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, which results in an unexpected frequency distribution. In addition, the acceptability judgment experiment ($N = 181$) shows that the third person plural suffix significantly reduces the acceptability of -AsI desideratives, corroborating the findings from the corpus. 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 account is supported by the presence of two competing forms whereas negative evidence account is supported by the anomalous relative frequency distribution of desideratives.

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

1. Introduction. Productivity is generally 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 rules of their language. They can readily inflect words they have never seen before for 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– a paradigm gap in the agreement paradigm of a construction we call *-AsI desideratives* in Turkish. -AsI desideratives are ungrammatical for most speakers of Turkish when inflected for third person plural agreement. Namely, there are two possible forms for 3PL desideratives that are acceptable to some speakers to varying degrees. However, for most speakers of Turkish neither of the forms exemplified in (1)[1] are acceptable.

(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 subject pronouns can be dropped in Turkish (Göksel & Kerslake 2005). What differs between the two sentences is the form of the desiderative suffix. Some people prefer (1-a) over (1-b) while some prefer (1-b) over

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

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

(1-a). Nevertheless, most speakers tend to reject both forms, meaning that there is a gap in the 3PL cell of the agreement paradigm of -AsI desideratives.

In what follows, we proceed with some background on -AsI desideratives and their agreement paradigm 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 morphological theories. Finally, section 5 concludes the paper.

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

Desiderative clauses can be taken as complement by only a small set of predicates such as *gel-* 'come', *git-* 'go', *kaç-* 'go.away' and *tut-* 'hold' in addition to the existential predicates *var* 'exist' and *yok* 'not exist'. The sentence as a whole conveys 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.

2.1. AGREEMENT PARADIGM. 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 suffixes from the possessive agreement paradigm on the verb. However, the desiderative constructions come with an exceptional pattern. Typically, the third person singular possessive suffix is -sI, which can be seen in a canonical example of nominalization with -mA in (3) and in a Genitive-Possessive construction in (4).

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

(4) Merve'nin tatlı-**sı**
Merve-GEN dessert-**3SG.POS**
'Merve's dessert'

However, we do not observe an additional -sI 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 (5).

(5) Merve-nin tatlı yap-**ası** var.
Merve-GEN dessert do-**DESID.3SG.POS** exist

'Merve feels like making a dessert.'

2.2. THE SOURCE OF 2 DIFFERENT 3PL FORMS. The complex suffix -AsI in third person singular agreeing desideratives such as in (5) can be decomposed into DESID and 3SG in two different ways. The first one is that -A is the allomorph of DESID and -sI 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 desiderative constructions (6-b).

(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. Nevertheless, the meaning of the 3SG desideratives is straightforward even without decomposing -AsI into smaller components. The only knowledge speakers need is that verbs bearing the desiderative suffix -AsI have to be inflected for person-number agreement. If the verb does not have an overt agreement marker right next to -AsI, then it is interpreted to agree with a 3SG subject. 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. In addition, not decomposing the 3SG desiderative suffix -AsI would not be problematic also for the comprehension and production of first person and second person agreeing desideratives since they are regular suffixes coming from the possessive paradigm.

When it comes to the effect of not decomposing the 3SG desiderative suffix -AsI on the comprehension and production of 3PL desideratives, non-decomposition approach becomes problematic. Leaving the complex -AsI as a whole unit without parsing it into two separate suffixes, we cannot know how to form 3PL desideratives. If speakers decompose *yapası* as in (6-a), there is no evidence for 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. Let's take the derivation of the root *yap* 'do' as an example for demonstration. If speakers adopt the parsing in (6-b), they would produce the output only in (7-b) for third person plural agreeing desiderative form of *yap*. However, if speakers adopt the parsing in (6-a), they can produce both forms in (7). If they decide that DESID allomorphy only occurs in the environment of 3SG and the exponent of DESID is -AsI in the environment of 3PL, then they would inflect *yap* as in (7-b) for DESID and 3PL features. However, if speakers hypothesize that DESID allomorphy is observed in the environment of third person feature regardless of the number feature, the form in (7-a) would be derived.

(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 since 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 the corpus search we ran. Thus, speakers need to parse the complex suffix -AsI in third person singular desiderative constructions into 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 with no definitive solution. 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. Hence, there is a systematic paradigm gap in the 3PL cell of the agreement paradigm of desideratives in Turkish.

In the next section, we bring in evidence from a corpus search and an acceptability judgment experiment to support 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 -AsI suffix due to the fact that it is often mistaken for the optative suffix -A and the third person singular suffix -sI, which happens to be the most common suffix in Turkish (Bilgin 2016). For 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).

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 desiderative 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 1), it can be observed that the relative frequency of 3PL desideratives

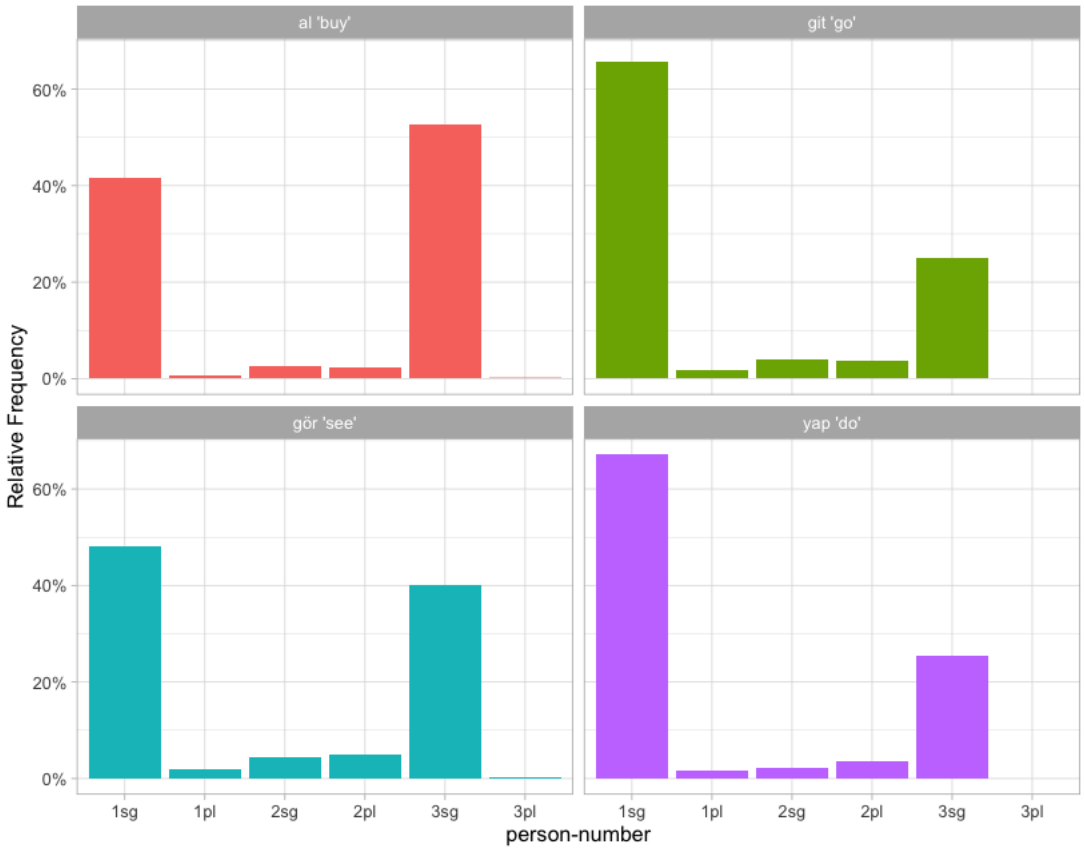

Figure 1. Relative frequencies of the desideratives of 4 frequent verbs with respect to their person-number agreement markers. 1SG and 3SG forms dominate the frequencies while only 2 instances of 3PL forms (*alasıları* and *göresileri*) are attested.

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, p.226).

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

Table 1. 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 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 stem of third plural agreeing desiderative verbs affects their grammaticality. Even though the intuitions of our native speaker consultants, including ourselves, do not predict a frequency effect on the grammaticality of third plural agreeing desiderative verbs, several studies in the paradigm gap literature have reported that low frequency stems reduce acceptability where there is a paradigm gap (Albright 2003, Sims 2006). Hence, we also test in this experiment if low frequency reduces the acceptability of third person plural agreeing desiderative verbs.

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, 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 in total. To manipulate the frequency of the main verb of the desiderative clause, 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 to 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 is manipulated to have two variants (a minimal pair): one with 3PL desiderative form and one with OTHER (1SG: 5 + 1PL: 4 + 2SG: 4 + 2PL: 5 + 3SG: 6 = 24) desiderative form. In addition, we controlled 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 3PL X FREQUENT

b. Haftaya bisiklet-i al-ası-∅ var-mış.
Next.week bicycle-ACC buy-DESID-3SG.POS exist-EVID OTHER X FREQUENT

c. Bugün ufaklığ-ı şımart-ası-ları var-mış.
Today little.one-ACC spoil-DESID-3PL.POS exist-EVID 3PL X INFREQUENT

d. Bugün ufaklığ-ı şımart-ası-n var-mış.
Today little.one-ACC spoil-DESID-2SG.POS exist-EVID OTHER X INFREQUENT

3.2.4. PROCEDURE. We divided 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 markers. 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. They first did a training session in which they rated 9 sentences that were constructed to have different levels of acceptability in order to enable them to use the whole range of options on the 7-point scale before starting 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 illustrates 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 high frequency and low frequency desiderative verbs.

In addition to the raw responses in Figure 2, we have also made use of within-subject standardized (z-score transformed) responses. Within-subject z-score responses are obtained by subtracting the mean of each subject from that subject's responses and then dividing them from the standard deviation of that subject's responses. This transformation is used for reducing scale bias in Likert-scale responses collected from multiple participants (Schütze & Sprouse 2013). However, z-score transforming Likert-scale data is shown to be problematic in statistical testing (Liddell & Kruschke 2018, Bürkner & Vuorre 2019). Hence, we used them only in the exploratory

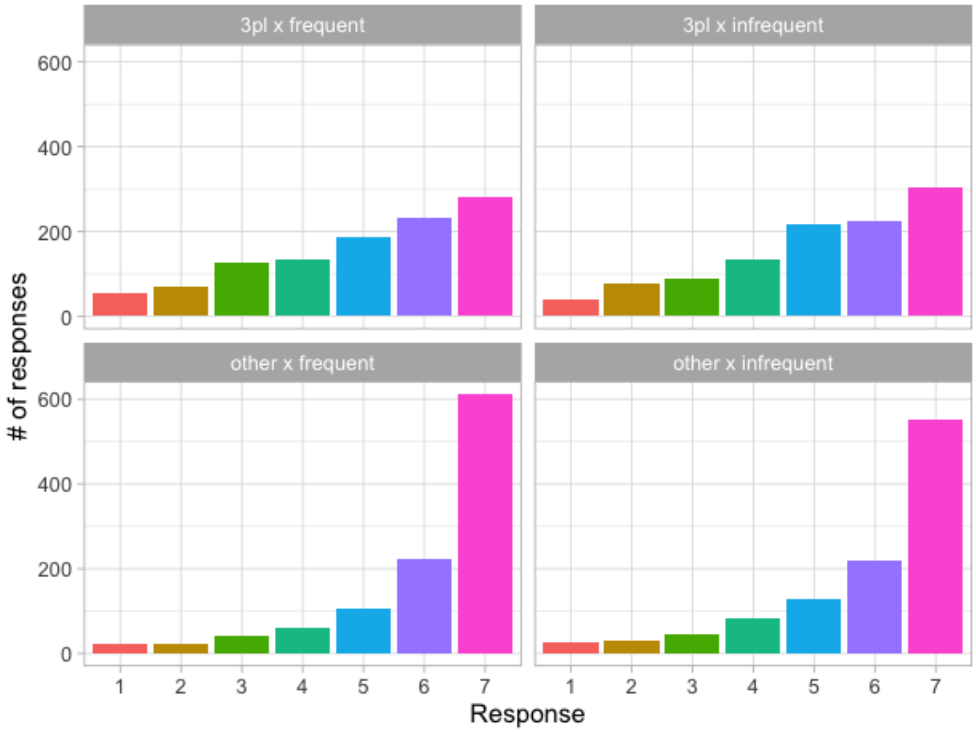

Figure 2. The distribution of the number of Likert-scale (1-7) responses by experimental condition, with 1 being "sounds unnatural" and 7 being "sounds natural".

data analysis to detect unexpected patterns, if any, 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 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 found that 2PL desiderative constructions are exclusively used with the Aorist suffix on the main predicate of the matrix clause. This adds the meaning of genericity to the sentence as in the case of generic *you* sentences in English. Since our experimental items consist only of an existential matrix predicate and the desiderative clause, and since existential predicates cannot bear the Aorist suffix on their own, our experimental items with 2PL marked desiderative verbs sound unnatural. Given that 2PL

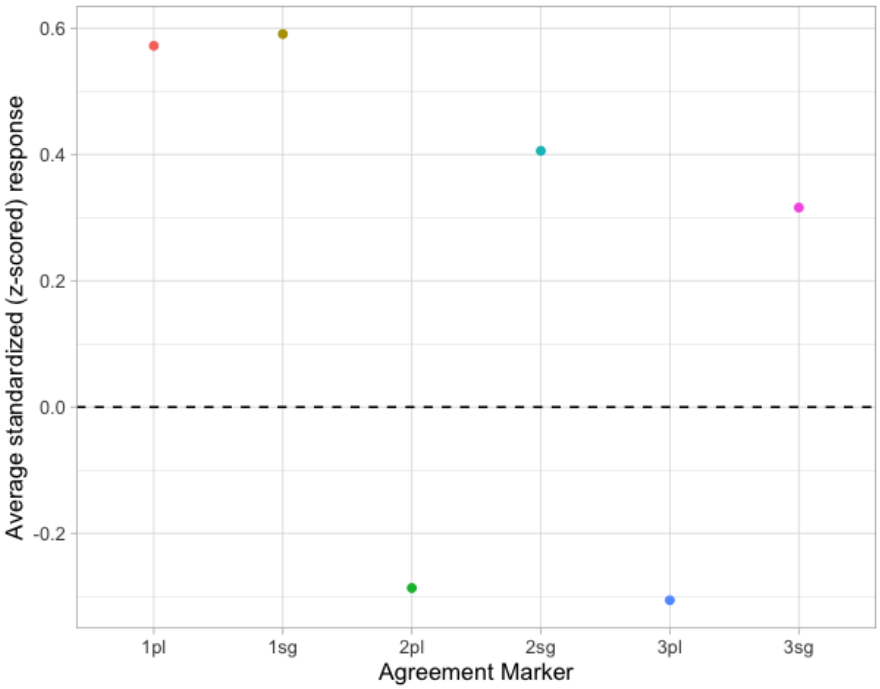

Figure 3. The average of the within-subject standardized responses for each person-number agreement marker. Standardizing has been done only to values of the experimental items, to the exclusion of fillers. The values on the y-axis represent the distance in standard deviation units from the mean, which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

desideratives are used and 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 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 ran an ordinal Bayesian model (Bürkner & Vuorre 2019) with by-item and by-subject mixed effects and intercepts. We modeled the frequency effect to be fixed by item since frequency was a between-item factor. Other than that, the model is a maximal mixed model (Baayen 2008).

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 -1.21 and -0.59, meaning that there is strong evidence in favor of our hypothesis that third person plural agreement 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 $= 0.19$, 95%-CI $= [-0.30, 0.69]$). For the interaction effect, the model found some evidence that there is an interaction between frequency and person-number agreement (M $= -0.37$, 95%-CI $= [-0.77, 0.02]$).

Figure 4 reports the output of the model which shows the probability of observing a specific response given a sentence belonging to an experimental condition. It can be seen that the

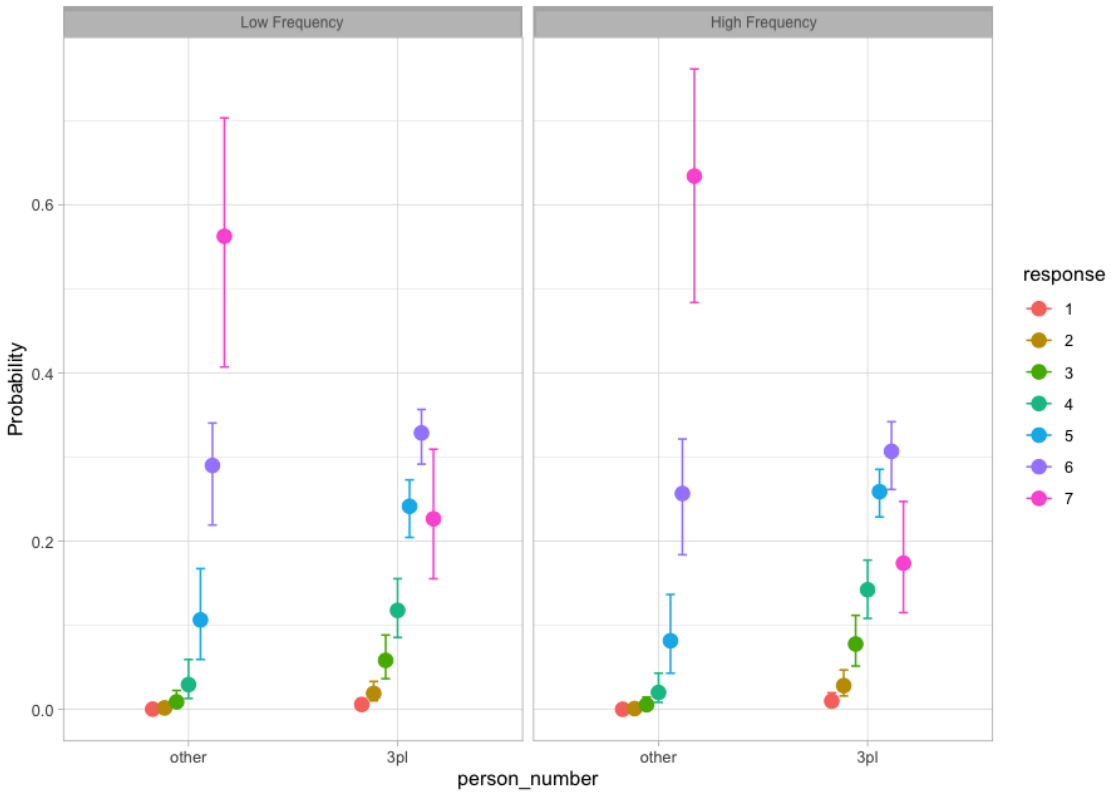

Figure 4. The probability of one of the four experimental conditions. The plots represent the probability of observing a specific Likert-scale response from 1 to 7 given a condition. Conditions (from left to right): {LOW-FREQUENCY X OTHER}, {LOW-FREQUENCY X 3PL}, {HIGH-FREQUENCY X OTHER}, {HIGH-FREQUENCY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

probability of giving a 7 to desideratives significantly drops regardless of frequency when the desiderative verb is marked with 3PL. Nevertheless, it can also be observed that the gap between the probabilities of giving a 7 response is wider with high frequency verbal stems that bear the desiderative suffix than with low frequency verbal stems. This stems from the negative effect of the interaction between frequency and agreement marking on the acceptability of desiderative clauses.

Finally, the intercept and the effect of 3PL on acceptability of desideratives vary a lot between participants. The mean of the standard deviation of the intercept across participants is 1.05 with a %95-CI $[0.91, 1.22]$. This means that the base acceptability of desiderative clauses varies a lot across participants. For the standard deviation of the effect of 3PL, its mean is 0.87 (%95-CI $= [0.71, 1.04]$). This result suggests that 3PL reduces acceptability more for some participants than for the 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 used in the experiment. 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 desider-

ative constructions systematically at the population level. But does this mean 3PL desideratives are ungrammatical?

In order to understand why people rate 3PL low, but not as low as completely ungrammatical sentences, we looked at the ratings of filler items. To remind you, every participant saw 24 grammatical and 24 ungrammatical filler sentences. 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).

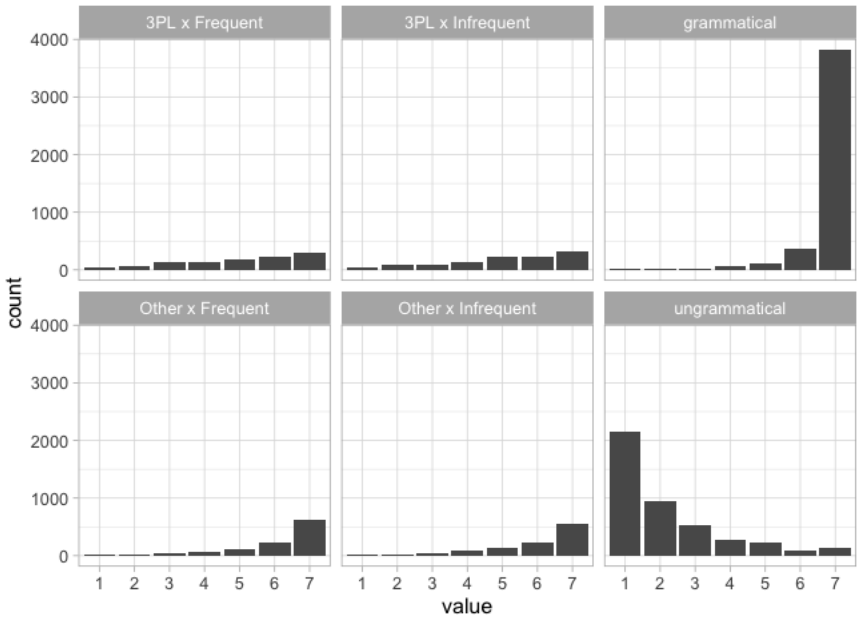

Figure 5. Distribution of responses for all sentences in the experiment including fillers, which are of two types: grammatical and ungrammatical.

Figure 5 illustrates that participants are indecisive when it comes to desiderative clauses: They dislike even those desiderative clauses in OTHER conditions that should be acceptable based on native speaker intuitions. This finding can be best explained by frequency effect. Studies investigating 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, it is predominantly the 1SG and 3SG forms that are attested. 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.

The anomalous frequency distribution of 3PL forms relative to OTHER forms in the desiderative paradigm suggests that speakers avoid using it. Since overt 3PL agreement is obligatory only in pro-drop environments and desiderative suffix is not very frequent overall (7001 out of 491 million tokens (Bilgin 2016:p.149), speakers have the means to avoid using 3PL desideratives. When they are faced with an online production task in spontaneous speech, it is a real life wug-

test for most speakers. In these circumstances, even if they can come up with either *yapaları* or *yapasıları* type 3PL desideratives (or most probably both), they feel uncertain regarding the well-formedness of these outputs.

There are some accounts proposed in the gap literature to explain the kind of uncertainty attested when producing -AsI desideratives. Some researchers propose that there is no productive rule (elsewhere) to produce the form in the gapped cell (Albright & Hayes 2003, Yang 2016) while some others suggest that speakers are sensitive to frequency information and hence avoid using the form in the gap even though they can produce the form and there is no justified phonological or semantic reason to not to do so (Daland et al. 2007). Thus, previous research shows that not every gap is the same and that some gaps are explained by one account while some are explained by another (Sims 2015, Baerman et al. 2010).

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 suggest that there is a competition. On the other hand, evidence for a negative evidence based account comes from the skewed relative frequency distribution attested in the corpus. Since 3PL desideratives have a drastically lower frequency than expected, speakers are highly likely to pick up on this anomaly –in line with the research in statistical learning (Aslin et al. 1998, Maye et al. 2002)– to avoid using the form. Hence, this process would also lead to the persistence of the desiderative gap in Turkish.

5. Conclusion. Based on native speakers intuitions, we proposed that third person plural agreeing forms of desiderative bearing verbs are ungrammatical. We proposed that this gap is a result of the irregularity in the third person singular agreeing desiderative verbs. 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 V+AsIları type, some others prefer V+Aları type as the phonological output of the 3PL desideratives. However, the 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. However, 3PL desideratives are not completely unacceptable either. This provides evidence 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 the existence of conflicting hypotheses for the grammatical form of 3PL desideratives. All in all, the 3PL desiderative gap in Turkish is compatible with both negative evidence accounts and competing forms accounts. In the absence of conclusive evidence favoring one hypothesis over the

other, we suggest 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: Core vocabulary for learners*. Routledge. https://doi.org/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*, Cascadilla Press.

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.

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. 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 judgments1. *Journal of Linguistics* 46(2). 273–330.

Baerman, Matthew, Greville Corbett & Dunstan Brown. 2010. Defective paradigms: Missing forms and what they tell us .

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/doi:10.1515/cllt-2012-0010. 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. https://doi.org/10.1177/2515245918823199.

Daland, Robert, Andrea D. Sims & Janet B. Pierrehumbert. 2007. Much ado about nothing: A social network model of russian paradigmatic gaps. In *Acl*, .

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. https://doi.org/10.1111/cogs.12335.

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/https://doi.org/10.1016/j.jml.2007.11.005. https://www.sciencedirect.com/science/article/pii/S0749596X07001271.

Göksel, Aslı & Celia Kerslake. 2005. *Turkish: A comprehensive grammar*. Routledge.

Gorman, Kyle & Charles Yang. 2019. *When nobody wins* 169–193. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-02550-2$_7$. https://doi.org/10.1007/978-3-030-02550-2_7.

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/https://doi.org/10.1016/S0010-0277(01)00157-3. https://www.sciencedirect.com/science/article/pii/S0010027701001573.

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*. The Ohio State University.

Sims, Andrea D. 2015. *Inflectional defectiveness* Cambridge Studies in Linguistics. Cambridge University Press. 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. The MIT Press. https://doi.org/10.7551/mitpress/9780262035323.001.0001.

Zehr, Jérémy & Florian Schwarz. 2022. Penncontroller for internet based experiments (ibex). https://doi.org/10.17605/OSF.IO/MD832. osf.io/md832.