

**Research Article**  

**Cite this article:** Uygun S, Clahsen H (2021). Morphological processing in heritage speakers: A masked priming study on the Turkish aorist. *Bilingualism: Language and Cognition* **24**, 415–426. <https://doi.org/10.1017/S1366728920000577>

Received: 3 March 2020  
Revised: 25 August 2020  
Accepted: 26 August 2020  
First published online: 10 November 2020

**Keywords:**

Turkish; morphology; aorist; priming; variability; processing

**Address for correspondence:**

Serkan Uygun, Email: [uygun@uni-potsdam.de](mailto:uygun@uni-potsdam.de)

**Abstract**

Previous research has shown that heritage speakers struggle with inflectional morphology. ‘Limitations of online resources’ for processing a non-dominant language has been claimed as one possible reason for these difficulties. To date, however, there is very little experimental evidence on real-time language processing in heritage speakers. Here we report results from a masked priming experiment with 97 bilingual (Turkish/German) heritage speakers and a control group of 40 non-heritage speakers of Turkish examining regular and irregular forms of the Turkish aorist. We found that, for the regular aorist, heritage speakers use the same morphological decomposition mechanism (‘affix stripping’) as control speakers, whereas for processing irregularly inflected forms they exhibited more variability (i.e., less homogeneous performance) than the control group. Heritage speakers also demonstrated semantic priming effects. At a more general level, these results indicate that heritage speakers draw on multiple sources of information for recognizing morphologically complex words.

**Introduction**

Heritage speakers are taken to be a subset of (unbalanced) bilinguals who acquired the heritage language as a first language in their homes but live in an environment in which another dominant community language is spoken (Scontras, Polinsky & Fuchs, 2018). They may be simultaneous bilinguals in which case heritage speakers acquire both the non-dominant and the dominant language concurrently, or sequential bilinguals in which case the dominant language comes in later in life. Typically, the dominant community language becomes a heritage speaker’s most familiar language when s/he starts school at around ages 5 or 6 (Montrul, 2012). Heritage speakers acquire the non-dominant language from reduced input and/or practice it less than non-heritage speakers of the same language. Hence, the study of heritage languages and their speakers should provide insights into the role of exposure and language use for language acquisition and language processing and help to answer the question of which aspects of linguistic knowledge and which mechanisms of language processing remain stable despite reduced exposure and practice, and which ones are less robust and more susceptible to adverse exposure/input conditions. Another hallmark of heritage speakers is the notable variability in their language abilities (e.g., Köpke & Schmid, 2004), which may pose a potential challenge for developing any general model of heritage languages and their speakers. Yet, the variability among heritage speakers may not be chaotic or random (Putnam, Kupisch & Pascual y Cabo, 2018) and may provide insights that are less directly available from non-heritage baseline languages and their speakers (Flores & Rinke, 2020).

Previous studies have reported that inflectional morphology is one of the linguistic domains that are particularly challenging for heritage speakers (e.g., Arabic: Benmamoun, Montrul & Polinsky, 2010; Korean: Choi, 2003; Russian: Gor & Cook, 2010; Gor & Vdovina, 2010; Spanish: Montrul, 2011). More specifically, heritage speakers were found to be less accurate on irregular than regular inflection and to commonly overapply regular morphological rules (Montrul & Mason, 2020). In a recent keynote article, Polinsky and Scontras (2020a) outlined a model of heritage language in which they attribute this so-called ‘resistance to irregularity’ (along with other properties of heritage language grammars) to two factors, (i) insufficient input, (ii) limited online resources for processing and memory. They specifically argued that ‘the limited nature of processing resources, combined with the added cost of operating in a non-dominant language, forces heritage speakers to draw on knowledge from other domains as they optimize their resource use’ (Polinsky & Scontras, 2020a, p. 14). As regards inflectional morphology, what they referred to as ‘other domains’ is in this case the overapplication of regular morphology which is supposed to reduce memory demands and to overcome the problem of insufficient input in heritage speakers’ language and language performance; see Polinsky and Scontras (2020a, p. 15). Unfortunately, however, there is to date very little evidence for these

© The Author(s), 2020. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.



claims. To argue for ‘processing limitations’ as the source of properties of heritage language, online experimental studies that tap into the mechanisms of language processing are required. Such experiments are, however, extremely scarce for heritage speakers as pointed out in Felser’s (2020) and Gürel’s (2020) commentaries of the Polinsky and Scontras (2020a) keynote article.

There are (to our knowledge) only four experimental studies that explore morphological processing in heritage speakers, two on Russian (Gor & Cook, 2010; Gor, Chrabaszc & Cook, 2019) and two on Turkish (Jacob & Kırkıcı, 2016; Jacob, Şafak, Demir & Kırkıcı, 2019). The results from these studies are mixed and do not yield a clear picture. Gor and Cook (2010) found (auditory) priming effects for regular and irregular Russian verb forms in a group of heritage speakers of Russian. Gor et al. (2019) reported results from a cross-modal priming experiment that yielded differences between default and oblique case markings in heritage speakers of Russian, significant priming effects for the (default) nominative case of Russian, but no effect for two of the oblique cases (genitive, instrumental). For Turkish, both Jacob and Kırkıcı (2016) and Jacob et al. (2019) found reliable masked priming effects for regular verb forms in heritage speakers and non-heritage L1 controls. However, while Jacob and Kırkıcı (2016) additionally obtained significant orthographic priming for the heritage speaker group only, there was no such effect in Jacob et al.’s (2019) study, in either participant group. It is hard to draw any strong conclusions from the results of these studies. Clearly more research is needed in this domain to properly assess claims about ‘limited processing resources’ in heritage speakers. It is true that Jacob et al. (2019) reported slower RTs in heritage speakers than in non-heritage control speakers, in line with limited processing resources in heritage speakers. However, according to Polinsky and Scontras (2020a), we would mainly expect irregular morphology to strain a heritage speaker’s processing system, and Jacob et al. (2019) as well as Gor et al. (2019) only dealt with regular morphology. Furthermore, Gor and Cook’s (2010) auditory priming study tested irregular morphology, but focused on comparing heritage speakers with L2 learners rather with non-heritage controls.

The current study aims at determining the mechanisms heritage speakers employ for processing both regularly and irregularly inflected word forms and potential differences to non-heritage speakers of the same language. We tested the processing of inflected word forms in Turkish, an agglutinating language with largely regular morphology. While previous experimental studies of Turkish morphology have examined regularly inflected forms only, our study is the first to also investigate the processing of irregular forms in Turkish. We employed the visual masked priming technique, which has been argued to be particularly sensitive to morphological cues for visual word recognition and lexical access. In addition to accuracy and speed, we also included a measure of variability to capture the potentially larger variability among heritage speakers. As will be shown, these two measures – taken together – will provide new insights into how the processing of morphologically complex words differs between heritage and non-heritage speakers.

### Mechanisms of morphological processing during word recognition

One of the most crucial mechanisms involved in the processing of morphologically complex words is MORPHOLOGICAL DECOMPOSITION, by which a morphologically structured word form is segmented

into its component parts, e.g., *walked* into [walk]+ed. Previous research from several studies using masked-priming experiments has provided evidence for morphological decomposition during visual word recognition and lexical access. In a masked priming experiment, prime words are presented for a very short period of time only, which typically prevents the prime words from being directly recognized. Instead, masked priming is supposed to tap into subliminal processes involved in visual word recognition. A considerable number of studies have reported that lexical decision times on targets following masked morphologically related primes tend to be faster than those following unrelated primes; see Marslen-Wilson (2007) for review. Note, however, that morphologically related prime-target pairs such as *walked* → *walk* also exhibit (orthographic/phonological) form and semantic overlap. Indeed, a number of experimental studies have found effects of orthographic form and/or semantic overlap on response times under masked-priming conditions (e.g., Feldman, O’Connor & Moscoso del Prado Martín, 2009; Nakano, Ikemoto, Jacob & Clahsen, 2016).

Different mechanisms have been proposed to account for morphological processing and specifically for masked-priming effects during word recognition and lexical access. One proposal posits a universal morpho-orthographic segmentation mechanism (‘affix stripping’) that automatically and subconsciously strips off affixes from their stems prior to lexical access (Taft & Forster, 1975; Taft, 2003). Under this view, facilitated target word recognition after morphologically related prime words is in essence a stem-repetition effect, resulting from morpho-orthographic decomposition. A radical alternative holds that the recognition of inflected or derived words does not make use of their morphological structure, but is instead determined by their form-level and semantic properties (Baayen, Milin, Durdevic, Hendrix & Marelli, 2011; Seidenberg & Gonnerman, 2000). Masked-priming effects under this view are due to orthographic and/or semantic overlap, rather than to morphological relatedness (Feldman et al., 2009; Feldman, Milin, Cho, Moscoso del Prado Martín & O’Connor, 2015). A third approach posits a dual-mechanism account according to which both sublexical morpho-orthographic decomposition and morpho-lexical activation of related lexical entries are involved in morphological processing; see Crepaldi, Rastle, Coltheart and Nickels (2010), Bosch, Verissimo and Clahsen (2019), Verissimo (2018). Morpho-lexical activation is thought to operate on the associatively-linked lexical representations of irregular forms, i.e., on lexically restricted exceptions that are likely to be learned as individual items and stored in lexical memory. These forms may be entirely suppletive (e.g., *are*), may contain unpredictable stem changes (e.g., *fell*) or potential affixes (e.g., *kept*). The crucial common property is lexical restrictedness. Masked-priming effects from regularly affixed word forms are explained in terms of the direct morpho-orthographic decomposition route, whereas the often smaller masked priming effects from irregular forms such as for *fell* → *fall* or *kept* → *keep* are attributed to the second (weaker) morpho-lexical priming route.

As regards Turkish, the language we examined here, there are only very few morphological processing studies. Hankamer (1989) speculated that due to its regular agglutinative morphology, the Turkish language processor should largely rely on morphological decomposition of complex words. Frauenfelder and Schreuder (1992), on the other hand, hypothesized that morphologically complex word forms, at least those with high surface frequency, may be stored as wholes rather than being decomposed. Evidence for this comes from the results of Gürel’s (1999)

unprimed lexical-decision study which compared multimorphemic nouns with ablative, locative or plural suffixes with frequency-matched monomorphemic nouns in monolingual Turkish speakers. The results revealed that nouns with the ablative suffix (which has the lowest frequency among the suffixes under investigation) yielded significantly longer RTs than the monomorphemic control items, whereas this was not the case for nouns with plural and locative suffixes. Gürel (1999) argued that a morphologically complex word in Turkish, depending on the frequency of a suffix, can also be recognized via the whole-word access route; see also Gürel and Uygun (2013) and Uygun and Gürel (2016) for related findings on Turkish and Lehtonen and Laine (2003) for Finnish, another agglutinating language.

Turkish morphology was also examined in four masked-priming studies. Kırkıcı and Clahsen (2013) examined regularly inflected verb forms and derived (deadjectival) nominals in (non-heritage) L1 speakers and in adult second language (L2) speakers of Turkish. They reported significant priming effects from both inflected and derived word forms in the L1 group, derivational (but no inflectional) priming in the L2 group, and no effects of orthographic overlap in either the L1 or the L2 group. Kırkıcı and Clahsen (2013) explained these priming patterns in terms of morphological decomposition, which they argued functions more efficiently for (non-heritage) L1 than L2 speakers. Jacob and Kırkıcı (2016) performed the same masked-priming experiment with a group of heritage speakers of Turkish. They found significant priming effects not only for inflected and derived prime words in the heritage group (parallel to L1 controls), but also for orthographically related prime-target pairs (unlike in the L1 controls). Jacob and Kırkıcı (2016, p. 308) concluded that heritage speakers "...rely more on (orthographic) surface form ...at the expense of morphological decomposition". In a follow-up study, Jacob et al. (2019) again tested for priming effects from regularly inflected verb forms and derived nominals in new groups of heritage speakers of Turkish and L1 controls. The results were parallel for both participant groups, significant priming in the two morphological conditions and no reliable priming in the semantic or the orthographic control conditions. Jacob et al. (2019) concluded that heritage speakers employ the same morphological decomposition mechanisms for processing inflected and derived word forms as non-heritage L1 controls. Note, however, that additional considerations are required to explain why this is apparently not the case for Jacob and Kırkıcı's (2016) group of heritage speakers, even though they come from the same cohort of Turkish/German bilinguals living in the Berlin/Potsdam area. We come back to this particular discrepancy in the *Discussion* section. The fourth study comes from Uygun and Gürel (2018) investigating the processing of one-suffix nouns with a plural, locative or ablative suffix in native, L1 English and L1 Russian speakers of Turkish. The results indicated priming effects for all participant groups, which the authors took as an indication of morphological decomposition during processing for both native and non-native speakers.

### The present study

The experiment reported below investigates the processing of inflected word forms of Turkish in bilingual heritage speakers (henceforth 'HS') of Turkish and a control group of non-heritage Turkish speakers (henceforth 'CTR'). Here, we adopt a broad notion of the term 'heritage speaker' that includes individuals who underwent incomplete acquisition and/or attrition. In the heritage language literature, there is a controversy as to whether

gaps in the heritage language grammar (relative to non-heritage speakers' grammars) are due to incomplete acquisition of the language in childhood, attrition of structures acquired early in childhood but lost due to lack of exposure and use later on, or both; see Montrul (2008) for a review. Our motivation for including both incomplete acquirers and attriters with varying ages of arrival into our study was to increase the heterogeneity of the sample in order for us to examine which aspects of morphological processing are variable and which are stable (within an otherwise heterogeneous group of participants).

The linguistic phenomenon under study is the aorist in Turkish which encodes habitual aspect or general present tense. Unlike previous experimental studies on the Turkish aorist that only investigated regular forms, the present study examined both regular and irregular forms. There are different exponents of the aorist; see Nakipoğlu and Ketrez (2006). Multisyllabic verb stems ending with consonants take variants of *-Ir* (*-ır, -ir, -ur, -ür*). The choice of the latter is determined by the rules of Turkish vowel harmony, e.g., *konuş-ur* "speaks", *düşün-ür* "thinks". Mono- and multisyllabic verb stems ending in a vowel take the suffix *-r*, e.g., *ye-r* "eats", *uyu-r* "sleeps". Monosyllabic verb stems ending with a consonant are most commonly suffixed with *-Ar* (*-ar, -er*). The choice between the latter two is again determined by vowel harmony, e.g., *kes-er* "cuts", *sor-ar* "asks". In addition, there are 13 monosyllabic verb stems, all of which are highly frequent, that are exceptional in their aorist form in that they require the suffix *-Ir* instead of the regular *-Ar*, e.g., *gör-ür* "sees", *var-ır* "arrives". Thus the irregularity in this particular case arises from an existing suffix exceptionally applied to a limited unpredictable set of monosyllabic stems.

With materials constructed from the aorist, it is possible to directly compare morphological processing of regular and irregular inflection, the first study of this kind for Turkish. We employed the visual masked priming technique, because this technique is (arguably) particularly sensitive to a word's internal morphological structure. Two morphological priming conditions, one with regular ('Morph/Reg') and the other with irregular ('Morph/Irreg') aorist forms of monomorphemic verbs as primes were designed. In addition, we included an orthographic control priming condition, with the same word-initial orthographic overlap as the morphological conditions, and a semantic control condition with prime-target pairs that were semantically closely related (similarly to those of the morphological conditions), but otherwise unrelated; see (1) for an example stimulus set of the different priming conditions.

(1)	Related Prime	Unrelated Prime	Target
'Morph/Reg'	<i>duyar</i> "hears"	<i>bekle</i> "wait"	<i>DUY</i> "hear"
'Morph/Irreg'	<i>gelir</i> "comes"	<i>zaman</i> "time"	<i>GEL</i> "come"
'Orthographic'	<i>devre</i> "period"	<i>üslup</i> "style"	<i>DEV</i> "giant"
'Semantic'	<i>kafa</i> "head"	<i>merkez</i> "center"	<i>BAŞ</i> "head"

Given previous results from several masked-priming studies with non-heritage L1 speakers, we expect that our non-heritage control group participants will show significant morphological



priming effects that are clearly dissociable from any facilitation (or inhibition) in the orthographic and semantic control conditions. As far as HS are concerned, we can derive two specific predictions from Polinsky and Scontras' (2020a) account of 'limitations of online resources' in HS. Firstly, recall that Polinsky and Scontras (2020a) refer to memory demands as one of the sources of HS' difficulties with irregular morphology and their overreliance on regular morphology in language production ('resistance to irregularity'). These memory demands are also likely to affect perception and recognition, resulting in less efficient processing of irregular forms (relative to regular ones). Consequently, we expect to find smaller priming effects for irregular than for regular aorist forms. Secondly, Polinsky and Scontras (2020a) hypothesized that due to limitations of the heritage grammar and of the corresponding grammatical processing resources, as well as the added cost of operating in a non-dominant language, HS draw on knowledge from other domains. Consequently, we expect to find effects of orthographic and/or semantic relatedness in HS in our experiment.

## Methods

### Participants

We tested two participant groups, (i) a HS group consisting of 110 Turkish–German bilinguals who had acquired Turkish from birth and were recruited from a large Turkish community residing in Berlin and Potsdam (Germany), and (ii) a CTR group consisting of 40 non-heritage L1 speakers of Turkish recruited and tested in Istanbul (Turkey). All participants completed the grammar section of the TELC test (<https://www.telc.net/tr.html>), which gave us a measure of their knowledge of Turkish grammar. They all received a small fee for taking part in our study. Prior to any data analysis, ten participants from the HS group who had less than 12 points (maximum score: 20) in the TELC test were excluded. During data cleaning, three more HS participants were excluded because of their low accuracy in the lexical decision task (< 70%). The data from the remaining 97 HS participants (53 women, mean age: 32.91, SD: 10.67) were further analyzed and compared to the 40 participants of the CTR group (36 women, mean age: 36.13, SD: 9.76). These 97 individuals of the HS group had a high TELC/Turkish score (mean: 18.55, SD: 2.02), albeit significantly lower ( $t = -6.182$ ) than the CTR group's score (mean: 19.88, SD: 0.40). Self-ratings of the HS' skill (on a 10-point scale) also revealed a high level of Turkish, for both speaking/listening (mean: 8.91, SD: 1.33) and for reading/writing (mean: 8.32, SD: 2.03). They also rated Turkish as being common in their daily language use (spoken Turkish: mean: 3.22, SD: 1.31; written Turkish: mean: 2.99, SD: 1.38; maximum score: 5). Concerning the HS' level of education, 31 individuals hold degrees from universities, 30 from high schools, 24 from vocational schools, and 12 participants from secondary schools. Amongst the CTR participants, ten individuals are high school graduates and the rest have university degrees. The HS' age of arrival in Germany varied considerably (mean age of arrival: 10.11, SD: 9.64); 38 HS participants were born in Germany.

### Materials

There were 24 pairs of morphologically related primes and targets, 12 with irregular and 12 with regular aorist forms as primes, and 24 unrelated prime-target pairs with no morphological, orthographic,

or semantic relation between the prime and target word. Irregular aorist forms contain the suffix *-Ir* in one of its variants (*-ir, ir, -ur, -ür*) – the latter determined by vowel harmony. Regular aorist forms contain the suffix *-Ar* in one of its variants (*-ar, -er*), again determined by vowel harmony. The corresponding target word forms for both irregular and regular aorist prime words were monosyllabic bare verb stems, which also function as second person singular imperative forms in Turkish. The morphologically related prime words were third person singular aorist forms, which do not contain any overt person or number suffix. Furthermore, in both aorist conditions the target words are fully contained in the related prime words, so that stem-priming effects for aorist forms can be directly determined without any potential interference from unprimed affixes or stem allomorphs appearing on the target word. Unrelated prime words were either nouns or verbs that did not bear any morphological, orthographic or semantic relation to the target. Target words in both prime conditions (related, unrelated) were presented in upper case and prime words in lower case, to minimize visual overlap between primes and targets; see example stimuli in (1) above.

Two control conditions were added to determine whether any priming effect in the morphological conditions could be due to orthographic and/or semantic overlap between primes and targets. In the orthographic control condition, there were 12 related and 12 unrelated prime-target pairs (taken from Kırkıcı & Clahsen, 2013); see 'orthographic' in (1) above for an example. The related prime-target pairs were morphologically and semantically unrelated but overlapped orthographically in the same way as the morphologically related prime-target pairs. Specifically, the target word was fully contained in the prime word in both the morphological and the orthographic conditions. We also calculated orthographic overlap ratios between prime and target using the Spatial Coding measure of Davis' (2000) Match Calculator, which revealed parallel mean orthographic overlap ratios for these conditions ('Irregular': 0.62; 'Regular': 0.58; 'Orthographic': 0.56). Note also that the bigram *-re* of the prime word *devre* 'period' (or any other grapheme sequence of a prime word that does not occur in the corresponding target) does not represent an existing suffix in Turkish. All target words in the orthographic condition were monosyllabic unaffixed nouns and their related primes were disyllabic unaffixed nouns. The unrelated primes consisted of nouns and adjectives that were not orthographically related to the target words.

The semantic control condition consisted of 20 related (10 antonyms and 10 synonyms) and 20 unrelated prime-target pairs; see example stimulus set 'semantic' shown in (1) above. The target words in the semantic condition were nouns, adjectives, and adverbs. Each target item was preceded by either a semantically related but morphologically and orthographically unrelated prime word (again a noun, adjective, or adverb), or an unrelated prime word. In order to ensure semantic relatedness, an online survey with 27 Turkish native speakers (who did not participate in the main experiment) was conducted in which they were asked to rate, on a five-point Likert scale (with 1 as the lowest degree of similarity in meaning), how related the prime-target pairs were in meaning. They rated both the related (20 items) and unrelated (20 items) prime-target pairs together with 45 filler pairs that were either semantically related or unrelated. The results for the target items confirmed that the semantically related primes received very high ratings (mean: 4.63, SD: 0.80) and the unrelated primes low semantic relatedness ratings (mean: 1.22, SD: 0.61).

**Table 1.** Item properties (means, standard deviations and 95% confidence intervals) of prime and target words**Table 1a.** Item properties of prime words

Condition	Prime Type	LF	WFF	Letter
Irregular	Related	<b>3.02</b> (0.42) [2.75, 3.29]	<b>1.17</b> (0.55) [0.82, 1.52]	<b>4.83</b> (0.39) [4.59, 5.08]
	Unrelated	<b>2.97</b> (0.37) [2.73, 3.21]	<b>1.02</b> (0.67) [0.59, 1.44]	<b>4.83</b> (0.58) [4.47, 5.20]
Regular	Related	<b>3.05</b> (0.31) [2.85, 3.24]	<b>1.20</b> (0.37) [0.96, 1.43]	<b>4.75</b> (0.45) [4.46, 5.04]
	Unrelated	<b>3.09</b> (0.21) [2.96, 3.23]	<b>1.13</b> (0.45) [0.84, 1.42]	<b>4.83</b> (0.58) [4.47, 5.20]
Orthographic	Related	<b>1.84</b> (0.29) [1.67, 2.01]	<b>1.28</b> (0.52) [0.95, 1.61]	<b>5.08</b> (0.29) [4.90, 5.27]
	Unrelated	<b>1.46</b> (0.63) [1.05, 1.86]	<b>1.20</b> (0.62) [0.80, 1.59]	<b>4.67</b> (0.65) [4.25, 5.08]
Semantic	Related	<b>2.06</b> (0.62) [1.77, 2.35]	<b>1.71</b> (0.64) [1.41, 2.01]	<b>5.10</b> (0.55) [4.84, 5.36]
	Unrelated	<b>1.60</b> (0.73) [1.26, 1.94]	<b>1.24</b> (0.69) [0.92, 1.57]	<b>5.40</b> (0.50) [5.17, 5.63]

**Table 1b.** Item properties of target words

Condition	LF	WFF	Letter
Irregular	<b>3.02</b> (0.42) [2.75, 3.29]	<b>1.15</b> (0.46) [0.86, 1.44]	<b>2.83</b> (0.39) [2.59, 3.08]
Regular	<b>3.05</b> (0.31) [2.85, 3.24]	<b>1.08</b> (0.53) [0.74, 1.41]	<b>2.75</b> (0.45) [2.46, 3.04]
Orthographic	<b>1.33</b> (0.44) [1.05, 1.61]	<b>1.01</b> (0.48) [0.70, 1.31]	<b>3.00</b> (0.00) [3.00, 3.00]
Semantic	<b>2.05</b> (0.78) [1.69, 2.42]	<b>1.73</b> (0.72) [1.39, 2.06]	<b>2.90</b> (0.48) [2.69, 3.11]

LF = lemma frequency; WFF = word-form frequency; Overlap = orthographic overlap between prime and target.

We also controlled the items for length and frequency. **Table 1** presents information on the experimental items, separately for the prime and the target words: (i) length (in letters); (ii) mean lemma and word-form frequencies per million of the experimental items, based on the Turkish ‘TS Corpus’ (<https://tscorpus.com/>, Sezer, 2017), which consists of nearly 5 million types and 500 million tokens taken from sources such as newspapers, social media, forums, blogs, and academic journals and books. The materials and data are publicly available at: <https://doi.org/10.17605/OSF.IO/3SJHA>.

By following a Latin Square design, the prime-target pairs were distributed over two presentation lists to ensure that no participant saw the same target more than once. A set of 114 fillers was added to the 56 experimental prime-target pairs (12 ‘irregular’, 12 ‘regular’, 12 ‘orthographic’, 20 ‘semantic’), resulting in a total of 170 trials per presentation list. The prime-target pairs in each list were pseudorandomized to eliminate any undesired priming effects across items. Out of the 114 filler pairs, 85 were nonwords so that half of the items in each version required a ‘no’ response. Nonwords were created by changing 2–3 letters of an existing word without violating the phonotactic rules of Turkish. The remaining 29 words that served as fillers were monomorphemic verbs. The proportion of targets preceded by a related prime constituted 20% of all trials in each list.

### Procedure

All participants were tested in a quiet lab room and were randomly assigned to one presentation list. The CTR group was tested in Istanbul, and the HS group in Berlin/Potsdam. Their response accuracy and reaction times (RTs) were measured by using the DMDX software (Forster & Forster, 2003). Prior to the experiment, all participants completed a demographic background questionnaire and consent form. After that, participants received detailed instructions about the procedure of the main experiment, that they had to respond to a set of words appearing on the computer screen by pressing, as quickly and accurately as possible, either a ‘Yes’ or ‘No’ button on a gamepad connected to the computer. Yes-responses were always elicited with the participants’ dominant hand. At the beginning of the experiment, participants were shown a number of practice items so that they become familiar with the experimental procedure. After the practice session, the participants were asked whether they could identify any words they had seen during the practice session. None of the participants reported that they had seen any of the primes. Following that, the main experiment started. For each trial, participants were first presented with a blank screen for 500 ms. This was followed by a standard forward mask (#####) consisting of the number of hashes equal to the number of letters of the prime presented in the center of the screen for 500 ms. This was followed by the prime – presented for 50 ms – immediately followed by the target. The target item remained on the screen for 500 ms and then disappeared, but participants were allowed to make a lexical decision up to 5000 ms after the target. After the experiment, all participants completed the TELC Turkish placement test.

### Data analysis

The experiment yielded accuracy scores and RTs on the participants’ word/non-word button presses. Analysis of the RT data was carried out on correct responses only. All incorrect responses and timeouts were excluded from the analysis (CTR group: 3.71%; HS group: 5.29%). In order to normalize the distribution and

**Table 2.** Back-transformed mean RTs in milliseconds (and standard errors) and accuracy scores for the four conditions and the two participant groups

	Irregular	Regular	Orthographic	Semantic
HS Group				
Prime Type	Unrelated	Unrelated	Unrelated	Unrelated
RT	605.67 (6.41)	621.03 (5.92)	642.33 (7.70)	638.26 (5.61)
Accuracy	91.4%	96.6%	90.8%	93.8%
Prime Type	Related	Related	Related	Related
RT	588.83 (5.34)	596.60 (5.60)	629.95 (8.15)	618.58 (5.35)
Accuracy	98.3%	96.1%	94.0%	93.5%
Priming magnitude	16.84	24.43	12.38	19.68
CTR Group				
Prime Type	Unrelated	Unrelated	Unrelated	Unrelated
RT	576.42 (8.47)	579.65 (6.61)	590.08 (9.09)	594.87(6.70)
Accuracy	92.9%	98.3%	96.5%	96.5%
Prime Type	Related	Related	Related	Related
RT	563.81 (7.17)	563.49 (6.82)	585.45 (9.16)	593.32 (6.35)
Accuracy	98.3%	97.9%	94.5%	96.1%
Priming magnitude	12.61	16.16	4.63	1.55

reduce the influence of outliers, the remaining data were log-transformed (Ratcliff, 1993). In addition, RTs exceeding two standard deviations above and below a participant's mean log RT across all correct trials were deemed outliers and removed (CTR group: 4.67%; HS group: 5.01%). Furthermore, two items from the orthographic condition had to be excluded due to coding error. The remaining log RT data were then analyzed by using R, an open-source programming language and environment for statistical computing (R Development Core Team, 2017).

Linear mixed-effects regression models were used to analyze the RT data (Baayen, Davidson & Bates, 2008). The models included 'Subject' and 'Target' as random effects, and as fixed effects 'Condition' (irregular, regular, orthographic, semantic), 'Prime type' (related vs. unrelated), and 'Group' (CTR vs. HS) and their interactions. The models were fitted using the package *lme4* (Bates, Mächler, Bolker & Walker, 2015). Parameters were estimated with restricted maximum likelihood. For main effects and overall interactions, we employed sum-coded contrasts (-0.5, 0.5) to the factors Condition, Prime Type and Group. Priming effects for a single condition were analysed with treatment contrasts and by releveling for each condition. By employing backwards elimination, we started with a maximum random-effects model for each analysis. When the model did not converge, it was gradually simplified until convergence was reached (Barr, Levy, Scheepers & Tily, 2013). In the simplification process, random slopes by Subject and Target for each fixed effect in the model were only retained if they improved the model fit significantly, which was measured using the Akaike Information Criterion (AIC). The best-fit model with the lowest AIC score for each analysis is reported in the *Results* section.

If we recall, the language of HS has been found to exhibit a large degree of inter and intra-speaker variability. This is likely to also affect language processing. We therefore assessed the

variability of the priming effects, by determining the priming magnitudes of individual participants within each condition and comparing them across the two participant groups, using Levene's test for statistical evaluation.

## Results

Table 2 provides mean target RTs (back-transformed) and accuracy scores by Condition, Prime Type and Group.

As regards accuracy, Table 2 shows that both groups had high scores of around 90% or more correct responses across all conditions and prime types. We therefore did not perform any further analyses on these accuracy scores.

With respect to the RT data, Table 3 presents the results from the best-fit model testing for morphological priming effects. The overall model included fixed effects for Prime Type (related, unrelated), Condition (regular, irregular), Group (CTR, HS) and their interactions (see Table 3a). This model indicated significant main effects of both Prime Type and Group, the former due to shorter target RTs for related than for unrelated prime types, and the latter due to faster RTs for the CTR than for the HS group. The results from the above model yielded similar priming effects for the two conditions in the two participant groups. However, a lack of an interaction does not rule out the possibility that any differences were not detected, for example, due to limited power in the overall data set. We therefore ran additional linear-mixed effect models separately for the two morphological conditions and the two participant groups; see Table 3b. These models confirm significant priming effects for both irregular and regular aorist forms in both participant groups.

The same analyses were performed on the orthographic and semantic control conditions; see Table 4a to 4d. There were again main effects of Group (see Tables 4a and 4c), due to faster

**Table 3.** Fixed effects from the models of the two morphological conditions

Fixed Effects	Estimate	Std. Error	t value
<b>(a) Overall Model</b>			
Intercept	6.378	0.017	367.736*
Main effect: Prime Type (Related vs. Unrelated)	-0.034	0.006	-5.638*
Main effect: Condition (Regular vs. Irregular)	0.002	0.020	0.117
Main effect: Group (CTR vs. HS)	-0.061	0.029	-2.107*
Prime Type (Related vs. Unrelated)*Condition (Regular vs. Irregular)	-0.005	0.013	-0.357
Prime Type (Related vs. Unrelated)*Group (CTR vs. HS)	0.008	0.012	0.644
Condition (Regular vs. Irregular)*Group (CTR vs. HS)	-0.011	0.012	-0.869
Prime Type (Related vs. Unrelated)*Condition (Regular vs. Irregular)*Group (CTR vs. HS)	0.002	0.025	0.085
<i>Formula in R: log (RT) ~ Prime Type * Condition * Group + (1 + Prime Type * Condition   subject) + (1 + Condition * Group   target)</i>			
<b>(b) Model Split by Group and Condition</b>			
<b>HS Irregular</b>			
Intercept	6.379	0.022	290.604*
Prime Type (Related vs. Unrelated)	-0.034	0.010	-3.437*
<i>Formula in R: log (RT) ~ Prime Type + (1 + Prime Type   subject) + (1   target)</i>			
<b>HS Regular</b>			
Intercept	6.392	0.020	322.330*
Prime Type (Related vs. Unrelated)	-0.040	0.009	-4.330*
<b>CTR Irregular</b>			
Intercept	6.364	0.033	191.358*
Prime Type (Related vs. Unrelated)	-0.028	0.014	-1.946*
<b>CTR Regular</b>			
Intercept	6.362	0.027	239.677*
Prime Type (Related vs. Unrelated)	-0.032	0.015	-2.140*

RTs in the CTR than the HS group. For the orthographic condition (Table 4a and 4b), there were no further main effects or interactions in any participant group. More interestingly, however, Table 4d shows a significant semantic priming effect for the HS but not for the CTR group. Summarizing the findings thus far, we found genuine morphological priming in the CTR group, i.e., significant priming in the morphological but not in the orthographic and semantic overlap conditions. By contrast, the HS group exhibited both morphological and semantic priming effects.

As a follow-up, we sought to determine potential differences between the morphological and semantic priming effects in the HS group; see Table 5. To do this, we fitted models with the fixed effects of Prime Type (related, unrelated), Condition (irregular/regular, semantic), and their interactions to the RT data of the HS group. Note also that there is a numerical difference in the priming magnitudes between regular (23.41ms) and irregular (15.45ms) prime types in the HS group. We therefore decided to compare them separately against the semantic priming effects. The results revealed significant main effects of Prime Type and Condition, due to shorter RTs for related than for unrelated prime types and shorter overall RTs in the morphological than the semantic condition. Furthermore, there was no significant interaction of Prime Type and Condition, reflecting the fact that

both morphological prime types (irregular and regular ones) as well as semantically-related prime types yielded significant priming effects in HS.

Finally, we assessed potential differences in the variability of the priming effects obtained for the two participant groups. There are, of course, different ways of measuring variability of response times. Here, we were specifically interested in the distribution of the priming magnitudes within the two participant groups for the different experimental conditions. To this end, we calculated priming magnitudes for each individual participant, by subtracting the log-transformed RTs for the related primes from the log-transformed RTs for the unrelated primes, separately for the two morphological and the two control conditions in the two participant groups. We found that the HS group showed more variability in the distribution of the priming magnitudes than the CTR group, but only for two of the four conditions (viz. 'irregular', 'semantic'). Levene's tests confirmed this contrast. There was significantly more inter-individual variability in the HS group's priming magnitudes than for the CTR group with respect to the irregular condition (CTR group: mean: 14.18, SD: 44.86; HS group: mean: 22.97, SD: 81.53;  $F = 4.206$ ;  $p = 0.042$ ) and the semantic condition (CTR group: mean: 2.61, SD: 45.30; HS group: mean: 12.33, SD: 70.93;  $F = 5.563$ ;  $p = 0.019$ ), but not for the regular and the orthographic conditions (regular: CTR

**Table 4.** Fixed effects from the models of control conditions (orthographic and semantic)

Fixed Effects	Estimate	Std. Error	t value
<b>(a) Overall Model for Orthographic Condition</b>			
Intercept	6.423	0.023	284.426*
Main effect: Prime Type (Related vs. Unrelated)	-0.011	0.011	-0.985
Main effect: Group (CTR vs. HS)	-0.094	0.033	-2.880*
Prime Type (Related vs. Unrelated)*Group (CTR vs. HS)	-0.003	0.022	-0.115
<i>Formula in R: log (RT) ~ Prime Type * Group + (1 + Prime Type   subject) + (1 + Group   target)</i>			
<b>(b) Model Split by Group</b>			
HS: Orthographic Condition			
Intercept	6.463	0.027	240.190*
Prime Type (Related vs. Unrelated)	-0.009	0.013	-0.710
<i>Formula in R: log (RT) ~ Prime Type + (1 + Prime Type   subject) + (1   target)</i>			
CTR: Orthographic Condition			
Intercept	6.369	0.029	218.170*
Prime Type (Related vs. Unrelated)	-0.012	0.018	-0.660
<b>(c) Overall Model for Semantic Condition</b>			
Intercept	6.421	0.020	321.800*
Main effect: Prime Type (Related vs. Unrelated)	-0.013	0.008	-1.612
Main effect: Group (CTR vs. HS)	-0.074	0.032	-2.321*
Prime Type (Related vs. Unrelated)*Group (CTR vs. HS)	-0.020	0.016	-1.237
<i>Formula in R: as for Orthographic Condition</i>			
<b>(d) Model Split by Group</b>			
HS: Semantic Condition			
Intercept	6.441	0.023	278.458*
Prime Type (Related vs. Unrelated)	-0.022	0.009	-2.407*
<i>Formula in R: log (RT) ~ Prime Type + (1 + Prime Type   subject) + (1   target)</i>			
CTR: Semantic Condition			
Intercept	6.383	0.025	259.263*
Prime Type (Related vs. Unrelated)	-0.002	0.012	-0.207

group: mean: 17.53, SD: 59.84; HS group: mean: 26.79, SD: 62.45;  $F = 0.048$ ;  $p = 0.827$ ; orthographic: CTR group: mean: 6.03, CTR SD: 67.35; HS group: mean: 3.38, SD: 82.21;  $F = 2.454$ ;  $p = 0.120$ ). These results indicate that HS' performance is not generally more variable, but that increased inter-individual variability within the HS group relative to the CTR group may be modulated by linguistic factors, in the present case by the different prime types. Finally, we also explored the role of age of onset of exposure to German as a potential mitigating factor on RT variability, which considerably varied amongst our HS participants, with a range of 0 to 31 years. To this end, we determined whether AoA (taken as a continuous variable) correlates with RT variability in the experimental conditions. We did not find any such correlations, even for those conditions (viz. 'irregular' and 'semantic') that yielded significant between-group differences in RT variability ('irregular':  $r = -.082$ ; 'semantic':  $r = .062$ ). We conclude that age of onset of exposure to German does not modulate individual RT variability in the heritage speaker group.

## Discussion

The main findings from the current study are differences between the two participant groups, HS vs. non-HS of Turkish, with respect to their priming patterns. Whilst the non-heritage CTR group speakers showed pure morphological priming effects (without any facilitation in the orthographic and semantic control conditions), the HS group displayed significant priming effects not only for morphologically but also for purely semantically related prime-target pairs. Furthermore, we found that the HS group exhibited significantly more inter-individual variability in their priming magnitudes than the CTR group in the irregular and semantic conditions, but not in the regular and orthographic conditions.

### Comparisons to previous research

For the non-heritage CTR group we found morphological priming effects that are clearly dissociable from facilitation due to orthographic or semantic prime-target overlap, which is in line



**Table 5.** Fixed effects from the models of the morphological vs. semantic conditions for the heritage speaker group

Fixed Effects	Estimate	Std. Error	t value
<b>Irregular vs. Semantic</b>			
Intercept	6.424	0.020	323.089*
Main effect: Prime Type (Related vs. Unrelated)	-0.028	0.007	-3.915*
Main effect: Condition (Irregular vs. Semantic)	-0.056	0.021	-2.721*
Prime Type (Related vs. Unrelated)*Condition (Irregular vs. Semantic)	-0.011	0.013	-0.841
<i>Formula in R: log(RT) ~ Prime Type * Condition + (1 + Prime Type * Condition   subject) + (1 + Condition   target)</i>			
<b>Regular vs. Semantic</b>			
Intercept	6.432	0.019	338.795*
Main effect: Prime Type (Related vs. Unrelated)	-0.031	0.006	-5.026*
Main effect: Condition (Regular vs. Semantic)	-0.040	0.017	-2.401*
Prime Type (Related vs. Unrelated)*Condition (Regular vs. Semantic)	-0.019	0.015	-1.300
<i>Formula in R: as above</i>			

with the results of many previous masked-priming studies with L1 speakers of different languages (Marslen-Wilson, 2007). For Turkish HS, there are two previous masked-priming studies (Jacob & Kırkıcı, 2016; Jacob et al., 2019) to which we can compare our findings. One outcome that has been obtained in the two previous studies as well as in our study is the significant priming effect for regularly inflected word forms. By contrast, orthographic priming was only obtained by Jacob and Kırkıcı (2016) and semantic priming only in our study. How can these discrepancies be explained?

While orthographic facilitation has been documented in a number of priming studies for non-native second-language speakers (e.g., Heyer & Clahsen, 2015), Jacob and Kırkıcı (2016) is the only study that reported significant orthographic priming effects for HS. Jacob and Kırkıcı (2016) attribute their finding to specific properties of their HS, who acquired their literacy skills in German and are therefore likely to devote additional processing resources and attention to orthography when reading in Turkish, which may lead to orthographic priming effects. This does not, however, apply to our participants who acquired their literacy skills either in parallel for German and Turkish or first in Turkish ( $n = 59$  participants). Moreover, our HS had high reading/writing scores for Turkish (mean: 8.32 out of 10), which makes it unlikely that written Turkish is particularly demanding for them. Note also that Jacob et al. (2019) did not find any orthographic priming effects in their Turkish HS group either.

For our HS group, we obtained a semantic priming effect whereas Jacob et al. (2019) did not for their HS group; Jacob and Kırkıcı (2016) did not have a semantic condition in their study. We believe this discrepancy between our results and those of Jacob et al. (2019) is due to differences in the experimental materials of the two studies. Perea and Rosa (2002) reported that word pairs which are truly semantically related through antonymy/synonymy yield significantly higher semantic relatedness scores than word pairs that are just associated in meaning. Consequently, we only included antonyms or synonyms as ‘semantically related’ in our semantic control condition. By contrast, one third of the related item pairs in Jacob et al.’s (2019) semantic condition were meaning associates, e.g., *postane* → *mektup* “post office → letter”. Hence, the reason as to why Jacob et al. (2019) did not find a reliable semantic priming effect

could be that their item pairs were semantically less directly related than the ones we used.

### Explaining the priming results

Our findings from the non-heritage CTR group, i.e., reliable morphological priming effects without any orthographic or semantic priming, are not in line with non-morphological accounts that attribute priming effects for morphologically related words to orthographic and/or meaning overlap (e.g., Feldman et al., 2009). Instead, the CTR group’s priming patterns are consistent with both ‘affix stripping’ and dual-mechanism accounts of morphological processing. According to ‘affix stripping’ (Rastle & Davis, 2008) morpho-orthographic decomposition would apply to both regular and irregular aorist forms of Turkish. Note that unlike in Indo-European languages (e.g., English, German) in which irregular forms often have stem changes (e.g., *walked* → *walk* vs. *fell* → *fall*), the irregularity in Turkish aorist forms appears within the affix; compare, for example, the regular pair *duyar* → *duy* “hear” with the irregular one *gelir* → *gel* “come”. Consequently, morphological (stem/affix) decomposition ([duy]-ar, [gel]-ir) may yield a parallel stem-repetition effect from both regular and irregular prime words in the Turkish aorist. According to dual-mechanism accounts (e.g., Bosch et al., 2019), both morphological decomposition and morpho-lexical activation are involved in the processing of inflected word forms, the former typically for regularly inflected and the latter for irregularly inflected ones. Our finding of (largely) similar priming magnitudes for regular and irregular prime words (16ms vs. 12ms) in the CTR group then means that both routes are functioning (almost) equally efficiently in non-heritage L1 speakers of Turkish.

Explaining the findings from the Turkish HS group appears to be straightforward, AT FIRST SIGHT. Reliable priming effects were found for both the two morphological conditions and the semantic one. What is common to these three conditions is that primes and targets are related in meaning, unlike in the orthographic control condition. Thus, it seems as if the HS’ priming effects are caused by semantic relatedness of primes and targets and that morpho-orthographic decomposition is not necessarily involved. Note, however, that in addition to mean priming

magnitudes we also examined their variability within the two participant groups for the different experimental conditions. We assume that both measures (viz. priming magnitudes and variability of priming) tap into the same underlying processing mechanisms, with the priming magnitudes signaling a participant group's average performance for a given condition and the variability measure revealing the distribution across its individuals. We argue that, by including the results from both these measures, we get a more comprehensive picture of the HS group's performance and the mechanisms involved.

If HS were indeed relying on semantic relatedness only (unlike CTRs), the HS group's pattern of results should be parallel to the one of the CTR group for the morphological conditions, albeit for different reasons (viz. semantic relatedness for the HS group, morphological relatedness for the CTR group), but different for the semantic condition, because only the HS group is supposed to rely on semantic relatedness. This is not what we found, however. It is true that the semantic condition yielded a significant priming effect for the HS group only as well as significantly more inter-individual variability than in the CTR group. As regards the morphological conditions, on the other hand, we obtained a parallel pattern of results for the HS and the CTR group in the regular but not the irregular condition. In particular, the finding that the irregular condition yielded significantly more variability in the HS than the CTR group goes against any account that posits the same underlying mechanism for the HS group's performance on the two morphological conditions.

Instead, we argue that the HS group's priming patterns in the two morphological conditions have different sources. The HS group is indistinguishable from the CTR group in the regular condition, in that both groups exhibited significant priming effects with the same level of inter-individual variability. Assuming that parallel performance for the two participant groups signals parallel processes, we conclude that, for regularly inflected prime words, HS employ the same morphological decomposition mechanism as non-heritage CTR speakers. Concerning the irregular condition, we obtained a significant priming effect for HS (as for CTR speakers), but, unlike in the regular condition, there was also significantly more variability for the HS than the CTR group. Increased inter-individual variability indicates less homogeneous performance of the HS group than the CTR group for the irregular condition. Here we offer an interpretation for this finding in terms of a dual-mechanisms account of the Turkish aorist in which regular aorist forms are derived from a morphological rule (*Add -Ar*) whereas irregular ones have idiosyncratic lexical entries. Viewed from this perspective, the processing of irregular aorist forms is more reliant on lexical memory than the decompositional processes involved in the processing of regular forms. The efficiency of processing entries from lexical memory is known to be dependent upon frequency, and frequencies for lexical entries may be highly variable for HS given their individual linguistic experience. Therefore, the increased variability in the irregular priming condition within the HS group may reflect differences in the strength of memory traces for irregularly inflected verbs amongst the individuals of the HS group.

### *Processing limitations in HS?*

According to Polinsky and Scontras (2020a) many properties of heritage languages can be attributed to 'processing limitations'. The idea is that performing in a non-dominant language

challenges the language processing system, for example, by causing additional memory demands. It should be noted, though, that given the paucity of experimental studies on HS, Polinsky and Scontras (2020b, p. 53) readily qualify their claims related to processing limitations as not more than 'conjectures'. What do the current findings have to say about 'processing limitations' in HS?

We derived two specific predictions from this account for our HS participants: firstly, smaller priming effects for irregular than for regular aorist forms – due to additional memory demands for irregular forms (relative to regular ones); and, secondly, orthographic and/or semantic priming effects in HS – due to HS drawing on knowledge from other domains 'to optimize their resource use' (Polinsky & Scontras, 2020a: 14). These predictions were only partially confirmed. The results from the mean priming magnitudes revealed the same significant priming effects for both morphological conditions and both participant groups (contra the first prediction). On the other hand, the results of the additional variability analysis showed a regular/irregular contrast, with significantly more variability in the heritage than the non-heritage group for the priming magnitudes from irregular aorist forms (but not from regular ones). We attributed this contrast to increased variability in the strength of memory traces for irregular aorist forms amongst HS. If this is correct, Polinsky and Scontras' (2020a) idea of memory demands as the source of difficulties with irregular morphology may hold for some HS but not for HS taken as one group. We also found partial support for the second prediction in that the HS showed significant semantic priming. The semantic priming effect indicates that HS efficiently recruit additional resources during morpho-lexical processing, viz. semantic associations, that the non-heritage speaker group does not employ, in line with Polinsky and Scontras' (2020a) proposal. However, the same reasoning should apply to orthographic information, for which we did not find any priming effect, either in the HS or the CTR group. Furthermore, we found that the semantic priming comes on top of efficient morphological priming in HS. It is therefore not the case that other sources compensate for reduced grammatical ones in HS. Hence, our results do not lend much support for the idea of general processing limitations in HS.

### *Limits of variability in morphological processing*

Most experimental research on morphological processing has focused on discovering general (potentially universal) mechanisms of language processing, even though it is also obvious that individual differences arising from a range of linguistic and non-linguistic factors (e.g., cognitive capabilities, language input and use, task demands) can lead to variability in language performance. This then raises the question of whether this variability is random or whether there are limits on variability in language processing, for example, patterns of consistent behavior that resist modulation by an individual's cognitive capabilities, the amount of exposure and practice, or by task demands. It has been proposed that heritage languages and their speakers offer a wide range of inter and intra-speaker variability across all domains of language. This also applies to the HS group we tested, which was considerably more heterogeneous than the non-heritage CTR group with respect to demographic measures, input, exposure, and practice/use of Turkish. The HS group should therefore allow us to explore the variability of morphological processing and its potential limits.

Our results do indeed show more variability in the HS than the CTR group, but only for the semantic condition and the irregular morphological one. Given the heterogeneity of the HS group, there may be a multitude of factors that contribute to more variability in the HS than the CTR group. The current study was not designed to identify these factors. Instead, we sought to determine how different kinds of linguistic cues, specifically morphological ones, affect the variability of priming effects in HS' word recognition. Our most interesting finding on this matter is that the HS group performed like the CTR group in the regular morphological condition, with significant priming effects and the same level of variability in the two participant groups. As both participant groups yielded parallel results for this condition, we argued that HS and CTR speakers rely on the same morpho-orthographic decomposition mechanism for regular forms. This mechanism is apparently highly robust despite increased variability for our HS participants (relative to non-heritage CTR speakers) with respect to irregular and semantic priming. Another way of looking at these results, specifically the contrast between regular and irregular inflection in the HS group, is in terms of their linguistic representations. We argued that regular aorist forms might be rule-based and irregular forms lexically represented. If this is correct, the HS results may be taken as a case in which the grammar (qua *-Ar* affixation) selectively constrains variability in language processing.

**Data Availability Statement.** Data and materials are available via the Open Science Framework site for this project: <https://doi.org/10.17605/OSF.IO/3SJHA>

**Acknowledgements.** This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project Number: 317633480 – SFB 1287 Project B04. The authors wish to thank João Veríssimo and Laura Ciaccio for helpful comments on an earlier version of this paper. We are also grateful for valuable *BLC* reviewer comments that have led to many improvements of our original version.

## References

- Baayen RH, Davidson DJ and Bates D (2008) Mixed-effects modelling with crossed random effects for subjects and items. *Journal of Memory and Language* 59, 390–412.
- Baayen RH, Milin P, Durdevic DF, Hendrix P and Marelli P (2011) An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review* 118, 438–482.
- Barr DJ, Levy R, Scheepers C and Tily H (2013) Random-effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68, 255–278.
- Bates D, Mächler M, Bolker B and Walker S (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67, 1–48.
- Benmamoun E, Montrul S and Polinsky M (2010) *White paper: Prolegomena to heritage linguistics*. Harvard University Press.
- Bosch S, Veríssimo J and Clahsen H (2019) Inflectional morphology in bilingual language processing: An age-of-acquisition study. *Language Acquisition* 26, 339–360.
- Choi H. W (2003) Paradigm leveling in American Korean. *Language Research* 39, 183–204.
- Crepaldi D, Rastle K, Coltheart M and Nickels L (2010) 'Fell' primes 'fall', but does 'bell' prime 'ball'? Masked priming with irregularly-inflected primes. *Journal of Memory and Language* 63, 83–99.
- Davis CJ (2000) Match Calculator. Software. <http://www.pc.rhul.ac.uk/staff/c.davis/utilities/matchcalc/index.htm>
- Feldman L, O'Connor P and del Prado Martín F (2009) Early morphological processing is morphosemantic and not simply morpho-orthographic: A violation of form-then-meaning accounts of word recognition. *Psychonomic Bulletin & Review* 16, 684–691.
- Feldman L, Milin P, Cho K., Moscoso del Prado Martín F and O'Connor P (2015) Must analysis of meaning follow analysis of form? A time course analysis. *Frontiers in Human Neuroscience* 9 (111).
- Felsler C (2020) Do processing resource limitations shape heritage language grammars? *Bilingualism: Language and Cognition* 23, 23–24.
- Flores C and Rinke E (2020) The relevance of language-internal variation in predicting heritage language grammars. *Bilingualism: Language and Cognition* 23, 25–26.
- Forster LI and Forster J. C (2003) DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers* 35, 116–124.
- Frauenfelder U and Schreuder R (1992) Constraining psycholinguistic models of morphological processing and representation. The role of productivity. In Booij G and van Marle J (eds), *Yearbook of morphology*. Dordrecht: Kluwer, pp. 165–183.
- Gor K and Cook S (2010) Nonnative processing of verbal morphology: In search of regularity. *Language Learning* 60, 88–126.
- Gor K and Vdovina T (2010) Frequency, regularity and input in second language processing of Russian verbal inflection. *Slavic and East European Journal* 54, 7–31.
- Gor K, Chrabaszcz A and Cook S (2019) A case for agreement: Processing of case inflection by early and late learners. *Linguistic Approaches to Bilingualism* 9, 6–41.
- Gürel A (1999) Decomposition: To what extent? The case of Turkish. *Brain and Language* 68, 218–224.
- Gürel A (2020) Towards a comprehensive model of heritage language development. *Bilingualism: Language and Cognition* 23, 27–28.
- Gürel A and Uygun S (2013) Representation of multimorphemic words in the mental lexicon: Implications for second language acquisition of morphology. In Baiz S, Goldman N and Hawkes R (eds), *Proceedings of the 37th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press, pp. 122–133.
- Hankamer J (1989) Morphological parsing and the lexicon. In Marslen-Wilson WD (ed), *Lexical representation and process*. Cambridge, MA: The MIT Press, pp. 392–408.
- Heyer V and Clahsen H (2015) Late bilinguals see a scan in scanner AND in scanda: Dissecting formal overlap from morphological priming in the processing of derived words. *Bilingualism: Language and Cognition* 18, 543–550.
- Jacob G and Kırkıncı B (2016) The processing of morphologically complex words in a specific speaker group: A masked priming study with Turkish heritage speakers. *The Mental Lexicon* 11, 308–328.
- Jacob G, Şafak DF, Demir O and Kırkıncı B (2019) Preserved morphological processing in heritage speakers: A masked priming study on Turkish. *Second Language Research* 35, 173–194.
- Kırkıncı B and Clahsen H (2013) Inflection and derivation in native and non-native language processing: Masked priming experiments on Turkish. *Bilingualism: Language and Cognition* 16, 776–791.
- Köpke B and Schmid M (2004) First language attrition: The next phase. In Schmid M, Köpke B, Keijzer M and Weilemar L (eds), *First language attrition: Interdisciplinary perspectives on methodological issues*. Amsterdam: John Benjamins, pp. 1–43.
- Lehtonen M and Laine M (2003) How word frequency affects morphological processing in monolinguals and bilinguals. *Bilingualism: Language and Cognition* 6, 213–225.
- Marslen-Wilson W. D (2007) Morphological processes in language comprehension. In Gaskell G (ed), *The Oxford handbook of psycholinguistics*. Oxford: Oxford University Press, pp. 175–193.
- Montrul S (2008) *Incomplete acquisition in bilingualism: Re-examining the age factor*. Amsterdam: John Benjamins.
- Montrul S (2011) Morphological errors in Spanish second language learners and heritage speakers. *Studies in Second Language Acquisition* 33, 163–192.
- Montrul S (2012) Is the heritage language like a second language? *EUROSLA Yearbook* 12, 1–29.
- Montrul S and Mason S. A (2020) Smaller vocabularies lead to morphological overregularization in heritage language grammars. *Bilingualism: Language and Cognition* 23, 35–36.

- Nakano Y, Ikemoto Y, Jacob G and Clahsen H** (2016) How orthography modulates morphological priming: Subliminal kanji activation in Japanese. *Frontiers in Psychology* 7 (316).
- Nakipoğlu M and Ketrez N** (2006) Children's overregularizations and irregularizations of the Turkish aorist. In Bamman D, Magnitskaia T and Zaller C (eds), *Proceedings of the 30th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadia Press, pp. 399–410.
- Perea M and Rosa E** (2002) The effects of associative and semantic priming in the lexical decision task. *Psychological Research* 66, 180–194.
- Polinsky M and Scontras G** (2020a) Understanding heritage languages. *Bilingualism: Language and Cognition* 23, 4–20.
- Polinsky M and Scontras G** (2020b) A roadmap for heritage language research. *Bilingualism: Language and Cognition* 23, 50–55.
- Putnam M, Kupisch T and Pascual y Cabo D** (2018) Different situations, similar outcomes: Heritage grammars across the lifespan. In Bayram F, Denhowska N, Miller D, Rothman J and Serratrice L (eds), *Bilingual cognition and language: The state of the science across its subfields*. Amsterdam: John Benjamins, pp. 251–279.
- R Development Core Team** (2017) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>
- Rastle K and Davis MH** (2008) Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes* 23, 942–971.
- Ratcliff R** (1993) Methods for dealing with reaction time outliers. *Psychological Bulletin* 114, 510–532.
- Scontras G, Polinsky M and Fuchs Z** (2018) In support of representational economy: Agreement in heritage Spanish. *Glossa: A Journal of General Linguistics* 3, 1–29.
- Seidenberg MS and Gonnerman L. M** (2000) Explaining derivational morphology as the convergence of codes. *Trends in Cognitive Sciences* 4, 353–361.
- Sezer T** (2017) TS corpus project: An online Turkish dictionary and TS DIY corpus. *European Journal of Language and Literature* 9, 18–24.
- Taft M** (2003) Morphological representation as a correlation between form and meaning. In Assink E and Sandra D (eds), *Reading complex words*. Amsterdam: Kluwer, pp. 113–137.
- Taft M and Forster K. I** (1975) Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior* 14, 638–647.
- Uygun S and Gürel A** (2016) Processing morphology in L2 Turkish: The effects of morphological richness in the L1. In Gürel A (ed), *Second language acquisition of Turkish*. Amsterdam: John Benjamins, pp. 251–279.
- Uygun S and Gürel A** (2018) Processing second language inflectional morphology. In Akıncı MA and Yağmur K (eds), *The Rouen Meeting Studies on Turkic Structures and Language Contacts*. Wiesbaden: Harrassowitz Verlag, pp. 189–204.
- Verissimo J** (2018) Taking it a level higher: The LEIA model of complex word recognition. Presented at the 24th Architectures and Mechanisms for Language Processing Conference (AMLaP), Berlin.