 being "sounds unnatural" and 7 being "sounds natural".

ing scale bias in Likert-scale responses collected from multiple participants (Schütze & Sprouse 2013). However, z-scored Likert-scale data is shown to be problematic in statistical testing (Lidell & Kruschke 2018, Bürkner & Vuorre 2019). Hence, we used them only in the exploratory data analysis to detect if there are unexpected patterns before proceeding with statistical tests.

Based on our hypothesis, we did not expect any person-number agreement marker other than the 3PL to cause ungrammaticality in desiderative constructions. Therefore, as explained in the Materials section, the items with desiderative verbs bearing agreement markers other than the 3PL are collapsed in the OTHER condition. However, not to introduce any confounds, we wanted to make sure that other agreement markers do not differ with respect to grammaticality in desiderative constructions before moving onto statistical analysis. To this end, we plotted the average within-subject z-scored response for each person-number agreement marker (Figure 3).

Figure 3 plots the average acceptability of the items, in standard deviation units, by the agreement marker in their desiderative suffix bearing verb. A z-scored response lower than 0 means that the item is less acceptable than the average desiderative clause in the experiment. In a theory assuming binary grammaticality, responses lower than 0 can be interpreted as ungrammatical whereas those above 0 can be interpreted as grammatical. As can be seen in Figure 3, 2PL and 3PL are both below 0 and they are virtually equally unacceptable. This is surprising since we predicted only 3PL to be ungrammatical.

We examined the 2PL items closer by a corpus search and find that 2PL desiderative constructions are exclusively used with the Aorist suffix, which adds the meaning of genericity to the sentence (Göksel & Kerslake 2005:296), on the main predicate of the matrix clause. Since each experimental item consists only of an existential matrix predicate and a desiderative clause, and that existential predicates cannot bear the Aorist suffix on their own, our experimental items with 2PL marked desiderative verbs sound unnatural. Given that 2PL desideratives are used and

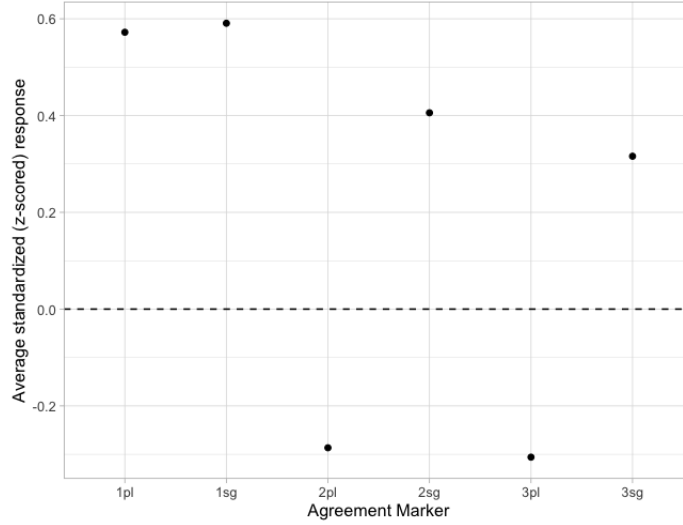


Figure 3. The average of the within-subject standardized responses for each person-number agreement marker. Standardizing has been done only to values of the experimental items, to the exclusion of fillers. The values on the y-axis represent the distance in standard deviation units from the mean, which is marked with 0. The mean represents the average acceptability of the items (desiderative clauses) in the experiment.

accepted by speakers even though 2PL desideratives in the experiment sound odd due to this unforeseen confound, we think it is well motivated to remove 2PL items from the analysis in order to get reliable results while testing our hypothesis.

After removing the experimental sentences with 2PL and their corresponding minimally different sentences that are marked with 3PL, we were left with 19 pairs of sentences out of 24 pairs in total. We modeled the Likert-scale responses as a function of agreement marker type (3PL-OTHER) and frequency (HIGH-LOW) with a maximally mixed ordinal Bayesian model (Bürkner & Vuorre 2019) including by-item and by-subject mixed effects and intercepts (Baayen 2008). We modeled the frequency effect to be fixed by item since frequency was a between-item factor.

The results show that the items in the 3PL condition is -0.91 standard deviations lower on the latent acceptability scale than the items in the OTHER condition. The 95% Credible Interval of this parameter is between -1.22 and -0.61, meaning that there is strong evidence in favor of our hypothesis that 3PL marker on the desiderative bearing verb significantly reduces the acceptability of desiderative constructions. According to the model, there is no main effect of frequency on the acceptability of desiderative clauses ($M = 0.16$, 95%-CI = $[-0.32, 0.65]$); however, there is an interaction effect: there is some evidence –though weak– that 3PL agreement marker reduces acceptability more when the verbal lexeme of a desiderative verb is high-frequency as opposed to when it is low-frequency ($M = -0.34$, 95%-CI = $[-0.73, 0.05]$).

The output of the model in Figure 4 shows the probability of observing a specific response given an experimental sentence. Regardless of frequency, the probability of giving a 7 to desideratives significantly drops when there is a 3PL marker on the desiderative verb. But, the difference between the probabilities of giving a 7 to desideratives in the 3PL condition and OTHER condition is larger when the verbal lexeme has high frequency, meaning that 3PL agreement marker reduces acceptability more when the inflected desiderative form has a high-frequency lexeme.

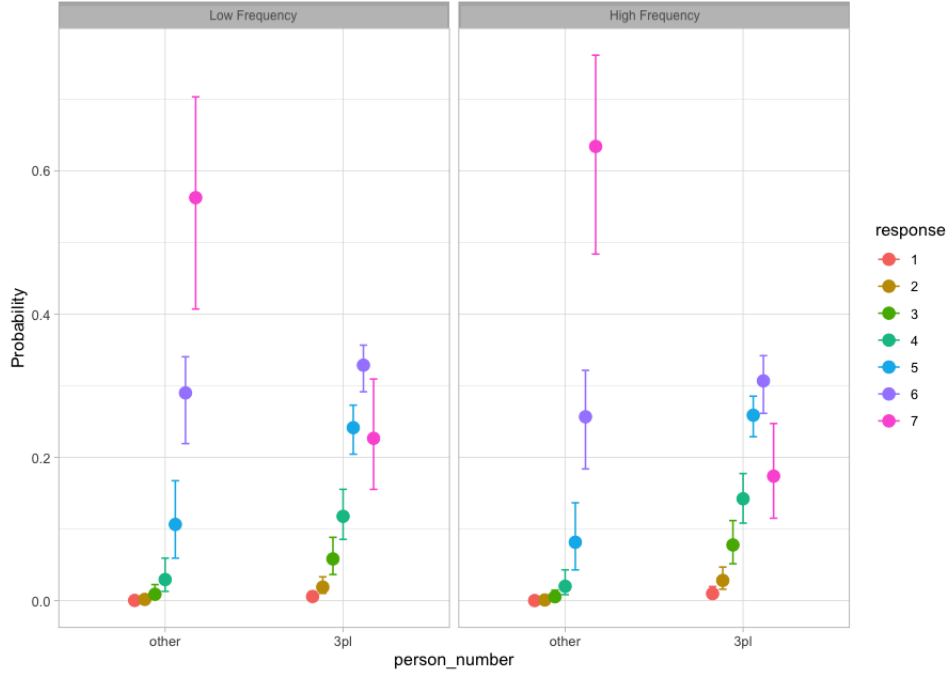


Figure 4. The probability of observing a specific Likert-scale response from 1 to 7 given a condition. Conditions (from left to right): {LOW-FREQUENCY X OTHER}, {LOW-FREQUENCY X 3PL}, {HIGH-FREQUENCY X OTHER}, {HIGH-FREQUENCY X 3PL}. The whiskers represent 95%-CIs.

Finally, the intercept and the effect of 3PL on the acceptability of desideratives significantly vary between participants. The mean of the standard deviation of the intercept across participants is 1.04 with a %95-CI [0.89, 1.21], meaning that the base acceptability of desideratives significantly varies across participants. The mean of the standard deviation of the effect of 3PL is 0.85 (%95-CI = [0.69, 1.03]), suggesting that 3PL agreement marker reduces acceptability more for some participants than for others.

4. Discussion. First, the experiment results suggest that the 3PL does not render a perfectly grammatical desiderative sentence ungrammatical: 3PL desideratives have gradient acceptability, ranging from 1 to 7 on the Likert-scale. There is also a lot of variation in native speaker judgments with regards to 3PL desideratives. However, in spite of the variation and gradience, there is strong evidence that 3PL significantly reduces the acceptability of desiderative constructions systematically at the population level. But does this mean that 3PL desideratives are ungrammatical?

In order to understand why people rate the acceptability of 3PL desideratives low, but not as low as completely ungrammatical sentences, we looked at the ratings of the filler items (each participant saw 24 grammatical and 24 ungrammatical fillers). When we probed into the ratings of the fillers, we noticed that the tendency to rate ungrammatical sentences higher on the scale is much more than the tendency to rate grammatical sentences lower (Figure 5).

Figure 5 illustrates that participants are indecisive when it comes to desiderative clauses: they dislike even the desiderative clauses in OTHER conditions that should be acceptable based on native speaker intuitions. This finding can be best explained by frequency effect. Studies investi-

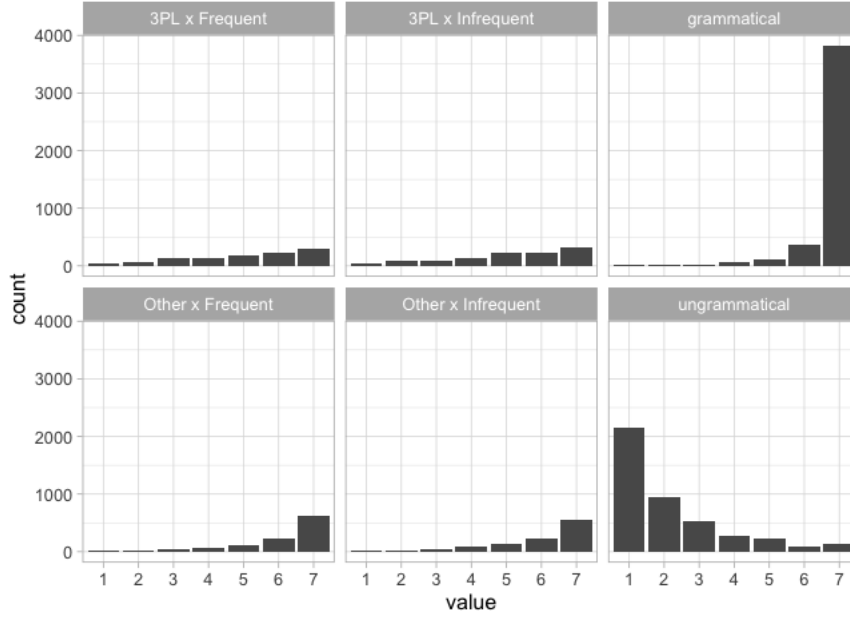


Figure 5. Distribution of responses for all sentences in the experiment including fillers, which are of two types: grammatical and ungrammatical.

gating the relationship between acceptability scores and corpus frequency have shown that more probable forms of a construction are rated higher than less probable forms (Dabrowska 2008, Divjak 2017, Bermel & Knittl 2012). The corpus results show that desiderative clauses are rare overall and when they are used, they predominantly occur in the 1SG and 3SG forms. Since the items in the OTHER condition also include 1PL, 2SG, and 2PL desiderative clauses, we can explain the low acceptability ratings of desideratives by the frequency effect. Hence, desiderative constructions are grammatical overall even though their acceptability scores are lower due to the frequency effect. However, when they occur with the 3PL agreement marker, their acceptability reduces significantly, which cannot be alluded simply to the frequency effect. Thus, we argue that there is enough evidence for the defectivity of 3PL desiderative forms.

Returning to the anomalous frequency distribution of 3PL forms in the desiderative paradigm reported in Table 3, it suggests that speakers avoid using 3PL desideratives, which is not very difficult given that (i) overt 3PL agreement is obligatory when the 3PL pronoun is dropped (Göksel & Kerslake 2005:117) and (ii) desiderative constructions are not very frequent overall (7001 tokens out of 491 million tokens (Bilgin 2016:149)) in Turkish. Hence, speakers can easily avoid using 3PL desideratives, which reduces the amount of 3PL desiderative in the input even further. Thus, when speakers end up in a situation where they need to produce a 3PL desiderative on the fly, it is a real life wug-test for most of them. Even if they can come up with either *yapaları* or *yapasıları* type 3PL desideratives in these circumstances, the feeling of uncertainty about the wellformedness of these forms would be unavoidable.

There are various theories that have been proposed to explain the kind of uncertainty attested when producing defective forms like -AsI desideratives. Some propose that defectivity is caused by the absence of a productive (i.e. elsewhere) rule to produce the form in the gapped cell (Albright & Hayes 2003; Yang 2016). For instance, Albright (2003) argues that defectivity is an

end result of low speaker confidence, which is observed when there is no productive rule to inflect especially infrequent stems. This account predicts that low frequency lexemes will be more defective than high frequency lexemes in the absence of a productive rule to inflect them. However, our results suggest the opposite: frequent lexemes are more defective when inflected for 3PL desiderative features. We argue that this finding is best explained by negative evidence accounts such as the one proposed by Daland et al. (2007): frequent lexemes are expected to occur more in the defective cells than infrequent lexemes. Every time they occur in the non-defective cells but not in the defective cells, negative evidence accumulates for the cells' defectiveness. However, since infrequent lexemes are rarely found in both defective and non-defective cells, speakers do not receive much negative evidence for their defectiveness when they do not occur in the defective cell. Thus, negative evidence accounts predict low-frequency lexemes to have higher acceptability than high-frequency lexemes in defective cells, which is what our results suggest.

However, for the desiderative gap, it is very likely that both competition based (Yang 2016; Albright & Hayes 2003) and negative evidence based explanations (Daland et al. 2007) are at play. The fact that there are two possible –though imperfect– forms for 3PL desideratives provides strong evidence that these two forms compete in the minds of the speakers. Although the winner has a degraded level of acceptability for a speaker, there is a winner for each and every speaker and the winner varies depending on the speaker. This observation strongly suggests that there is a competition between the two forms. Moreover, evidence for a negative evidence based account comes from the anomalous relative frequency distribution in the corpus and the reduced acceptability of frequent lexemes inflected for 3PL desiderative features. Since 3PL desideratives have a drastically lower frequency than expected, speakers can pick up on this anomaly, in line with the findings in the research on statistical learning (Aslin et al. 1998, Maye et al. 2002).

5. Conclusion. Based on native speakers intuitions, we proposed that third person plural agreeing forms of verbs that bear the desiderative suffix are ungrammatical. We proposed that this gap is due to the irregularity in the third person singular agreeing desiderative verbs and the anomalous frequency distribution of the defective forms. We argued that there are (at least) two different likely decompositions that speakers can posit for the complex suffix *-AsI* in 3SG desideratives. We demonstrated how different decompositions lead to two conflicting hypotheses regarding the form of the 3PL desideratives: while some speakers prefer *X+AsI+lArI* type, some others prefer *X+A+lArI* type as the phonological output of the 3PL desideratives. However, speaker intuitions about the acceptability of the possible forms of 3PL desideratives show that they are uncertain about the well-formedness of even the form they prefer. This, we argue, is due to their implicit knowledge that there are two competing forms for 3PL desideratives based on different parsings of 3SG desideratives and that there is no conclusive evidence for the grammaticality of one form or the ungrammaticality of the other.

Evidence from corpus frequencies and an acceptability judgment experiment ($N = 181$) support the hypothesis that third person plural agreeing desiderative verbs have significantly lower acceptability ratings compared to desiderative forms marked with other person-number agreement suffixes. Moreover, the experiment results also suggest that there is some evidence that 3PL desiderative forms of high-frequency lexemes are more defective than 3PL desiderative forms of low-frequency lexemes. Although further experimentation is needed to test the robustness of this effect, we argue that this finding provides evidence that speakers are sensitive to negative evidence in the input. Finally, 3PL desideratives are not completely unacceptable, which sug-

gests that the gap is not an instance of complete ungrammaticality, but instead it is an instance of speaker uncertainty about the correctness of the form, similar to what Albright (2003) argues to be the case for Spanish gaps. Overall, our results suggest that speakers are implicitly aware of both the anomalous frequency distribution and the existence of conflicting hypotheses for the grammatical form of 3PL desideratives. We argue that both the absence of a reliable rule and the atypical frequency distribution contribute to the existence and persistence of the 3PL desiderative gap in Turkish.

References.

- Aksan, Yeşim, Mustafa Aksan, Ümit Mersinli & Umut Ufuk Demirhan. 2016. *A frequency dictionary of turkish (1st ed.)*. Routledge. <https://doi.org/10.4324/9781315733302>.
- Albright, Adam. 2003. A quantitative study of spanish paradigm gaps. In *WCCFL 22: Proceedings of the 22nd West Coast Conference on Formal Linguistics*, Cascadia Press. <https://web.mit.edu/albright/www/papers/Albright-WCCFL22.pdf>.
- Albright, Adam & Bruce Hayes. 2003. Rules vs. analogy in english past tenses: A computational/experimental study. *Cognition* 90. 119–61. [https://doi.org/10.1016/S0010-0277\(03\)00146-X](https://doi.org/10.1016/S0010-0277(03)00146-X).
- Aslin, Richard N., Jenny R. Saffran & Elissa L. Newport. 1998. Computation of conditional probability statistics by 8-month-old infants. *Psychological Science* 9(4). 321–324. <https://doi.org/10.1111/1467-9280.00063>.
- Baayen, R. H. 2008. *Analyzing linguistic data: A practical introduction to statistics using r*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511801686>.
- Bader, Markus & Jana Häussler. 2010. Toward a model of grammaticality judgments. *Journal of Linguistics* 46(2). 273–330. <https://doi.org/10.1017/S0022226709990260>.
- Baerman, Matthew, Greville G. Corbett & Dunstan Brown. 2010. *Defective Paradigms: Missing Forms and What They Tell Us*. British Academy. <https://doi.org/10.5871/bacad/9780197264607.001.0001>.
- Bermel, Neil & Luděk Knittl. 2012. Corpus frequency and acceptability judgments: A study of morphosyntactic variants in czech. *Corpus Linguistics and Linguistic Theory* 8(2). 241–275. <https://doi.org/10.1515/cllt-2012-0010>.
- Bilgin, Orhan. 2016. *Frequency effects in the processing of morphologically complex turkish words*: MA thesis. <https://doi.org/10.13140/RG.2.2.20856.44809>.
- Bürkner, Paul-Christian & Matti Vuorre. 2019. Ordinal regression models in psychology: A tutorial. *Advances in Methods and Practices in Psychological Science* 2(1). 77–101. <https://doi.org/10.1177/2515245918823199>.
- Dabrowska, Ewa. 2008. The effects of frequency and neighbourhood density on adult speakers' productivity with polish case inflections: An empirical test of usage-based approaches to morphology. *Journal of Memory and Language* 58(4). 931–951. <https://doi.org/10.1016/j.jml.2007.11.005>.

- Daland, Robert, Andrea D. Sims & Janet Pierrehumbert. 2007. Much ado about nothing: A social network model of Russian paradigmatic gaps. In *Proceedings of the 45th annual meeting of the association of computational linguistics*, 936–943. Prague, Czech Republic: Association for Computational Linguistics. <https://aclanthology.org/P07-1118>.
- Divjak, Dagmar. 2008. *On (in)frequency and (un)acceptability*. 213–233. Lodz Studies in Language Peter Lang Verlag.
- Divjak, Dagmar. 2017. The role of lexical frequency in the acceptability of syntactic variants: Evidence from that-clauses in polish. *Cognitive Science* 41(2). 354–382. <https://doi.org/10.1111/cogs.12335>.
- Gorman, Kyle & Charles Yang. 2019. *When nobody wins* 169–193. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-02550-2_7.
- Göksel, Aslı & Celia Kerslake. 2005. *Turkish: A comprehensive grammar (1st ed.)*. Routledge. <https://doi.org/10.4324/9780203340769>.
- Göz, İlyas. 2020. *Yazılı türkçenin kelime sıklığı sözlüğü* Türk Dil Kurumu Yayınları. Türk Dil Kurumu Yayınları.
- Liddell, Torrin & John Kruschke. 2018. Analyzing ordinal data with metric models: What could possibly go wrong? *Journal of Experimental Social Psychology* 79. 328–348. <https://doi.org/10.1016/j.jesp.2018.08.009>.
- Maye, Jessica, Janet F Werker & LouAnn Gerken. 2002. Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82(3). B101–B111. [https://doi.org/10.1016/S0010-0277\(01\)00157-3](https://doi.org/10.1016/S0010-0277(01)00157-3).
- R Core Team. 2021. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna, Austria. <https://www.R-project.org/>.
- RStudio Team. 2020. *Rstudio: Integrated development environment for r*. RStudio, PBC. Boston, MA. <http://www.rstudio.com/>.
- Schütze, Carson & Jon Sprouse. 2013. Judgment data. In Devyani Sharma Robert J. Podesva (ed.), *Research methods in linguistics*, chap. 3, 27–50. Cambridge: Cambridge University Press. <https://doi.org/10.1017/cbo9781139013734.004>.
- Sims, Andrea D. 2006. *Minding the gaps: Inflectional defectiveness in a paradigmatic theory*: dissertation. http://rave.ohiolink.edu/etdc/view?acc_num=osu1157550938.
- Sims, Andrea D. 2015. *Inflectional defectiveness* (Cambridge Studies in Linguistics). Cambridge University Press. <https://doi.org/https://doi.org/10.1017/CBO9781107053854>.
- Yang, Charles. 2016. *The Price of Linguistic Productivity: How Children Learn to Break the Rules of Language*. The MIT Press. <https://doi.org/10.7551/mitpress/10842.001.0001>.
- Zehr, Jérémy & Florian Schwarz. 2022. Penncontroller for internet based experiments (ibex). <https://doi.org/10.17605/OSF.IO/MD832>.