

## Processing Turkish case markers: Implications for the case containment hypothesis

Metehan Oğuz & Elsi Kaiser\*

**Abstract.** We investigate the processing of accusative-marked NPs in Turkish, compared to genitive- and locative-marked NPs. We use psycholinguistic methods (a lexical decision task) to test competing predictions derived from frequency effects in lexical processing on the one hand, and morphosyntactic theories of case containment on the other hand. Our experimental outcomes support a case containment approach to Turkish morphology.

**Keywords.** word processing; case marking; lexical decision; case containment

**1. Introduction.** One growing area in the syntax/morphosyntax literature is research on the case containment hypothesis, according to which some morphosyntactic cases are in a structural (hierarchical) containment relationship (Caha 2009, 2019; Moska & Smith 2015; Zompì 2017; Smith et al. 2019; Davis 2023; a.o.). Very briefly, the main claim of the case containment hypothesis is that structural (morphosyntactic) case is internally complex, and more complex cases structurally contain the less complex ones. Caha (2009) argues for the containment hierarchy in (1), where accusative case structurally contains nominative, and is structurally contained by genitive. Similarly, genitive is contained by dative and so on.

(1) [ [ [ [ [ [ NOM ] ACC ] GEN ] DAT ] INSTR ] COM ]

The main evidence for the case containment hypothesis comes from a cross-linguistic ban on ABA patterns (\*ABA) on case-driven suppletion patterns, and cross-linguistic case syncretisms, i.e., two distinct cases surfacing with the same morphological form.

A different containment hierarchy as in (2), following dependent case theory (Baker 2015), is proposed by Smith et al. (2019), where dependent cases (e.g. accusative) structurally contain unmarked cases (e.g. nominative) and are contained by oblique cases (e.g. locative).

(2) [ [ [ UNMARKED ] DEPENDENT ] OBLIQUE ]

Under both views, the proposed containment is on the abstract/structural level, and is not necessarily represented in the surface (morphological) form. For example, according to the case containment hypothesis, the English pronoun *us* structurally contains its nominative counterpart *we*, even though *us* does not morphologically (or phonologically) contain *we* (e.g. Caha 2009). However, recent studies unearth interesting data from languages like Khanty, Kalderaš Romani (Smith et al. 2019), Turkish (Türk & Caha 2021), and Balkar (Davis 2023), where structural containment is proposed to be overtly reflected in the morphological structure of the language. Davis (2023) proposes that the locative and ablative cases in (Karachay-)Balkar (Turkic, Russia) morphologically contain accusative case in a specific morphological environment. Davis demonstrates that oblique case markers in Balkar undergo a phonological change when they appear

---

\*We thank Zuzanna Fuchs, Travis Major, the audience at the 2024 LSA Annual Meeting, and the members of USC Psycholinguistics Lab group for their valuable and constructive feedback. More information about the experimental stimuli is available upon request from the first author. Authors: Metehan Oğuz, University of Southern California ([moguz@usc.edu](mailto:moguz@usc.edu)) & Elsi Kaiser, University of Southern California ([emkaiser@usc.edu](mailto:emkaiser@usc.edu)).

after the third person possessee suffix *-sI*. Table 1 summarizes his observations (from Davis 2023: 3).

	NOM	ACC	LOC	ABL	DAT	GEN
TYPICAL	-∅	-nI	-dA	-dAn	-nA/-gA	-nI
POSSESSED	-∅	-n	-ndA	-ndAn	-nA/-gA	-nI

Table 1. Balkar case morphology

More concretely, Davis shows that the regular accusative marker *-nI* shows up in the form *-n* in possessed NPs when it appears after the third person possessee suffix. He also shows that the locative (*-dA*) and ablative (*-dAn*) markers are preceded by an *n* when they are attached to a possessed NP, surfacing as (*-ndA*) and (*-ndAn*), respectively. Davis proposes that this *n* preceding these suffixes is indeed the accusative marker *-nI*, surfacing as *-n* after possessed NPs. Davis suggests that this is an overt (morphological) realization of case containment, where oblique cases like locative and ablative contain accusative case, and he argues that Balkar data provides strong evidence for the case containment hypothesis.

The current study shifts the focus to a highly related language: we experimentally tests the predictions of morphological containment analyses for Turkish, and explore the implications of Davis’ (2023) analysis of Balkar for Turkish data.

1.1. TURKISH CASE MORPHOLOGY. Table 2 shows the standard view of Turkish morphology. The regular markers are observed with the root *ev* ‘house’. Since Turkish does not allow vowel clusters, an epenthetic *y* is inserted before accusative and dative markers with roots like *kedi* ‘cat’. Also, similar to the Balkar data discussed in Davis (2023), whenever a case marker appears after the third person possessee marker *-(s)I*, an *n* is inserted between the suffix and the case marker. Interestingly, the same *n* appears between the vowel-final word *kedi* and genitive marker as well.

Case	‘house’	‘cat’	‘cat.3POSS’
NOM	<i>ev</i>	<i>kedi</i>	<i>kedi-si</i>
ACC	<i>ev-i</i>	<i>kedi-yi</i>	<i>kedi-si-ni</i>
GEN	<i>ev-in</i>	<i>kedi-nin</i>	<i>kedi-si-nin</i>
DAT	<i>ev-e</i>	<i>kedi-ye</i>	<i>kedi-si-ne</i>
LOC	<i>ev-de</i>	<i>kedi-de</i>	<i>kedi-si-nde</i>
ABL	<i>ev-den</i>	<i>kedi-den</i>	<i>kedi-si-nden</i>

Table 2. Turkish morphological system in standard terms

Türk & Caha (2021) provide a case containment analysis of Turkish, highlighting the overlap in accusative and genitive suffixes, and decompose the genitive *-In* into accusative *-I* and genitive *-n*. They propose that the genitive in Turkish overtly (morphologically) contains the accusative, providing evidence for the case containment hypothesis. They note that the morphological containment is not observed with vowel-final roots (e.g. *kedi*), because an epenthetic *y* is inserted before the accusative marker, while an *n* is inserted before the genitive marker with a vowel-final root like *kedi* ‘cat’. Türk and Caha do not discuss the containment possibilities of other cases in detail but imply that other overt containment instances can be found in Turkish.

As seen in Table 2, Turkish case morphology with possessed NPs (e.g. *kedi-si* ‘cat-3POSS’) is almost identical to the Balkar case morphology in Table 1. Considering this, we adopt Davis’s (2023) analysis of Balkar for Turkish and suggest that accusative case is morphologically

contained in more complex (structurally higher) case markers in Turkish after possessive marker *-(s)I*. Considering that accusative can surface as *-nI* in some environments in Turkish (i.e. after *-(s)I*), we suggest that vowel-final roots are not problematic for the containment proposal of Türk & Caha (2021), and genitive case-marked vowel-final roots also show that genitive case marker overtly contains accusative (in *-nI* form). Under the case containment hypothesis then, the case morphology of Turkish can be represented as in Table 3.

	Case	‘house’	‘cat’	‘cat’ (pos- sessed)
a.	NOM	ev	kedi	kedi- <i>si</i>
b.	ACC	ev-i	kedi-yi	kedi- <i>si-ni</i>
c.	GEN	ev-i-n	kedi-ni-n	kedi- <i>si-ni-n</i>
d.	DAT	ev-e	kedi-ye	kedi- <i>si-n-e</i>
e.	LOC	ev-de	kedi-de	kedi- <i>si-n-de</i>
f.	ABL	ev-den	kedi-den	kedi- <i>si-n-den</i>

Table 3. Turkish morphological system predicted by the case containment hypothesis

The standard view of Turkish morphology (Table 2) and the case containment based (CC-based) view (Table 3) make the same predictions regarding the complexity of accusative-marked NPs. An accusative-marked noun like *evi* ‘house.ACC’ consists of two morphemes (*ev-i*), and an accusative-marked possessed NP like *kedisini* ‘cat.3POSS.ACC’ consists of three morphemes (*kedi-si-ni*) under both views.

However, the two views of Turkish morphology differ regarding the morphological complexity of genitive case-marked NPs and other case-marked NPs that carry the suffix *-sI*. According to the standard accounts, a genitive-marked NP like *evin* ‘house.GEN’ consists of two morphemes (*ev-in*). However, according to the CC-based view, the same noun consists of three morphemes (*ev-i-n*). Similarly, a locative-marked possessed NP like *kedisinde* ‘cat.3POSS.LOC’ consists of three morphemes (*kedi-si-nde*) according to the standard view, but according to the CC-based view, it consists of four morphemes (*kedi-si-n-de*), with the insertion of accusative *-n*.

So, according to the CC-based view, accusative-marked NPs (two morphemes) and genitive-marked NPs (three morphemes) differ in morphological complexity: the genitive case-marked NPs are morphologically more complex than their accusative case-marked counterparts. Similarly, according to the CC-based morphological system, locative case-marked possessed NPs are morphologically more complex (four morphemes) than their accusative case-marked counterparts (three morphemes). On the other hand, according to the standard view of Turkish morphology, accusative case-marked NPs and their genitive counterparts are equal in morphological complexity (two morphemes), and accusative case-marked possessed NPs also have the same morphological complexity as their locative counterparts (three morphemes).

1.2. PREDICTIONS FOR WORD PROCESSING. An increase in morphological complexity is known to increase processing effort, as reflected in longer reaction times (RTs) in a lexical decision task (e.g. Gillon, Kehayia & Taler 1999). Given this, a CC-based view of Turkish morphology (Table 3) predicts that genitive-marked NPs in Turkish would elicit longer RTs in a lexical-decision task than accusative-marked NPs, everything else being equal, since the genitive-marked NPs are morphologically more complex. With the same reasoning, it is predicted that locative-marked possessed NPs in Turkish would elicit longer RTs in a lexical-decision task than their accusative case-marked counterparts.

However, if the standard view (Table 2) is correct – i.e., there is no difference in morphological complexity between accusative and genitive/locative case-marked NPs – we do not expect to see any differences in how quickly they are recognized in a lexical decision task. In the experiments reported in this paper, we test these predictions regarding the morphological structure of Turkish, comparing how quickly native speakers recognize NPs with different case markers in a lexical decision task.

Furthermore, there are two approaches to the case containment hypothesis that differ in their predictions regarding genitive-marked NPs. Let us consider these two approaches in more depth. Even though the case containment hypothesis-based morphological system presented for Turkish in Table 3 predicts genitive case to contain accusative case, as proposed by Türk & Caha (2021), the nature of genitive case in case containment is currently unclear. This is because Smith et al.’s structure (2) is based on dependent case theory (DCT) and the status of genitive (and dative) in DCT can vary across languages (Baker 2015). If genitive functions as a lexical (or inherent) case in a given language, genitive must not get into a containment relationship, according to (2). Moreover, if genitive functions as an unmarked case in a language, it must not contain any other cases, but must be contained in all dependent and oblique cases in that language. For Turkish, Satik (2021) proposes that genitive case functions as an unmarked case. If Satik’s analysis is correct, the DCT based case containment hypothesis (2) predicts that Türk & Caha’s (2021) decomposition of genitive into accusative and genitive is not possible, and hence that genitive case-marked NPs are *not* more complex than their accusative counterparts. Only if Caha (2009) is correct is this decomposition possible.

1.3. EARLIER WORK ON MORPHOLOGICAL DECOMPOSITION. The large literature on lexical processing generally suggests that speakers decompose complex words into the morphemes that they contain, a process named morphological decomposition (Taft & Forster 1975; Forster & Davis 1984; Taft 2004; Amenta & Crepaldi 2012; Marantz 2013; Coch, Hua & Landers-Nelson 2020; a.o.). These findings suggest that – other things being equal – upon encountering a morphologically complex word like *nationality*, speakers decompose it into its morphemes (*nation*, *-al*, and *-ity*) and process these morphemes individually, which increases the processing effort associated with the word. Simplex (monomorphemic) words cannot, by definition, be decomposed. They are processed directly, which gives them an advantage regarding processing effort compared to complex words. As a result, complex words lead to significantly longer processing times (slower RTs) compared to simplex words.

Morphological decomposition (and thus RT slowdown) is observed with complex words with both inflectional (Marslen-Wilson & Tyler 1998; Tyler, Marslen-Wilson, & Stamatakis, 2005; Bozic et al. 2007; a.o.) and/or derivational morphology (Marslen-Wilson et al. 1994; Rastle et al. 2004; a.o.). Many previous studies used a lexical decision task, where participants indicate, as fast as possible, whether a visually-presented letter string is a possible word in the language under investigation (e.g. Taft 1979; Vannest & Boland 1999; Coch et al. 2020). Recognition of a word requires processing and accessing the word form, and thus RTs reflect how long it takes the participant to process the target word. Though other factors also impact RTs in lexical decision tasks, it is generally assumed that an increase in morphological complexity leads to an increase in RT (e.g. Gillon, Kehayia & Taler 1999).

One of the other factors that affects word processing is frequency (Forster & Chambers 1973; Gernsbacher 1984; Balota & Chumbley 1984; Balota & Chumbley 1985; a.o.). Frequent words are recognized and processed faster than infrequent words (e.g. Forster & Chambers 1973; Segui, Mehler, Frauenfelder, & Morton 1982; Monsell 1991; a.o.).

With morphologically complex words, the concept of frequency is also more complicated: There are at least three types of frequency that affect lexical decision RTs: root/base frequency, affix frequency, and surface frequency (Vannest et al 2005; Kuperman, Bertram, & Baayen 2010; Coch et al. 2020; a.o.). Root/base frequency is generally described as the total number of occurrences of one root in any morphological form (e.g. *color*, *colorful*, *colorless*, *coloring*), while surface frequency is the number of occurrences of a specific form (e.g. *colorless*). Affix frequency is described as the total number of occurrences of a particular affix in the language regardless of which root it is attached to (e.g. *-less*, in *colorless*, *helpless*, *hopeless*). Previous research has shown that increase in any of these frequency types make word processing easier and leads to faster RTs (Baayen, Dijkstra & Schreuder 1997; Alegre and Gordon 1999; Joanisse & Seidenberg 1999; a.o.). Additionally, some word processing models (e.g. automatic decomposition models, Taft & Forster 1975, dual processing and race models Niemi et al. 1994; Schroeder & Baayen 1995) predict that if a particular complex word form (e.g. *obesity*) has a high enough surface frequency, it can be processed as a whole, rather than being decomposed into its morphemes (here: *obese*, *-ity*) (Bradley 1980; Vannest & Boland 1999; a.o.), in which case these words should not exhibit an RT slowdown relative to simplex words (e.g. Burani & Caramazza 1987; Frauenfelder & Schreuder 1992; Schreuder & Baayen 1995).

The findings regarding how morphological complexity and frequency affect RTs have been replicated by studies conducted with Turkish speakers as well (Gürel 1999; Kırkıcı & Clahsen 2013; Gacan 2014; Bilgin 2016; a.o). In a relatively recent lexical decision task experiment, Bilgin (2016) reports that Turkish native speakers decompose morphologically complex words into morphemes, and the RT to target words are affected by the frequency of the suffix (or suffix templates) of the word.<sup>1</sup> In addition, other studies, using primed lexical decision tasks, found that Turkish native speakers showed morphological priming effects for complex words with inflectional or derivational suffixes (Kırkıcı & Clahsen 2013; Gacan 2014; Jacob & Kırkıcı 2016; Jacob, Şafak, Demir & Kırkıcı 2019; Eldem 2021; a.o) and for compounds (Özer, 2010). So, work in Turkish confirms the findings in the general literature, showing that Turkish speakers decompose complex words into smaller units (morphemes) in the course of word processing.

**2. Methods.** We tested different approaches to the Turkish morphological system (the case containment hypothesis vs. the standard view) by using lexical decision to see how Turkish speakers process accusative, genitive, and locative case-marked nouns. Using two experiment versions, we compared accusative-marked NPs to genitive-marked NPs (Version 1), and accusative-marked possessed NPs to their locative-marked counterparts (Version 2). The experiment was conducted remotely over the internet using PCIBex (Zehr & Schwarz 2018).

**2.1. PARTICIPANTS.** 156 adult Turkish speakers participated and could enter a lottery to win a 100 Turkish Lira gift card. Five participants were excluded from subsequent analyses because they did not reach the pre-determined minimum lexical decision accuracy of 80%. The overall lexical decision accuracy of the remaining 151 participants was ~96.2%. Participants were automatically assigned into one of the two experimental lists.

**2.2. MATERIALS AND DESIGN.** 32 target words were used in each experiment version. Target words consisted of 16 noun pairs that differed in what case marker they carried (e.g. *accusative*

---

<sup>1</sup> Bilgin (2016) investigates different levels of complexity (2-morphemes, 3-morphemes, 4-morphemes, and more) and controls for the frequency of suffix templates (suffixes that frequently appear together). He finds that the frequency of the suffix template has significant effects on the RT on lexical decision. Since the current study does not use suffix templates, this finding of Bilgin is not discussed in detail.

vs *genitive* in Version 1). The targets were controlled for *root frequency* and *surface frequency*, as defined in Section 1. Frequency data was gathered from the TS Corpus Project (Sezer & Sezer, 2013) which is based on the BOUN Corpus (Sak, Güngör & Saraçlar, 2008). Within each target pair (e.g. a noun with *accusative* vs *genitive* marker in Version 1, *accusative* and *locative* in Version 2), the two forms of the noun had similar surface frequencies to the best extent in order to avoid effects of surface frequency on processing times.

Since highly frequent words can be processed without morphological decomposition, we avoided words with very high root or surface frequencies. Also, to minimize effects of word length we only used 4-5-letter roots. Half (16) of the target items were 4-letter nouns; the other half were 5-letter nouns. Thus, we had four conditions: (1) 4-letter-ACC, (2) 4-letter-GEN/LOC, (3) 5-letter-ACC and (4) 5-letter-GEN/LOC. Version 1 tested accusative- and genitive-marked NPs (ACC/GEN) and Version 2 tested accusative- and locative-marked NPs (ACC/LOC).

Target words mostly consisted of concrete nouns such as tools, animals, and flowers. All target words had vowel-final roots (e.g. *fıçtı* ‘barrel’, *vişne* ‘cherry’, *karga* ‘crow’), because consonant-final words with accusative case are highly ambiguous in Turkish between an accusative reading (e.g. *kalem-i* ‘pencil-acc) and a third person possessive reading (e.g. *kalem-i* ‘his/her pencil’). A similar ambiguity exists with the genitive marked nouns, which can be parsed as a regular genitive (e.g. *kuzu-nun* ‘sheep-GEN’) or second person possessive suffix + genitive (e.g. *kuzu-n-un* ‘sheep-2SG.POSS-GEN’). However the possessive reading is very highly marked, and probably not accessible without previous contextual information.<sup>2</sup>

Different from ACC/GEN version, the targets in the ACC/LOC version had the third person possessee marker *-(s)I* in between the root and the case marker (accusative or locative), as the proposed morphological containment is only observed after this suffix (e.g. *sehpa-sı-nda*).

In addition to 32 target words, the study included 96 fillers. Of these, 32 were real Turkish words. Ten were control items that were either bimorphemic (e.g. *çiçek-lik* ‘vase’) or tri-morphemic (e.g. *çiçek-çi-m* ‘my florist’), which shared same roots (e.g. *çiçek* ‘flower’) and had the same total letter length. These items were compared to each other to check if the design of the current study detects RT differences between bimorphemic and tri-morphemic words. In addition, twelve monomorphemic filler words which had similar endings as the case-marked nouns (e.g. *bateri* ‘drums’ and *pelerin* ‘cloak’ for accusative and genitive, respectively) were used. These words were also controlled for word length and frequency as much as possible. If the methodology works, monomorphemic items should elicit shorter RTs than bimorphemic ones, and bimorphemic items should elicit shorter RTs than tri-morphemic ones.

The remaining 64 fillers were non-words used to balance the yes/no answer rate in the lexical decision task. Some of these pseudo-words had real Turkish roots but made-up suffixes or had made-up roots with real Turkish suffixes. The rest were non-words generated by the Wuggy non-word generator (Keuleers & Brysbaert, 2010).

---

<sup>2</sup>Though the unwanted possessive reading is highly marked and unavailable, this in principle could raise concerns regarding our results. To address this concern, participants were given four nouns in different word forms at the end of the study and were asked to use them in sentences. One of these nouns was in genitive-marked form. No participant, out of 156, used the genitive-marked noun with a possessive reading. As we expected, they all interpreted the genitive-marked noun as genitive, suggesting that they most likely interpreted these nouns as genitive in the main experiment as well. Moreover, even if the ambiguity of genitive items could impact the results of the ACC/GEN version, this concern does not apply to the ACC/LOC version of the experiment, as none of the items used in that version are ambiguous.

2.3. PROCEDURE. Participants started the study after confirming that they were at least 18 years old. Participants were instructed that would see letter strings in their screen and that their task is to indicate for each string, as quickly as possible, it is a real word of Turkish or not. The main study was preceded by 6 practice items. On each trial, a fixation cross appeared on the top center of the screen for 750 milliseconds (ms), and it was replaced by the test word after 750 ms. The participants pressed ‘K’ (for *Kelime* ‘word’) if they thought the string is a Turkish word, or ‘D’ (for *Değil* ‘not’) if they thought the string is not a word. A blank screen was shown between trials for 500 ms. The reaction time was calculated as the difference between (i) the time that the stimulus appeared on the screen, and (ii) the time that the participant pressed a key (D/K). Stimulus presentation order was randomized for each participant.

Each stimulus was presented for a maximum of 3 seconds. If a participant did not make a decision after 3 seconds, the stimulus disappeared and the next stimulus appeared on the screen, and the participant response was recorded as ‘missed’. If a participant missed three responses during the experiment, a warning message appeared before the next item was shown, asking the participant to be faster.

2.4. PREDICTIONS. The main predictions of the current study are based on (i) findings of earlier studies on word processing, as well as (ii) predictions of the case containment hypothesis analyses regarding the Turkish morphological system.

As explained in Section 1, earlier studies found that lower root and surface frequencies result in longer RTs in lexical decision tasks. The target items used in the current study were controlled for root and surface frequencies to the best extent possible. So, any effects of surface and root frequencies are expected to be consistent across conditions and should not affect the main comparisons between case-marked forms (ACC vs GEN, and ACC vs LOC).

In addition, suffix frequency is known to impact RTs in lexical decision tasks, especially in agglutinative languages like Turkish (e.g. Bilgin, 2016). Data retrieved from the BOUN Corpus (~491.3 million tokens) reveals that locative (23,710,559 tokens) is the most frequent suffix among the three that are compared in the current study, followed by genitive (19,847,485 tokens) and accusative (13,849,303 tokens) respectively. Based on this, if traditional accounts of Turkish morphology (Table 2) are followed, we predict accusative case-marked items to be processed slower than their genitive counterparts in ACC/GEN version of the study, and their locative counterparts in ACC/LOC version. This prediction is shown in the second column of Table 4.

On the other hand, the case containment hypothesis based analyses of Turkish morphology (Table 3) posits that locative and genitive case-marked NPs are more morphologically complex than accusative NPs, and thus predict these items to be processed slower. However, as discussed in Section 1, there are two main approaches to the Case Containment Hypothesis, possibly making different predictions regarding genitive case. In particular, while the two case containment approaches make the same predictions for Version 2 (ACC/LOC), as shown in Table 4, their predictions diverge for Version 1 (ACC/GEN). According to Caha’s (2009) approach to case containment (CCH, see (1)), genitive contains accusative so genitive-marked NPs are morphologically more complex, and should be elicit longer RTs (Table 4). But according to the Dependent Case Theory approach to case containment (DCT-CCH) in (2), genitive in Turkish does not contain accusative, so under this view genitive-marked NPs are *not* morphologically more complex than accusative-marked NPs. Assuming both are morphologically equally complex, we may find that genitive-marked NPs are processed faster than the accusative NPs in Version 1 (ACC/GEN) because genitive is more frequent than accusative case. These predictions are summarized in Table 4 (CCH = case containment hypothesis, DCT-CCH = dependent case

theory based case containment hypothesis). It is important to note effects of suffix frequency and morphological complexity could in theory cancel each other out, such that accusative- and genitive-marked NPs are processed equally fast, which could potentially support CCH/DCT-CCH approaches.

	traditional	CCH	DCT-CCH
Version 1 (ACC/GEN)	GEN<ACC	ACC<GEN	GEN<ACC
Version 2 (ACC/LOC)	LOC<ACC	ACC<LOC	ACC<LOC

Table 4. Summary of predictionns (< implies shorter/faster RT)

2.5. DATA TRIMMING AND ANALYSIS. Trials with incorrect lexical decisions were excluded from analysis (2.7% of data). RTs below 250 ms were also removed (~0.2% of the remaining data).

Version 1 and 2 used mostly the same control items (monomorphemic, bimorphemic, trimorphemic real words), so we analyzed them together to increase power. When analyzing RTs to target items, we analyzed Version 1 (ACC/GEN) and Version 2 (ACC/LOC) separately, because the targets differed in many ways (e.g. total length, morphological complexity). Thus, we report three main analyses: (i) Version 1 (ACC/GEN), (ii) Version 2 (ACC/LOC), and (iii) the control items of both versions combined. Recall that the analysis of control items tests if the findings of earlier studies regarding basic morphological complexity effects are replicated in the current study, which would verify the validity of our method.

Outlier RTs for individual participants were defined as any RTs more than two standard deviations away from the mean RT for that participant. These outliers were excluded from analysis, which removed ~5.6% of the data in Version 1 (ACC/GEN), ~4.9% in Version 2 (ACC/LOC), and ~5.2% in the combined control items analysis. In subsequent analyses, RTs and frequencies were log-transformed. The best fitting Linear Mixed-Effect Regression (LMER) models (Bates et al. 2015) were built for each analysis using *lme4* in R. Case (accusative vs genitive or accusative vs locative) and root length (4 vs 5 letters) were included in all models as fixed effects. Word length, root frequency, and surface frequency (and relevant interactions) were included if they improved the model fit. Models started with with fully crossed and fully specified fixed and random effects, and were reduced via model comparison; only effects that contributed significantly ( $p < 0.05$ ) were included. (All models included random intercepts for subject and item.)

### 3. Results.

3.1. VERSION 1 - ACCUSATIVE VS GENITIVE. As can be seen in Figure 1, which shows mean RTs for each item type in Version 1, genitive NPs have longer mean RTs than accusative NPs regardless of root length. The best fitting model for the comparison of target items in Version 1 includes effects of item order, case (*accusative* vs *genitive*), and surface frequency as fixed effects on log-transformed RTs. The model reveals that genitive case-marked NPs (mean = 767 ms) were processed slower than accusative-marked NPs (mean = 744 ms) ( $\beta = .02$ ,  $SE = .01$ ,  $t = 2.201$   $p < .05$ ). Also, as expected, the model showed an effect of item order (later items elicit faster RTs than earlier items, a very common finding,  $\beta = -.0015$ ,  $SE = .0001$ ,  $t = -13.035$ ,  $p < .001$ ). We also find effects of surface frequency (words with higher surface frequent elicit faster RTs,  $\beta = -.013$ ,  $SE = .005$ ,  $t = -2.7$ ,  $p < .05$ ). Perhaps surprisingly, root frequency does not improve model fit.

3.2. VERSION 2 - ACCUSATIVE VS LOCATIVE. Figure 2 shows the mean RTs for each item condition in Version 2. In contrast to Version 1, we now see that the mean RTs for accusative and



locative cases depend on root length. Descriptively, with 4-letter roots, accusative NPs have longer mean RTs (775 ms) than locative NPs (762 ms). With 5-letter roots, it is the locative NPs that have longer mean RTs (826 ms) than accusative NPs (788 ms). The best fitting statistical model for the comparison of target items in Version 2 includes effects of item order, case (*accusative vs locative*), root frequency, root length, and interactions between case and root frequency as fixed effects on log-transformed RT.

As expected, later items elicit faster RTs (item order;  $\beta = -.001$ ,  $SE = .0001$ ,  $t = -15.8$ ,  $p < .001$ ). Words with more frequent roots elicit faster RTs than less frequent roots ( $\beta = -.027$ ,  $SE = .012$ ,  $t = -2.224$ ,  $p < .05$ ). Surface frequency does not improve model fit; we leave a closer investigation of surface vs. root frequency effects and their possible relations for future work.

The model shows that although locative-marked NPs (mean= 793 ms) are slower than accusative-marked NPs (mean= 781 ms), this difference is not statistically significant ( $\beta = -.007$ ,  $SE = .0226$ ,  $t = -0.33$ ,  $p = .25$ ). The case x root length interaction is marginally significant ( $\beta = .054$ ,  $SE = .032$ ,  $t = 1.717$ ,  $p = .097$ ), suggesting that RTs elicited by accusative- vs locative-marked NPs depend on whether the root has five or four letters, which is theoretically unexpected. Furthermore, given the observed effect of root frequency, we checked the root frequencies of the 4- and 5-letter items using Welch's two-sample t-test. The test shows that root frequency is higher in 4- than in 5-letter items ( $t(2157.6) = 18.752$ ,  $p < .001$ ). Then, the two item groups were tested for surface frequency, showing that surface frequency was also significantly higher in 4-letter words than 5-letter words ( $t(1894.3) = 17.228$ ,  $p < .001$ ).<sup>3</sup> The high root and surface frequencies of 4-letter items might have elicited very fast RTs and thus *masked* other effects that could come from other factors, such as the accusative vs locative distinction.

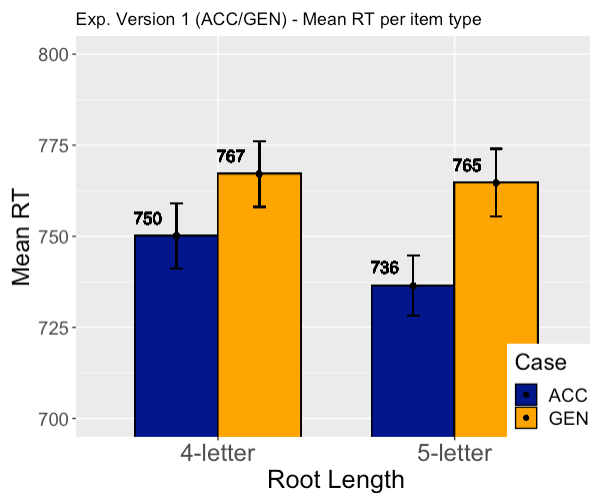


Figure 1. Mean RT per item type in Version 1

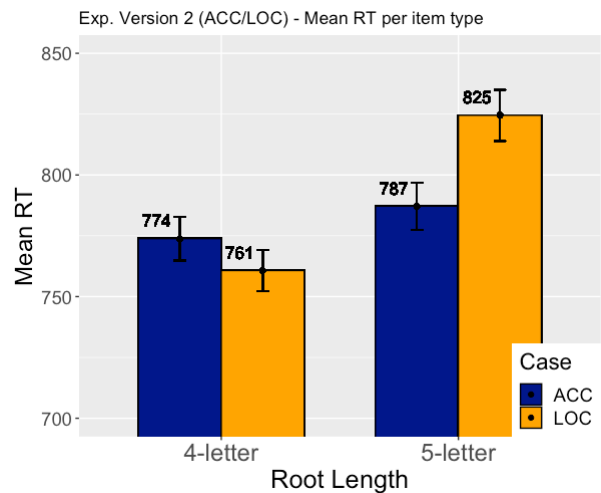


Figure 2. Mean RT per item type in Version 2

Considering this, and the marginal interaction between *case* and *root length* predictors, we conducted separate analyses of the 5-letter and 4-letter items.

The best fitting model for the 5-letter items includes effects of item order and case (*accusative vs locative*) on log-transformed RTs. As expected, the results again show effects of item

<sup>3</sup> The same comparison was done with 4-letter and 5-letter items in ACC/GEN version as well, which revealed no differences of root frequency ( $t(2146) = -0.30$ ,  $p > .05$ ) or surface frequency ( $t(2105.6) = -0.47$ ,  $p > .05$ ) between the 4-letter and 5-letter items. This suggests that a similar frequency problem did not exist in ACC/GEN items.

order ( $p < .001$ ). In addition, the model revealed that locative-marked NPs (mean=826 ms) had slower RTs than accusative-marked NPs (mean= 788 ms) ( $\beta = .048$ ,  $SE = .0203$ ,  $t = 2.367$ ,  $p < .05$ ).

The best fitting model for the analysis of the 4-letter items includes effects of item order, case (*accusative vs locative*), surface frequency and word length. The results showed effects of item order only ( $\beta = -.001$ ,  $SE = .0001$ ,  $t = -10.86$ ,  $p < .001$ ). No other effects were significant. Thus, the RTs to the 4-letter items in Version 2 (ACC/LOC) do not reflect effects that normally arise in a lexical decision task (e.g. surface frequency). This supports the idea that the high frequency of the 4-letter items used in ACC/LOC version of the experiment might have masked any other effects that we could expect to see.

To make sure that a similar issue did not exist in Version 1 (ACC/GEN), we compared the mean root frequencies of items in Version 1 to the 4-letter items in Version 2 (ACC/LOC). As Figure 3 shows, 4-letter items in Version 2 have the highest mean root frequency. Welch's two-sample t-tests reveal that 4-letter items used in Version 2 have higher root frequency than the 4-letter ( $t(1050) = 63.45$ ,  $p < .001$ ) and than the 5-letter items ( $t(1095) = 37.187$ ,  $p < .001$ ) used in Version 1. This supports the idea that a possible reason we fail to find meaningful effects with the 4-letter items in Version 2 (ACC/LOC) is due to their high surface frequency which may mask other effects.

3.2. ANALYSES OF CONTROL ITEMS. The experiment also included control items (real words) with three morphological complexity levels: monomorphemic, bimorphemic, and tri-morphemic. RTs to these items are analyzed in pairs (monomorphemic vs bimorphemic, and bimorphemic vs tri-morphemic), as a sanity check to check whether we can detect higher RTs for morphologically more complex words (e.g. Gillon et al. 1999). If yes, this would indicate that our participants are paying attention and the task is working as expected, i.e., that we are in a position to use the RTs to draw conclusions about morphological complexity in the accusative, genitive and locative conditions.

3.3.1. MONOMORPHEMIC VS BIMORPHEMIC ITEMS. Each version included twelve monomorphemic control words and five bimorphemic control words (Versions 1 and 2 are analyzed together, Section 2.4). The best fitting model for this comparison included the effects of item order, complexity, root length, and root frequency, and interactions between complexity and root length as the fixed factors, with participant and item as random factors. In addition to effects of item order ( $\beta = -.001$ ,  $SE = .00007$ ,  $t = -14.133$ ,  $p < .001$ ), root frequency ( $\beta = -.02$ ,  $SE = .004$ ,  $t = -5.71$ ,  $p < .001$ ), root length ( $\beta = -.028$ ,  $SE = .0098$ ,  $t = -2.892$ ,  $p < .01$ ) and total length ( $\beta = -.02$ ,  $SE = .0058$ ,  $t = -3.404$ ,  $p < .001$ ), the model shows that **monomorphemic words elicit faster RTs than bimorphemic words** ( $\beta = -.028$ ,  $SE = .049$ ,  $t = -5.79$ ,  $p < .001$ ). In addition, we find an interaction of complexity and root length ( $\beta = -.064$ ,  $SE = .01$ ,  $t = 6.442$ ,  $p < .001$ ), such that simplex words show a bigger RT slowdown than complex words as root length increases. This is expected, considering that long monomorphemic word forms are generally infrequent in agglutinative languages, and that frequency affects RTs. Also, this interaction seems to be responsible for the fact that monomorphemic items (mean= 801 ms) on average have slower RTs than bimorphemic items (mean=773 ms). The results show that when the effects of root length (and thus root frequency) are separated from complexity, monomorphemic items are processed faster than bimorphemic items. This replicates the findings of earlier studies and shows that our approach can detect RT differences between monomorphemic and bimorphemic nouns.

3.3.2. BIMORPHEMIC VS TRI-MORPHEMIC ITEMS. Five bimorphemic and five tri-morphemic control items were used in each experiment version. The best fitting model for the comparison between bimorphemic and tri-morphemic items included the effects of item order, complexity (bimorphemic or tri-morphemic), root length, root frequency, and an interaction between root length and root frequency as the fixed factors. The model shows main effects of item order ( $\beta = -.001$ ,  $SE = .0002$ ,  $t = -7.339$ ,  $p < .001$ ), root length ( $\beta = .11$ ,  $SE = .037$ ,  $t = 3.114$ ,  $p < .05$ ), and an interaction between root length and root frequency ( $\beta = -.012$ ,  $SE = .0041$ ,  $t = -2.909$ ,  $p < .05$ ). More importantly for our current purposes, the results show that **bimorphemic items** (mean= 808 ms) **have faster RTs than tri-morphemic items** (mean= 838 ms) ( $\beta = -.044$ ,  $SE = .0166$ ,  $t = -2.672$ ,  $p < .05$ ). Thus, echoing what we saw with mono- vs bimorphemic words, we again find that increased morphological complexity yields longer RTs, in line with the general assumption that RTs increase with morphological complexity (e.g. Gillon et al., 1999).

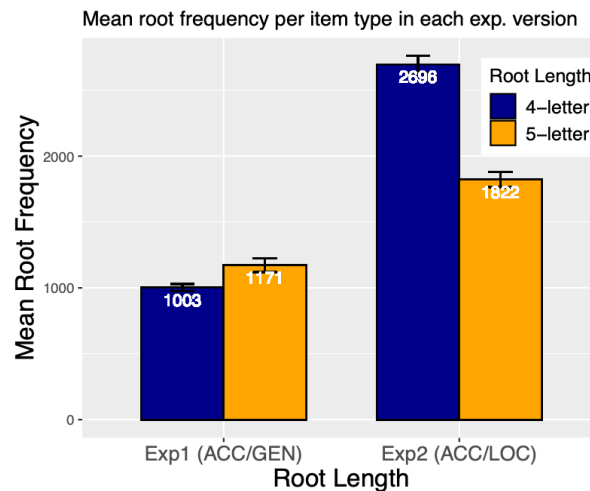


Figure 3. Mean root frequency per root length in each experiment version

In sum, the control item data replicates findings of earlier studies and shows that our method allows us to reliably detect an increased in RTs as morphological complexity increases.

**4. Discussion.** This work tested the predictions of the Case Containment Hypothesis based analyses of Turkish morphology (e.g. Türk & Caha 2021), which posit that genitive and locative case-marked NPs have more complex morphological structures than accusative case-marked NPs. We ran two versions of a lexical decision task: Version 1 compared NPs with accusative vs genitive case; Version 2 compared NPs with accusative vs locative case. Our results show that participants were slower to process *genitive* than *accusative* case-marked NPs, even though the genitive marker is more frequent than the accusative marker in Turkish (Sezer & Sezer 2013). We also find that, when NPs have 5-letter noun roots, participants were slower to process *locative* than *accusative* case-marked NPs, although the locative marker is more frequent than the accusative marker. However, this effect was not found with 4-letter nouns. We suggest that the unexpected lack of a case effect with 4-letter nouns may be due to frequency effects.

Specifically, we suggest that there may be two frequency-related answers to this question. The first, and most plausible, explanation is that the 4-letter nouns in ACC/LOC version were processed very quickly due to their high frequency (compared to other items) which may have masked effects of the case manipulation and the associated morphological complexity. The second (and related) possibility is that 4-letter nouns in ACC/LOC version of the study might be

frequent enough to be processed as a whole, rather than being decomposed into morphemes, as predicted by some word processing models (e.g. Frauenfelder & Schreuder 1992; Schreuder & Baayen 1995). This idea is supported by the fact that no other expected effects (e.g. case, frequency) were significant in 4-letter item condition. The current study was not designed to directly compare these two possibilities, but the asymmetry observed between 4-letter items and 5-letter nouns in the ACC/LOC version (Version 2) emphasizes the importance of stimuli selection in experimental work and should be addressed in future work.

Though future research with more carefully selected set of target items is needed to further explore these issues, the results of the current study support the Case Containment Hypothesis based analyses of Turkish morphological system (Table 3): The predictions made based on this morphological system are borne out in our reaction-time results (except for the 4-letter condition in the ACC/LOC version, which might be due to frequency effects). On the other hand, none of the predictions made based on traditional accounts of Turkish morphology system (Table 2) are supported by our results. If a traditional view of morphology is assumed, both genitive and locative should be processed faster than accusative due to their higher frequency in Turkish.

Moreover, by using two separate experimental versions, and comparing two different case markers (genitive and locative) to accusative, the current study was able to test the predictions of two approaches to the case containment hypothesis. The containment of accusative in genitive is predicted by Caha's (2009) approach to CCH as in (1), but contrasts with the DCT-based approach in (2) (Smith et al. 2019). Ultimately, our experimental results support a morphological analysis based on Caha's approach. To make our results compatible with a DCT-based CCH system, one would need to argue that genitive case functions as an oblique case in Turkish, and as a result contains the accusative (dependent). This contrasts with earlier work (i.e. Satik 2021).

Our results for control items (nouns in nominative case) are in line with the general assumption that an increase in morphological complexity elicits longer RTs in a lexical decision task (e.g. see Gillon, Kehayia & Taler 1999 on morphological complexity, see Taft & Forster 1975; Forster & Davis 1984 or mono- vs. bimorphemic nouns in English). We compared RTs to monomorphemic, bimorphemic and tri-morphemic nouns and found increased morphological complexity yields longer RTs, validating the assumptions made in the literature. Importantly, this also shows that our internet-based method and remote participation yields RT data that detects meaningful differences in morphological complexity.

**5. Conclusion.** The current study has implications for theoretical work on syntax and morphology as well as psycholinguistic work on word processing.

On the theoretical side, our findings provide experimental support for the case containment hypothesis, regarding the abstract structure of morphosyntactic case features. Our results are in line with, and predicted by, the case containment hypothesis account of the morphological system of Turkish (e.g. Türk & Caha, 2021), and contrast with the standard accounts, where there is no containment between any case markers. To the best of our knowledge, the current study is the first one to experimentally test and confirm the predictions made by the Case Containment Hypothesis based analyses of the Turkish morphological system. Moreover, because we created two separate versions of the experiment (ACC/GEN and ACC/LOC), the current study individually tested predictions made by two separate containment structures proposed for the case containment hypothesis (Caha 2009 vs. Smith et al. 2019). The results are in line with Caha's (2009) approach to the case containment hypothesis, where genitive case structurally contains accusative, and do not seem to follow from the dependent case theory based approach to the case containment hypothesis proposed by Smith et al. (2019). However, we acknowledge the need

for future work on the status of genitive case in Turkish, which may shed more light on these issues.

On the experimental side, if the morphological system posited by the case containment hypothesis is at play, our results can have implications for psycholinguistic work on word processing and morphological decomposition. Even if some suffixes are standardly considered as one morpheme (e.g., genitive *-In* in Turkish), if the case containment approach is on the right track, these suffixes might actually underlyingly consist of two (or more) morphemes, which can affect the processes by which words are processed and recognized.

## References

- Amenta, Simona & Davide Crepaldi. 2012. Morphological processing as we know it: An analytical review of morphological effects in visual word identification. *Frontiers in Psychology* 3(232). 1–12. <https://doi.org/10.3389/fpsyg.2012.00232>.
- Baayen, R. Harald, Ton Dijkstra & Robert Schreuder. 1997. Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language* 37(1). 94–117.
- Baker, Mark. 2015. *Case: Its principles and parameters*. Cambridge University Press. <https://doi.org/10.1017/CBO9781107295186>.
- Balota, David A. & James I. Chumbley. 1984. Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, 10(3). 340.
- Balota, David A. & James I. Chumbley. 1985. The locus of word-frequency effects in the pronunciation task: Lexical access and/or production?. *Journal of Memory and Language*, 24(1). 89-106.
- Bates, Douglas, Martin Maechler, Ben Bolker, Steven Walker. 2015. Fitting Linear Mixed-Effects Models Using ‘lme4’. *Journal of Statistical Software*, 67(1). 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bilgin, Orhan. 2016. *Frequency effects in the processing of morphologically complex Turkish words*. Ankara: Middle East Technical University MA thesis. <http://dx.doi.org/10.13140/RG.2.2.20856.44809>.
- Bozic, Mirjana, William D. Marslen-Wilson, Emmanuel A. Stamatakis, Matthew H. Davis & Lorraine K. Tyler. 2007. Differentiating morphology, form, and meaning: Neural correlates of morphological complexity. *Journal of Cognitive Neuroscience* 19 (9). 1464–1475. <https://doi.org/10.1162/jocn.2007.19.9.1464>.
- Bradley, David. 1980. Lexical representation of derivational relation. In Mark Aronoff & M-L. Kean (eds.), *Juncture*, 37–55. Saratoga, CA: Amma Libri & Co.
- Burani, Cristina & Alfonso Caramazza. 1987. Representation and processing of derived words. *Language and Cognitive Processes*, 2(3–4). 217–227.
- Caha, Pavel. 2009. *The nanosyntax of case*. University of Tromsø dissertation.
- Caha, Pavel. 2019. *Case competition in nanosyntax: A study of numerals in Ossetic and Russian*. Berlin: Language Science Press.
- Coch, Donna, Jianjun Hua & Allison Landers-Nelson. 2020. All morphemes are not the same: accuracy and response times in a lexical decision task differentiate types of morphemes. *Journal of Research in Reading* 43. 329–346. <https://doi.org/10.1111/1467-9817.12306>.

- Davis, Collin. 2023. The morphology of case and possession in Balkar: Evidence that oblique cases contain accusative. *Languages* 8(1). 50. <https://doi.org/10.3390/languages8010050>.
- Eldem, Esra. 2021. *Morphological decomposition in heritage Turkish speakers in the US: Evidence from a masked priming experiment*. Long Beach: California State University MA thesis.
- Forster, Kenneth I. & Susan M. Chambers. 1973. Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*. 12(6). 627–635. [https://doi.org/10.1016/S0022-5371\(73\)80042-8](https://doi.org/10.1016/S0022-5371(73)80042-8).
- Forster, Kenneth I. & Chris Davis. 1984. Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 10(4). 680. <https://doi.org/10.1037/0278-7393.10.4.680>.
- Frauenfelder, Uli H. & Robert Schreuder. 1992. Constraining psycholinguistic models of morphological processing and representation: The role of productivity. In *Yearbook of Morphology 1991*, 65–183. Dordrecht: Springer Netherlands.
- Gacan, Pınar. 2014. *The morphological processing of derived words in L1 Turkish and L2 English* Ankara: Middle East Technical University MA thesis.
- Gernsbacher, Morton Ann. 1984. Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General* 113(2). 256–281. <https://doi.org/10.1037/0096-3445.113.2.256>.
- Gillon, Brendan, Eva Kehayia & Vanessa Taler. 1999. The mass/count distinction: Evidence from on-line psycholinguistic performance. *Brain and Language* 68(1–2). 205–211
- Gürel, Ayse. 1999. Decomposition: To what extent? The case of Turkish. *Brain and Language* 68(1), 218–224. <https://doi.org/10.1006/brln.1999.2085>.
- Jacob, Gunnar & Bilal Kırkıcı. 2016. The processing of morphologically complex words in a specific speaker group. *The Mental Lexicon* 11. 308–328
- Jacob, Gunnar, Duygu Şafak, Orhan Demir & Bilal Kırkıcı. 2019. Preserved morphological processing in heritage speakers: A masked priming study on Turkish. *Second Language Research* 35(2). 173–194. <https://doi.org/10.1177/0267658318764535>
- Joanisse, Marc. F. & Mark S. Seidenberg. 1999. Impairments in verb morphology after brain injury: A connectionist model. *Proceedings of the National Academy of Sciences* 96(13). 7592–7597. <https://doi.org/10.1073/pnas.96.13.7592>.
- Keuleers, Emmanuel & Marc Brysbaert. 2010. Wuggy: A multilingual pseudoword generator. *Behavior Research Methods* 42. 627–633 <https://doi.org/10.3758/BRM.42.3.627>
- Kırkıcı, Bilal & Harald Clahsen. 2013. Inflection and derivation in native and non-native language processing: Masked priming experiments on Turkish. *Bilingualism: Language and Cognition* 16(04). 776–791.
- Kuperman, Victor, Raymond Bertram & R. Harald Baayen. 2010. Processing trade-offs in the reading of Dutch derived words. *Journal of Memory and Language*, 62. 83–97. <https://doi.org/10.1016/j.jml.2009.10.001>
- Marantz, Alec. 2013. No escape from morphemes in morphological processing. *Language and Cognitive Processes* 23. 905–916. <https://doi.org/10.1080/01690965.2013.779385>.
- Marslen-Wilson, William D. & Lorraine K. Tyler. 1998. Rules, representations, and the English past tense. *Trends in Cognitive Science* 2. 428–436.
- Marslen-Wilson, William D., Lorraine K. Tyler, Rachelle Waksler & Lianne Older. 1994. Morphology and meaning in the English mental lexicon. *Psychological Review* 101(1). 3–33.

- Monsell, Stephen. 1991. The nature and locus of word frequency effects in reading. In Derek Besner & Glyn W. Humphreys (eds.), *Basic processes in reading: Visual word recognition*. 148–197. London: Routledge.
- Niemi, Jussi, Matti Laine & Juhani Tuominen. 1994. Cognitive morphology in Finnish: Foundations of a new model. *Language and Cognitive Processes*. 9(3). 423–446.
- Özer, Sibel. 2010. *Morphological priming in Turkish nominal compound processing*/ Ankara: Middle East Technical University MA thesis.
- Rastle, Kathleen, Matthew H. Davis & Boris New. 2004. The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin & Review* 11. 1090–1098.
- Sak, Haşim, Tunga Güngör & Murat Saraçlar. 2008. Turkish language resources: Morphological parser, morphological disambiguator and web corpus. *Advances in Natural Language Processing* 6. 417–427.
- Satik, Deniz. 2020. Turkic genitive case and agreement asymmetries. Manuscript. <https://ling.auf.net/lingbuzz/004981>
- Schreuder, Robert & R. Harald Baayen. 1997. How complex simple words can be. *Journal of Memory and Language* 37. 118–139.
- Segui, Juan, Jacques Mehler, Uli Frauenfelder & John Morton. 1982. The word frequency effect and lexical access. *Neuropsychologia* 20(6). 615–627.
- Sezer, Taner. 2017. TS Corpus Project: An online Turkish dictionary and TS DIY corpus. *European Journal of Language and Literature* 9(1). 18–24.
- Smith, Peter W., Beata Moskal, Ting Xu, Jungmin Kang & Jonathan D. Bobaljik. 2019. Case and number suppletion in pronouns. *Natural Language and Linguistic Theory* 37. 1029–1101. <https://doi.org/10.1007/s11049-018-9425-0>
- Taft, Marcus. 1979. Recognition of affixed words and the word frequency effect. *Memory and Cognition* 7(4). 263–272. <https://doi.org/10.3758/BF03197599>.
- Taft, Marcus. 2004. Morphological decomposition and the reverse base frequency effect. *Quarterly Journal of Experimental Psychology* 57A. 745–765.
- Taft, Marcus & Kenneth I. Forster. 1975. Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, 14(6). 638–647.
- Türk, Utku & Pavel Caha. 2021. Nanosyntactic analysis of Turkish case system. *Proceedings of the Workshop on Turkic and Languages in Contact with Turkic* 6. 5051. <https://doi.org/10.3765/ptu.v6i1.5051>.
- Tyler, Lorraine K., William D. Marslen-Wilson & Emmanuel A. Stamatakis. 2005. Differentiating lexical form, meaning and structure in the neural language system. *Proceedings of the National Academy of Sciences* 102. 8375–8380.
- Vannest, Jennifer & Julie E. Boland. 1999. Lexical morphology and lexical access. *Brain and Language* 68. 324–332.
- Vannest, Jennifer, Elissa L. Newport, Aaron J. Newman & Daphne Bavelier. 2011. Interplay between morphology and frequency in lexical access: The case of the base frequency effect. *Brain Research* 1373. 144–159.
- Veselinova, Ljuba N. 2006. *Suppletion in verb paradigms*. Amsterdam: John Benjamins.
- Zehr, Jeremy & Florian Schwarz. 2018. PennController for Internet Based Experiments (IBEX). <https://doi.org/10.17605/OSF.IO/MD832>
- Zompi, Stanislao. 2017. *Case decomposition meets dependent-case theories*. University of Pisa MA thesis.