Language Sciences 104 (2024) 101632

Contents lists available at ScienceDirect

Language Sciences

journal homepage: www.elsevier.com/locate/langsci

Print exposure leads to individual differences in the Turkish aorist☆

Tan Arda Gedik [a,b]

[a] FAU Erlangen-Nürnberg, Language & Cognition Lab, Bismarckstraße 6, 91054 Erlangen, Germany
[b] Bilkent University, Department of Psychology, 06800 Ankara, Turkey

ARTICLE INFO

Article history:
Received 3 August 2023
Received in revised form 21 February 2024
Accepted 5 March 2024
Available online 18 March 2024

Keywords:
Individual differences
Print exposure
Morphological productivity
The Turkish aorist
Nonce-verb conjugation

ABSTRACT

Several studies have established that not all native speakers extract the same generalization for a given construction due to speaker internal or external reasons, challenging a widely held assumption in linguistics. While there is a considerable number of studies investigating individual differences in grammatical knowledge in other languages, very little is known about how L1 Turkish speakers might manifest such differences in their linguistic knowledge. This is the first study to examine individual differences in the constructional representation of the Turkish aorist in adult L1 Turkish speakers. The aorist is known to be irregular and pose acquisition problems, especially when combined with monosyllabic sonorant ending verbs. The variants of the Turkish aorist have different corpus frequencies across spoken and written modalities. The study investigates to what extent differences in print exposure would lead to differences in how L1 Turkish speakers would apply the construction to monosyllabic-sonorant ending nonce-verbs. Based on the results, people with more written language experience extracted a more sensitive rule that applies to monosyllabic-sonorant ending nonce-verbs, such that they produced more -Ir than -Ar. Contrastingly, people who read less used more -Ar ($r = -0.35$), and print exposure accounted for roughly 12% of the variance. Our findings are compatible with usage-based approaches and suggest that print exposure-borne differences are pervasive in linguistic knowledge, adding to the growing body of evidence that challenges the convergence hypothesis.

1. Background

There is a general consensus that first language (L1) acquisition is a process that occurs largely unconsciously and automatically (e.g., Ellis, 1996; Paradis, 2009). Learning L1 grammar is assumed to rely on implicit learning mechanisms, which are thought to be separate from IQ scores or other variables such as linguistic aptitude or learning styles (Feldman et al., 1995; Reber et al., 1991). This belief, that L1 acquisition is not influenced by external (i.e., properties of input) or learner-internal (i.e., nonverbal IQ, motivation) factors, and other beliefs such as ease of acquisition and poverty of the stimulus (see Dąbrowska, 2015 for a detailed discussion) in generativist approaches arguably paved the way for what is called the convergence hypothesis. As conventional wisdom in linguistics, it proposes that all L1 speakers are uniformly successful in

☆ Reference (Self) Gedik, Tan Arda (in preparation). Literacy at Work: Ultimate Native Language Attainment. Book Manuscript.

E-mail address: tan.gedik@fau.de.

https://doi.org/10.1016/j.langsci.2024.101632

ultimate L1 attainment without showing much empirical evidence (see, for example, Bley-Vroman, 2009: 179; Chomsky, 1965, p. 11, 1975, p. 11; Crain and Lillo-Martin, 1999, p. 9; Crain et al., 2009, p. 124; Herschensohn, 2009, p. 264; Lidz and Williams, 2009, p. 177; Montrul, 2009, p. 4; Nowak et al., 2001, p. 114; Seidenberg, 1997, p. 1600).

However, recent studies show that it may not be the case. Individual differences in L1 knowledge are pervasive and exist on several different levels of linguistic systems such as knowledge of vocabulary, collocations, morphology, syntax, and pragmatics (e.g., see Dąbrowska 2012; Kidd et al., 2018 for a review) mostly because of print exposure. Research also shows that grammar in L1 speakers develops well into adulthood (Hartshorne et al., 2018). For instance, testing morphological productivity, Dąbrowska (2008) shows that L1 Polish speakers displayed varying levels of rule extraction for the Polish dative, and speakers that are more educated in Polish appeared to extract an across-the-board rule, whereas those with lower educational qualifications extracted local generalizations which only apply to a certain class of nouns.

These findings were argued to be a result of the facilitatory effect of print exposure as a result of being educated, because a) a more educated would arguably need to read more and b) written language contains more complex vocabulary and grammar than spoken language (e.g., Roland et al., 2007), and enhances linguistic knowledge in L1 on various levels, e.g., vocabulary (e.g., Stanovich and Cunningham, 1992), collocations (Dąbrowska, 2014, 2018), morphosyntax (Dąbrowska et al., 2022; Dąbrowska, 2018; Street, 2017, 2020), and production of passives, subject and object relatives (Montag and MacDonald, 2015).

As mentioned before, L1 learning is assumed to rely on implicit learning mechanisms, which are believed to be independent from IQ (Feldman et al., 1995; Reber et al., 1991). Nonverbal IQ is a learner-internal variable that is influenced by genetics and environmental factors such as socioeconomic status, stress, or diet (Lervåg et al., 2019), and is responsible for reasoning. However, research shows that L1 learning relies mostly on explicit learning mechanisms (i.e., explicit memory) (Llompart and Dąbrowska, 2020) or lays on a gradient spectrum between explicit and implicit learning mechanisms (Divjak et al., 2022), and that grammar is strongly correlated with nonverbal IQ, as well as with other levels of linguistic knowledge such as collocations, vocabulary, and inflectional morphology (Dąbrowska, 2018; Dąbrowska et al., 2022; Dąbrowska et al., 2023), however it should be noted that research on this is relatively scarce and controversial, thus it should be considered with caution.

Secondly, literacy is acquired through schooling where students are instructed to read written materials. Experience with reading increases metalinguistic awareness, such that speakers who read more perform better at tasks requiring the manipulation of linguistic knowledge in their L1 (e.g., Warren-Leubecker and Carter, 1988). This metalinguistic awareness may result in a more controlled, and reliable production of the L1 (Li and Wu, 2015), that is more sensitivity to distributions or co-occurrences of a construction as a result of more exposure, although research on this is controversial and more research is needed to establish its validity.

Furthermore, print exposure facilitates more across-the-board generalizations as well as further abstraction of a construction. For instance, Street and Dąbrowska (2014) show that people with more experience with written materials (as a result of having high academic attainment) can readily interpret more passive constructions than people with less experience (as a result of having low academic attainment): it is the degree of entrenchment of these verb-general and verb-specific representations that both groups acquired. In this case, the more educated group had entrenched these representations more. Contrastingly, people with less linguistic experience performed poorly when the passive construction stimuli involved less prototypical verbs, suggesting that their representation of the passive may only include verb-specific instantiations of the English passive. A similar finding was obtained in Street (2020), while both high and low academic attainment speakers were faster with prototypical items in the English passive construction (e.g., *attack*), reaction times for non-prototypical items (e.g., *bite*) were modulated by education, with higher academic attainment speakers showing shorter reaction times and being more accurate. This difference in education, once again, arguably results in how much speakers practice reading, leading to variance in print exposure as a concomitant effect. Thus, print exposure (as a by-product of education) leads to more abstract representations of the passive in both studies.

As a late-blooming construction, the Turkish aorist proves difficult to children during acquisition and they start to reliably produce it at the age of 7; 9 (Nakıpoğlu and Ketrez, 2006), and this interestingly coincides with the start of schooling. Schooling is important for native language attainment for two reasons; a) schooling and the concomitant acquisition of literacy modulate nonverbal IQ, and b) schooling provides grounds for exposure to written materials. As a recent meta-analysis shows, each year of schooling increases nonverbal IQ anywhere between 1 and 5 points (Ritchie and Tucker-Drob, 2018).

Pulling the strands together, it is not surprising that children start producing the Turkish aorist more reliably as they develop and are schooled. From a usage-based perspective, schooling-borne print exposure modulates the entrenchment and statistical preemption for the Turkish aorist, such that children learn how to generalize the suffix enough and avoid over-generalizing. This is based on the assumption that there are more types and tokens of the Turkish aorist in written materials than in spoken language, as written language contains more diverse input than spoken language (e.g., Biber, 1995; Cunningham and Stanovich, 1998; Roland et al., 2007).

As a matter of fact, a quick similarity search that looks up the suffix on the Turkish National Corpus (Aksan et al., 2012), a well-balanced corpus, shows that the aorist occurs 389,398 times in written, and 8533 times in spoken modalities. Because written language is lexically more diverse (e.g., Stanovich and Cunningham, 1992), it is safe to assume that written language contains more type and tokens of the aorist. Furthermore, because this experience with written materials increases nonverbal IQ and metalinguistic awareness, it could also help speakers provide the correct suffix form when prompted as a result of

heightened awareness of phonological environments and attention to such nuances, both as a concomitant result of reading in L1.

Nonverbal IQ may assist in paying attention to phonological environments in L1 that require the use of one form of the aorist over the other. This is an assumption based on L2 speakers where nonverbal IQ appears to fine-tune phonological discrimination in L2 learning (Georgiou, 2023; Georgiou and Giannakou, 2024), however research needs to be conducted with L1 speakers. Interestingly, Dąbrowska et al. (2023) found a statistically significant strong correlation between performance on a nonce verb inflection task in Spanish and Raven's Colored Progressive Matrices ($r = 0.63$, $p < 0.001$). The nonce verbs included both -ar and -er ending items, and semi-illiterates conjugated -er ending nonce verbs in 1st person with much less accuracy than -ar ending counterparts for the same paradigm (10% < 65%), late-literates outperformed semi-literates, and high-literates outperformed the other two groups. It is important to remember that IQ was highly correlated with group in this study. Therefore, a combination of all these factors (education, nonverbal IQ, experience among others) point to an increased linguistic experience and sensitivity to reliably producing the correct suffix when needed. Similarly, in light of evidence from previous usage-based studies (e.g., Dąbrowska, 2012, 2013, 2015, 2018, 2019), the choices of speakers on nonce verb inflection tasks should be influenced by IDs in linguistic input or cognitive machinery (i.e., nonverbal IQ).

As for the aorist, based on corpus evidence, Nakipoğlu and Michon (2020) argue that "the type frequency of -Ar … renders this possible. At the adult stage, participants' overwhelming -Ar use with sonorant-ending nonce-roots strongly confirms the unfolding of the abstraction and the default status of -Ar" (Nakipoğlu & Michon, p. 34). In other words, their claim is that adult Turkish speakers converge on -Ar as the linguistic representation of the aorist in combination with a novel sonorant ending monosyllabic verb in their mind.

Based on the Turkish National Corpus (Aksan et al., 2012), statistically a person who reads more on average is exposed to both -Ar and -Ir types; whereas a person with less print exposure is likely to hear more -Ar and fewer -Ir types. In other words, a person who does not read as much will predominantly have access to -Ar types and tokens, and 4 times less frequent -Ir tokens and its types (see Table 1).

Table 1 shows an estimated frequency of how many times the aorist, either in -Ir or -Ar form occurs in written or spoken modality (see methodology for how this was calculated). Thus, it is safe to assume that speakers with more print exposure should diverge from the local extraction of -Ar and use -Ir more often when faced with a nonce-sonorant ending monosyllabic verb. Whereas speakers with less print exposure should in theory converge more on -Ar because of fewer -Ir types and tokens.

Table 1
Estimated corpus frequencies derived from true positive rate of the first 200 lines in the Turkish National Corpus for monosyllabic + sonorant ending verbs ending in -Ar or -Ir.

		Written (per mil)	Spoken (per mil)	Total (per mil)
Monosyllabic + sonorant + the aorist suffix				
-ar	type	47	47	94
	token	37,694 (758.97)	2386 (2353)	40,080 (790.254)
-ır	type	14	14	28
	token	73,834 (1286.66)	596 (587.75)	74,430 (1467.53)

Note: Type count was taken from Nakipoğlu and Üntak (2008) and not included in the calculation of positive rates. Otherwise, there would have been 426 types of verbs that occur with -Ir in the spoken modality, which does not reflect the reality.

There have only been a few studies investigating the representation of Turkish constructions in L1 Turkish children (Nakipoğlu and Ketrez, 2006; Yıldız and Nakipoğlu, 2012), and one study that investigated the representation of the Turkish aorist in adult L1 Turkish speakers' minds (Nakipoğlu and Michon, 2020). To the researcher's knowledge, this is the first study to investigate IDs in adult Turkish L1 speakers. To fill this gap, this study is a partial replication of Nakipoğlu and Michon (2020). In this study as a modification to Nakipoğlu and Michon (N&M), print exposure is also controlled. Nonverbal IQ was not controlled in the current study as it could not have been operationalized in an online platform due to copyright issues.

The findings of the study will stimulate further discussions and assumptions made about the representation of constructions in the minds of native Turkish speakers. It also provides further converging evidence against the convergence hypothesis.

1.1. Usage-based approaches to L1 learning

Usage-based approaches to language learning, regardless of native or non-native learning, postulate that it is a frequency-driven journey with environmental or genetic factors playing an important role in predicting the outcome or the uniformity (e.g., Bybee, 2010; Kidd et al., 2018). There is a large body of evidence showing that L1 speakers activate high frequency words or other linguistic structures faster, and are better at recalling them (e.g., Divjak, 2019; Goldberg, 2019). In contrast to modular or generativist approaches (e.g. Pinker, 1997; Ullman, 2006), usage-based approaches suggest that language learning is a skill that is dependent on various cognitive machinery, i.e., perception, attention, phonological memory, explicit learning skills, abstraction, and potentially nonverbal IQ (e.g., Tomasello, 2003; Dąbrowska, 2018), although the full extent of the relationship is not clear. As such, the full mastery of a construction depends on both learner-internal factors as well as linguistic experience or the quality of the input (Divjak and Catherine Caldwell-Harris, 2015). Interestingly, there is evidence that humans differ in

their cognitive machinery drastically due to socioeconomic factors such as nutrition, diet, stress as well as genetic factors (Gruszka et al., 2010). For instance, nonverbal IQ, which appears to be correlated with language skills in L1 up to a certain extent (e.g., Dąbrowska, 2018; Dąbrowska et al., 2022), is heavily influenced by nutrition, diet and stress (Lervåg et al., 2019). Thus, differences in this machinery should theoretically result in differences in the representation of linguistic knowledge.

1.2. Individual differences in L1 knowledge

One of the major predictors of IDs in ultimate native language attainment is print exposure. It has been consistently shown that people who read more have bigger receptive vocabularies (e.g., Cunningham and Stanovich, 1998), and a more enhanced grasp of the grammar of the language they speak (e.g., Dionne et al., 2003; Huttenlocher, 1998). For instance, L1 English speakers with more print exposure produce more passives, and object relative clauses (Montag and MacDonald, 2015). Written language is known to consist of more complex structures (e.g., Roland et al., 2007; Özel et al., 2016). For instance, written Turkish is two times more likely to contain the Turkish passive construction, and around two times more likely to display the Turkish object relative construction. Similarly, written Turkish shows more complex words, i.e., words with more suffixes, than spoken Turkish (Özel et al., 2016). All of this variation is imperative for the continuation of L1 development, even later in adulthood. Studies show the importance of sustained exposure to L1 in children and adolescents (Nippold et al., 2005; Kaplan and Berman, 2015), and recently a study by Hartshorne et al. (2018) show that adult grammatical proficiency in L1 speakers seems to keep increasing until the age of 30.

Experience with written language is also interconnected to literacy or education –education impacts it, but is not necessarily a proxy for print exposure. Previous studies have shown a link between performing poorly on various constructions in English and having a low academic attainment (LAA, 12 years of formal education on average); whereas high academic attainment (HAA, 17 years of formal education on average) participants always outperformed LAA participants on tasks requiring knowledge of passives, relative clauses, and quantifiers (Dąbrowska, 2014, 2018; Street, 2017, 2020). This is because education and print exposure are related, as people who continue with higher education typically read more written materials or will have exposure to social circles that read more. This is usually the opposite for LAA people. Most LAA participants tend to work in manual labor jobs and have little or no time or interest in reading (Street and Dąbrowska, 2010).

This difference in experience with written language becomes clearer in experimental settings. For instance, Street and Dąbrowska (2014) show that the English passive representations of LAA participants were more likely to remain at a lexically-specific level than reaching a general, across-the-board generalization. In contrast, HAA participants performed at ceiling on both lexically-specific instantiations of and abstract schemas of the English passive. Similarly, Dąbrowska (2018) shows significant correlations between print exposure and performance on a picture selection task. Participants who had more education fared better at tasks which involved matching spoken sentences with a picture. One interesting finding was that good nonverbal IQ skills could compensate for a lack of print exposure and vice versa.

In another study, Dąbrowska (2008) tested the Polish dative construction in L1 Polish native speakers. The Polish dative has many form-meaning mappings because of gender and number declensions as well as the phonological environment of a word. Dąbrowska (2008) shows a strong and statistically significant correlation ($r = 0.48$) between participants' vocabulary knowledge, education and their target responses on a nonce-word conjugation task to elicit the Polish dative. Moreover, participants with less vocabulary knowledge opted to use only one of the forms of the Polish datives. This, as Dąbrowska argues, is expected because people who read more have bigger vocabularies, and a larger vocabulary size suggests having experienced more variation of a particular construction, in this case, the Polish dative construction. Therefore, Dąbrowska's conclusion is that people receive more education (who may have more experience with language) might arrive at more sensitive generalizations, whereas those who have less experience with written language extract only local generalizations and apply it across the board.

These findings are theoretically very interesting because for decades linguistics has assumed the convergence hypothesis as mentioned earlier, i.e., the idea that all native speakers are uniformly successful in arriving at overarching representations or global rules. In most cases, this idea was not supported with any empirical evidence (e.g., see for instance: Bley-Vroman, 2009: 179; Chomsky, 1965, p. 11, 1975, p. 11; Crain and Lillo-Martin, 1999, p. 9; Crain et al., 2009, p. 124; Herschensohn, 2009, p. 264). However, recent studies such as the ones mentioned here challenge this idea of attaining a form that supersedes others across all L1 speakers. It appears that people who read more may converge on what looks to be a similar language representation, although to what extent it is similar is up for debate considering IDs in cognitive machinery across people (Gruszka et al., 2010). Thus, language might be both a social phenomenon, accounting for the patterns in a speech community, and a cognitive one, accounting for patterns with IDs in speakers' minds (Dabrowska, 2020).

2. The Turkish aorist

The Turkish aorist construction is used on its own to realize tense/aspect/modality and has three allomorphs, the -r, -Ar and the -Ir forms[1] (e.g., Nakipoğlu and Ketrez, 2006). Furthermore, two of these three forms have different realizations in accordance with vowel harmony, e.g., -Ar, -er; -Ir, -ir, -ur, -ür. Acquisition of forms that have multiple functional pairings and a

[1] Capital letters indicate that the vowel can change in accordange with vowel harmony.

set of variation in the form make them particularly difficult for children to learn. Monosyllabic Turkish verbs combine with -Ar but monosyllabic verbs that end in a sonorant (e.g., /l/, /n/, /r/) take on -Ir; multisyllabic verbs take -Ir, with a few exceptions preferring -Ar (Nakipoğlu and Michon, 2020). Therefore, it is fair to call the Turkish aorist construction 'a construction that requires multiple local generalizations', as arriving at a superordinate construction that applies across the board requires many hours of exposure.

Previous studies on Turkish child language acquisition show that children tend to struggle with reliably producing the correct allomorph (i.e., -Ir or -Ar) until the age of 5, but this trend sometimes appears to last until the age of 7 or 8 (Nakipoğlu and Ketrez, 2006). In a seminal study, Nakipoğlu and Michon (2020) tested adult L1 Turkish speakers using a nonce-verb elicitation task and found that the majority of the speakers relied predominantly on using -Ar with nonce-verbs, even in sonorant ending nonce-verbs, and produced very few instances of -Ir, because the authors argue that ''an already attained symbolic abstraction … of -Ar takes the center'' (Nakipoğlu and Michon, 2020, p. 34).

While this may certainly be the case for some speakers, who rely on already existing memorized clusters to further generalize with -Ar (because -Ar in monosyllabic verbs is more frequent than -Ir), this may not be true for all L1 Turkish speakers. This is because not all L1 speakers display the same level of sensitivity to linguistic structures due to differences in education, linguistic exposure, and domain general cognitive mechanisms (e.g., memory, perception, attention among others, see Gruszka et al., 2010) (see discussion above). For instance, Dąbrowska (2013) shows in long distance dependencies in English, English native speakers display varying levels of verb-specific sensitivity when judging the grammaticality of such structures (i.e., some speakers find long distance dependencies with *think* more grammatical than some others). As such variation in this has been shown to result in individual differences in morphosyntactic knowledge in L1 (e.g., Kidd et al., 2018).

Reliably supplying the aorist also coincides with the start of schooling. Schooling provides explicit instruction in L1, which is also an important factor in L1 learning (e.g., Llompart and Dąbrowska, 2020), as well as more print exposure. As mentioned previously, the effects of print exposure on ultimate native language attainment are pervasive. Thus, the decreasing number of errors in producing the aorist in Turkish speaking children might be also connected to the quality of the input received in childhood.

Nakipoğlu and Michon's (2020) study follows the discussion on the single and dual-mechanisms (rule-free/analogy-based or rule-based/ analogy-free, respectively) for the acquisition and processing of inflectional morphology (e.g., Plunkett and Marchman, 1993; Prasada and Pinker, 1993). They created 168 nonce-verbs with a sonorant ending across 16 rhyme patterns. As discussed extensively in Nakipoğlu and Michon (2020), different rhyme patterns or n-grams with specific vowel–consonant combinations also trigger differing values of -Ir or -Ar responses from the participants. For instance, C-ur, C-ör, and C-öl are among the top patterns that trigger more -Ir responses than -Ar responses in Nakipoğlu and Michon (2020). They argue that frequency effects are in charge and speakers are sensitive to these distributions, as in these rhyme patterns, -Ir occurs with more types and tokens than -Ar. They also present patterns that should not trigger any -Ir because they do not occur with -Ir (e.g., C-ür, C-ır, C-ir, C-or, C-ıl, C-ül). Nakipoğlu and Michon (2020) also mention that /v/ as the onset of the nonce verb and /u/ in the vowel position in the model trigger the highest numbers of -Ir, which is explained by the high type token ratio verbs such as *vur* (to hit). Based on the low number of -Ir endings elicited overall, the authors argue that speakers might converge on -Ar as the attained symbolic representation of the Turkish aorist.

Although the single-mechanism does account for type-token frequencies up to a certain extent, approaching the Turkish aorist from a usage-based constructionist perspective (e.g., Goldberg, 2019) can account for a combination of both rule and analogy-based accounts of constructions or inflectional morphology. In a construction grammar approach, form and meaning are fused together and speakers are assumed to learn language via domain-general cognitive abilities such as perception, attention, and working memory among others (e.g., Goldberg, 2006). Construction grammarians believe that speakers start with lexically-specific templates and reach overarching generalizations or rules with more exposure (Divjak, 2019).

As such, L1 Turkish speakers possibly start with a limited number of -Ar verbs and reach a statistically more sensitive representation of the suffix (e.g., most multisyllabic verbs + -Ir, most monosyllabic sonorant ending verbs + -Ir). However, due to many individual differences in linguistic input or learner internal machinery, the extraction of this overarching representation might be slightly different across speakers, and thus lead to differences in representations. Taking a constructionist perspective also allows researchers to test the effects of print exposure, as was the case in previous individual differences studies (e.g., Dąbrowska, 2008, 2013, 2018; Street and Dąbrowska, 2014).

In the end, based on the differences of corpus frequencies of monosyllabic sonorant ending verbs that take -Ar and -Ir, we hypothesize that Turkish speakers who read more will be more likely to use -Ir, because they will show more sensitivity to the statistical distributions of -Ir and -Ar in monosyllabic sonorant ending verbs.

3. Methodology

3.1. Participants

Due to the COVID-19 pandemic, participants were recruited online from Turkish Facebook groups and the study was administered using Google Forms, whereas N&M's study was conducted in person using pen and paper. The study was designed in accordance with the Helsinki Declaration and the participants were given a consent form at the beginning of the

study. Data points from those who agreed to voluntarily participate in this study were recorded. Participants in this study were BA, MA, and PhD students. BA students were in different years, some had just started college, and some were senior students.[2] Data on gender, place of birth, and place of residence was not collected, as the Turkish aorist is not dialect or gender specific. The dataset is available in appendix C.

3.2. Procedure and materials

The Turkish National Corpus (Aksan et al., 2012) was used for the study to calculate the differences of verbs ending in -Ar and -Ir across written and spoken modalities. The TNC is a 50-million-word, well-balanced corpus that was developed in line with the British National Corpus. Using regex with different variations of the aorist suffix (e.g., -ar, -er, -ır, -ir, -ur, -ür), we retrieved a sample of two hundred lines and calculated a true positive rate across spoken and written modalities for monosyllabic sonorant verbs that occur with -Ar and -Ir. The similarity index in the corpus was dispreferred because it does not compute the frequencies of -Ar, -Ir, or -r separately. Because the TNC only allows access to the first 2005 lines when viewing results, the only way to observe the frequencies of the relevant endings was through a true positive rate analysis using the first 200 lines from the 2005 lines.

The stimuli were recreated in line with the methodology of Nakipoğlu and Michon (2020) (N&M) to test speakers' knowledge or -Ir and -Ar and were ensured to look and sound Turkish. To do this, in a norming study 20 native speakers were given a set of nonce consonant-vowel-sonorant monosyllabic verbs to judge how Turkish-like the verbs were (see appendix A for the nonce-verbs used in the experiment). The experimental stimuli followed c-v-sonorant; C-V-n, C-V-l, C-V-r; using /a/, /e/, /o/, /ö/, /u/, /ü/, /i/, and /ı/. Each rhyme pattern had 3 sets (e.g., 3 sets of C-V-n; C-V-l; C-V-r), $3 \times 6 \times 8 = 72$. Nonce verbs that resembled real words were discarded, which left 62 experimental stimuli. All 62 experimental stimuli +10 control verbs followed the consonant-vowel-sonorant pattern. Data from participants who did not engage were removed. Each participant took around 20 min to complete the survey. The number of experimental stimuli was kept at 62 instead of 168 in the original study. This is because finding voluntary participants online is difficult and adding additional stimuli would create fatigue and increase the chance of guessing. The rhyming patterns were kept from the original study although they were not the focus of this study. Stimuli were randomly assigned to 7 tenses as conducted by N&M (2020, p. 25), then presented in a context with adverbs that would trigger the aorist.

a. Ipek dülmeyi sever. '*İpek likes to dül.*'
Ipek her gece dül____. '*İpek every night dül____*'

Data was coded for machine readability and was analyzed using R (R Core Team, 2021). There were a total of 87 participants prior to data cleaning and pre-processing. Data from participants who did not answer all 10 of the control stimuli and participants whose overall production was lower than 90% were discarded. Control stimuli were real Turkish monosyllabic verbs to ensure that speakers were attending to the study. This 90% production rate translated to the successful completion of 55 out of 62 nonce-verb conjugations. Reasons for failing to conjugate a nonce verb were misunderstanding the task and supplying a different form (i.e., -DI, the simple past tense suffix), or failing to supply the correct suffix on the control stimuli. In the end, there was data from 51 participants with 90% or more accuracy.

3.2.1. Print exposure: reading time & attitudes questionnaire

Print exposure data were collected by means of a previously developed and piloted self-reported reading time and attitudes questionnaire, based on a similar questionnaire used by Dąbrowska (2014, 2019). While other methods of collecting print exposure data are arguably more reliable (i.e., author recognition tasks) because people tend to give socially desirable answers on print exposure questionnaires (e.g., Acheson et al., 2008; Stanovich and Cunningham 1992; Stanovich and West 1989), a Turkish author recognition task had not been developed at the time of this study. The print exposure questionnaire has a total of 5 questions. While four questions ask for the same information, two of them are formulated to see how much they read in work or school contexts, and the other two ask how much they read on their own initiative. Participants are asked to rate how often they read emails, messages, newspapers, books as fiction and non-fiction, social media, comic books among others (see appendix B). A wide variety of genres were included because written language is not uniform and differs in complexity depending on the format or genre (e.g., Biber, 2009). The final question asks participants to rate statements about reading. All answers were transformed into numerical codes, 0 through 6, with higher numbers indicating more frequency of reading. Scoring on the final question was reversed for negative answers, with 0 through −6. The maximum score a person could obtain on the print exposure questionnaire was 180. In the end, the scores were summed to create the 'print exposure' score. In a previously piloted study where the same questionnaire was used to collect data from 81 L1 Turkish speakers, the Cronbach's alpha was found to be 0.89. Such a number indicates reliability and internal-consistency among questionnaire items.

[2] One of the reviewers pointed out that the participant sample in this study may be WEIRD, and it might be problematic to generalize our findings from WEIRD to non-WEIRD Turkish speakers. Turkish universities have a tier system and Turkey's education system does not necessarily equate university enrollment with high education levels. Thus, it is more difficult to draw a parallel between WEIRDness and being a university student. However, we do acknowledge that the pattern we see here might be very different for illiterate speakers or speakers who had very little formal schooling.

3.3. Regression analyses

The effect of the three predictors (education, age, print exposure) on the preference of the aorist suffix in the nonce-verb conjugation task was calculated and measured by using standard multiple regression modeling in R (R Core Team, 2021). The codes used are available in appendix C. The initial model contained all three predictors (age, education, and print exposure as independent variables), but nonsignificant predictors were removed one after the other after running a log likelihood test, starting with the predictor of the highest order interaction, as suggested by Crawley (2010). If the predictor did not improve the fit of the model, it was removed. This process was done for all the predictors. To interpret the model coefficients more effectively, the relative importance of each predictor was measured using the lmg metric, computed by the relaimpo package in R (see Grömping, 2007). This metric is calculated by averaging the sum-of-squares obtained from all possible orderings of the predictors in the model, and is thought to be analogous to a squared semi-partial correlation. Larson-Hall (2010) argues that it quantifies the variance explained by each predictor in the model.

4. Results

After removing ineligible data points, there were a total of 3104 suffix productions from 51 participants. Of the elicited suffixes, 62.37% (1936) was -Ar, and 37.62% (1168) was -Ir. While many speakers provided -Ar, an important number of participants also supplied -Ir. (see Fig. 1).

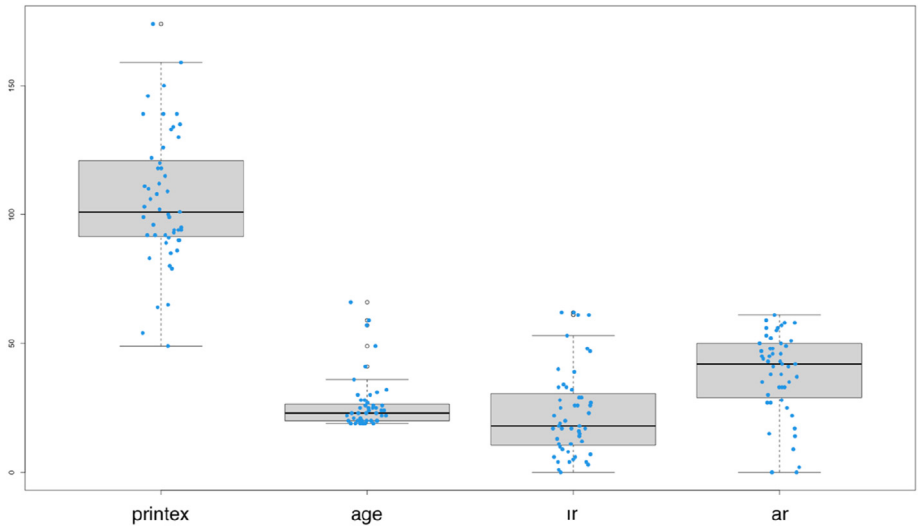

Fig. 1. Boxplots of variables showing individual differences.

Table 2 shows the proportions of each variable. As it is seen, there is considerable individual variation in print exposure, and the preference of the suffixes. -Ar on average is used more often than -Ir, but there is also variation of 16 points on average in supplying either of those suffixes.

Table 2
Proportions of variables, standard deviations, ranges, and interquartile ranges.

Variables	Mean	Standard Deviation	Range	Interquartile Range
Print Exposure	106	25.79	49–174	29.5
Education	1.76	0.929	1–4	1
Age	26.2	10.52	19–66	6.5
Ar	37.96	16.64	0–62	21
Ir	22.9	16.90	0–61	20

Fig. 2 visualizes the correlations presented in Table 3. Pearson correlations point to a statistically significant negative correlation between -Ar and print exposure ($r = -0.348$, $p = 0.01$), and a statistically significant positive correlation between -Ir and print exposure ($r = 0.35$, $p = 0.01$). Figs. 3 and 4 show the relationship between -Ar, -Ir and print exposure, respectively. -Ar and -Ir are very strongly correlated as the participants only had a binary choice; they could either provide -Ir (or its various forms in accordance with vowel harmony) or -Ar (or its various forms in accordance with vowel harmony). This is quite normal, and as a result, the correlations between *-Ar*printexposure* and *-Ir*printexposure* are very close to one another.

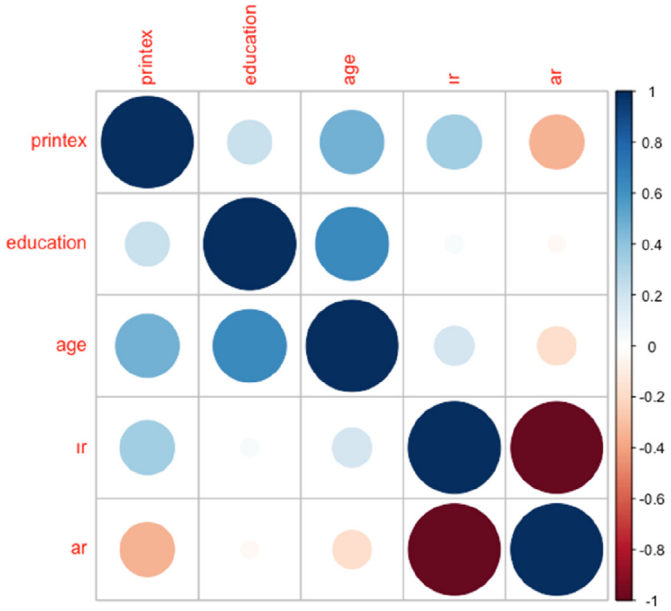

Fig. 2. Correlation matrix between all the variables.

Table 3
Pairwise correlations between all the variables.

Variables	Print exposure	Education	Age	Ar	ır
Print exposure	1.000	0.229	0.477***	−0.348**	0.350**
Education	0.229	1.000	0.630****	−0.034	0.038
Age	0.477***	0.630****	1.000	−0.176	0.184
Ar	−0.348**	−0.034	−0.176	1.000	−0.994
Ir	0.35**	0.038	0.184	−0.994****	1.000

Asterisks indicate different levels of statistical significance: $p < 0.0001$ '****'; $p < 0.001$ '***', $p < 0.01$ '**'.

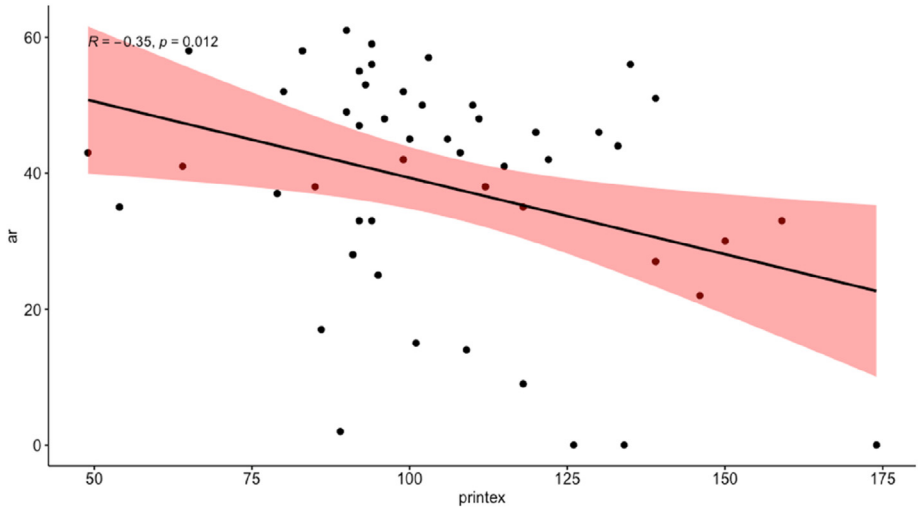

Fig. 3. The relationship between print exposure and -Ar.

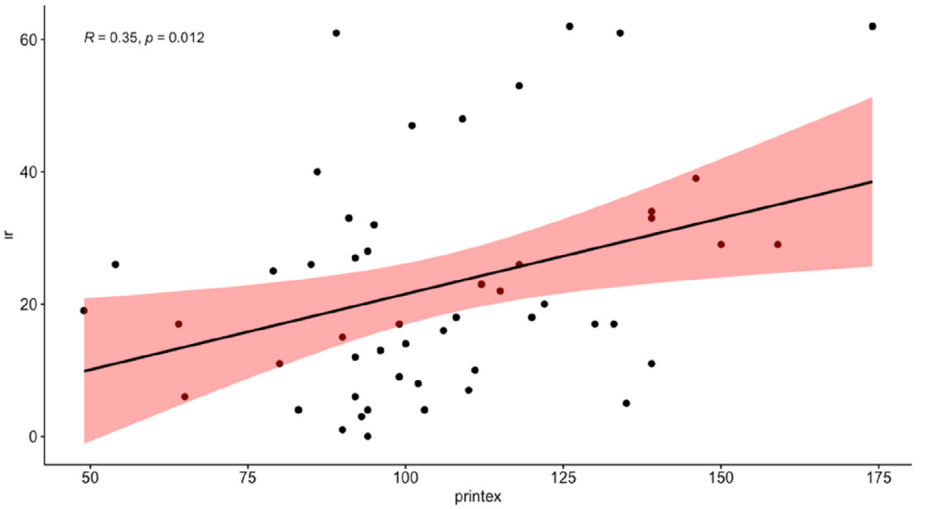

Fig. 4. The relationship between print exposure and -Ir.

Age, as expected from our sampling, is very strongly correlated with education ($r = 0.63$, $p = 0.0001$). This is also expected since people tend to attain further degrees as they mature. Age also showed a statistically significant correlation with print exposure ($r = 0.47$, $p = 0.0004$). Considering that age is strongly correlated with education in this sample, the correlation between print exposure and age is suggestive of the fact that people who have a higher academic attainment read more often in Turkish. Education is also weakly correlated with print exposure ($r = 0.22$, $p > 0.5$). The reason for a weak correlation might be due to the way in which education was operationalized in this study. Participants were asked to indicate their highest educational attainment from a choice of four options (i.e., High school, BA, MA, PhD). If education had been operationalized as the number of years spent in formal schooling, the correlation may have been stronger (see Dąbrowska, 2018 for an example).

As it was explained in 4.3, non-significant interactions and predictors were removed one after the other unless they improved the fit of the model statistically significantly. There were no significant interactions. This process only left one predictor in both models, namely print exposure. Tables 4 and 5 present the final model for -Ar and -Ir, respectively. As is clear from both Tables 4 and 5, print exposure displays statistical significance at $p < 0.05$, and it accounts for 12.1% in both models. This similar percentage across both final models is because of the binary nature of the experiment and the coding of the responses. One striking finding based on the coefficients is, as print exposure increases, speakers produce fewer instances of -Ar, and more instances of -Ir.

Table 4
Final model for -Ar.

Variable	Estimate	Standard error	*t* value	Pr(>|t|)	lmg
Intercept	61.77440	9.41565	6.561	3.18e-08***	
printex	−0.22474	0.08639	−2.602	0.0122*	0.121
Model R^2					0.121

Asterisks indicate different levels of statistical significance: $p < 0.001$ '***', $p < 0.05$ '*'.

Table 5
Final model for -Ir.

Variable	Estimate	Standard error	*t* value	Pr(>|t|)	lmg
Intercept	−1.36940	9.56057	−0.143	0.8867	
printex	0.2290	0.08772	2.611	0.0119*	0.121
Model R^2					0.122

Asterisks indicate different levels of statistical significance: $p < 0.05$ '*'.

4.1. Discussion & implications

The results so far show that print exposure is by far the most important predictor in the regression model. Such that print exposure accounts for roughly 12% of the individual differences in outcome of the suffix. Therefore, there is some evidence that even in a highly literate group of L1 adult Turkish speakers (as most of the participants in this study were undergraduate students or graduate students), literacy-related individual differences can be observed. Put differently, production of the aorist suffix shows frequency-driven trends as a result of being exposed to more -Ir instances in reading combined with

spoken language than they arguably experience -Ar instances. This production in return suggests that L1 Turkish speakers who read more extract a more sensitive representation to the fact that the aorist occurs with a monosyllabic sonorant ending verb. Contrastingly, those that read less extract a less sensitive generalization to the statistical distribution of the aorist. This fits well with the findings of Dabrowska (2008), in which she shows that speakers with more linguistic experience extracted rules that applied across-the-board.

Age and education were added as predictors because older speakers will have been exposed to more language. A conservative estimate of 8 h of daily exposure would add up to 2912 h per year. A 20-year-old (58,240 h in 20 years) will differ significantly from a 60-year-old speaker (174,720 h in 60 years). Furthermore, age has been reported to be correlated with diachronic change in other constructions (see Bamyacı, 2016 on optional plural marking), however this does not seem to be the case for the aorist. Similarly, education was included in line with previous studies (Dabrowska, 2008), as people with more education will have read more written materials. That is, we believe education impacts print exposure but it cannot be a proxy for it. Interestingly, age and education proved to be insignificant. It would be interesting to see how much these two predictors would influence the outcome in the wider Turkish population (i.e., participants with little formal education or illiterates).

With -Ir being more frequent in the written modality, a speaker who reads more should ideally encounter more of the -Ir form than someone who is older but does not read as much. As per our results, age does not appear to be a significant predictor. This once again underlines the importance of quantity and quality of input in the L1 and undermines the idea of a final adult state that speakers arrive at with maturation (Elman, 2001; Seidenberg and MacDonald, 1999). It also provides more evidence that the effects of print exposure are over and beyond maturational effects. If anything, L1 speakers appear to exhibit pervasive individual differences as a result of extracting different generalizations from ambient input (spoken and/or written language). This might prove problematic for generative approaches that offer maturation as the only driving force.

As pointed out in Nakipoğlu and Michon (2020), the consonant of the sonorant monosyllabic verb that is similar to the consonants used in -Ir types, the vowels in -Ir types, and the trigrams seem to trigger various levels of -Ir and -Ar. These similarity indexes, as referred to by Nakipoğlu and Michon (2020), were not carried out in the present study for three reasons. (1) It is time consuming to identify the corpus frequencies of each rhyme used in the study; (2) individual differences are pervasive and as such each rhyme pattern may reveal differing levels correlations with print exposure; and (3) statistically speaking, fitting in multiple rhyme patterns in a model would be difficult, and may not yield observable results given the sample size. Nevertheless, this does not invalidate our findings as we provide a general outlook on individual differences in the representation of the aorist suffix, whereas looking at these similarity factors may reveal more fine-grained individual differences.

Among many variables that may affect how much an L1 speaker attunes to these three factors mentioned in Nakipoğlu and Michon (2020), there are the effects of print exposure as these similarity indexes will occur at different frequencies in spoken and written modalities. Then, there is language aptitude or other forms of tests that can measure how much a speaker is sensitive to minor details in language which correlates with grammatical performance in L1 (see for instance Dąbrowska, 2018; Skehan and Ducroquet, 1988). Finally, there are individual differences in phonological memory skills that can affect later vocabulary and morphosyntax development (e.g., Kidd and Donnelly, 2020).

While the current study did not test for these factors or the similarity indexes, accounting for these variables is important as they may shape how (in)sensitive an L1 Turkish speaker might become to the local extraction of -Ar for the monosyllabic consonant-vowel-sonorant constructions when combined with the aorist. A similar example is provided by Dąbrowska (2013) on long distance dependencies in L1 English speakers, where she shows that not all speakers are equally sensitive to the same rule or did not extract the same rule.

A similar case may be made here, it is possible to suggest that our findings are a result of (including but not limited to) individual differences in print exposure. Thus, based on our findings, we can suggest that some L1 Turkish speakers extract different generalizations of the aorist suffix when it is combined with monosyllabic consonant-vowel-sonorant verbs.

This is the first study to investigate individual differences in L1 adult Turkish speakers. Our results provide further converging evidence for a) L1 learning is a usage-based phenomenon, b) this frequency can be modulated by print exposure, c) L1 learning continues well into adulthood, and d) differences in L1 linguistic input results in individual differences. Our study confirms and provides further psycholinguistic evidence that language learning is frequency-driven (e.g., Bybee, 2010). This frequency can be facilitated by reading as written materials contain more complex language (e.g., Özel et al., 2016; Roland et al., 2007). Therefore, individual print exposure scores should theoretically be correlated with production tasks that test constructions with differing frequencies across spoken and written modalities.

Based on our corpus count of the aorist suffix and results, our hypothesis was confirmed. That is, people who read more would be more likely to provide the -Ir variant as written language combined with spoken language would lead to a frequency advantage in arriving at a more sensitive generalization for C + V + sonorant verb constructions when combined with the aorist. Whereas speakers with less written language experience would see fewer -Ir instances, and more -Ar instances, and this would lead to a less sensitive, local generalization.

The results show that -Ar is not necessarily the attained symbolic generalization for every L1 adult Turkish speaker, rather, it is a local generalization that is represented in the minds of speakers that read less. People who read more tend to show more sensitivity to the fact that -Ir may also occur with nonce-monosyllabic sonorant ending verbs, and as such provide more -Ir. Interestingly, this also shows that even in adulthood, language learning or abstraction is still ongoing as a result of varying linguistic experience, otherwise such individual differences would not have been observed. Finally, previous studies on

individual differences use author recognition tasks (e.g., Llompart and Dąbrowska 2020; Dąbrowska 2018, 2019), as questionnaires tend to attract more socially desirable answers (e.g., Acheson, 2008). In light of this, it is quite interesting to see a correlation and variance accounting of this size given the fact that print exposure was assessed using a questionnaire as at the time of the study a Turkish author recognition task had not yet been developed.

These findings are also interesting from a theoretical perspective. Most linguists claim that L1 acquisition is completed by age 3, 4, or 5 (Hirsh-Pasek and Golinkoff, 1999; Hoff, 2009; Pinker, 1995) and that L1 speakers converge on the same grammar (e.g., Chomsky, 1965). The results of this study provide further converging evidence that even in adulthood speakers' generalization may be modulated by means of print exposure, a proxy for more varied linguistic input. A similar finding was demonstrated in Dąbrowska (2018) in which she found that well-into adulthood, L1 English speakers displayed statistically significant correlations between vocabulary, collocations, and grammar, and print exposure. Similarly, the results of this study also contribute to the argument that not all L1 speakers converge on the same grammar, and some show differences as a result of print exposure (e.g., Dąbrowska, 2018; Huettig and Pickering, 2019).

If all participants in this study had converged on the same linguistic representation, then there would have been no correlations or variance accounted for by print exposure. The idea is that while some L1 speakers may extract all there is to extract about a particular linguistic structure, a significant majority of the speakers also extract local or less sensitive generalizations of it and some may never arrive at a detailed representation with all there is to represent, for instance as evidenced by Street and Dąbrowska (2014).

This is important for theoretical considerations in language acquisition theories, especially for views engendered by generativist approaches. This also applies to usage-based studies whereby such findings are readily predicted in the theory, but are not entertained. Therefore, both schools of thought should aspire to have more theoretical and experimental plausibility. Generativist approaches also need to explain how individual differences in adult morphosyntactic knowledge can be accounted for in a universal grammar framework. Specifically, if some speakers extract local and some others extract more general schemas, how does this fit in with top-down approaches of universal grammar?

It is important to be careful with the implications of this study. While our research does not explicitly convey the notion that 'the more you read, the better your command of the language,' we acknowledge that it is crucial to prevent any misinterpretation that may arise from our findings. Indeed, the complexities of Turkish linguistic rules, particularly those related to the aorist suffix, make it apparent that exposure to written material can contribute to a deeper understanding of these structures. However, it is equally valid to recognize that language proficiency is a multifaceted construct and that formal reading alone does not encompass the entirety of linguistic knowledge. Importantly, individual differences in grammar may result from different reasons and some grammatical structures may be more likely to display individual differences than others. Thus, one question that linguists should try to answer is to what extent individual differences are equally distributed across grammatical structures.

While the role of nonverbal IQ in conjugation is a question, it was not controlled for in this study as copyright issues would have rendered an online study impossible. However, based on a pioneering recent study done by Dąbrowska et al. (2023), nonverbal IQ, measured by using Raven's Colored Progressive Matrices, is highly correlated with conjugating nonce-verbs in Spanish correctly in accordance with tense, aspect, and the correct variant of a suffix ($r = 0.63$, $p < 0.001$). However, in a regression analysis, nonverbal IQ did not show any statistical significance and only accounted for a small percentage in the model because highly literate participants had high IQ scores in their study design. While it is difficult to predict how the Turkish dataset regression analyses would have appeared with nonverbal IQ in the model, based on Dąbrowska et al. (2023), nonverbal IQ would potentially appear to be correlated but account for a limited percentage of variance. Another argument is that because in Spanish suffixes are conjugated in terms of number, gender, tense and aspect, nonverbal IQ might be correlated as it is a proxy measurement of reasoning (e.g., Garcia-Navarro et al., 2020), and conjugating a verb to agree with multiple factors arguably requires some level of reasoning as evidenced by the correlation in Dąbrowska et al. (2023). However, for the Turkish aorist, especially in an experimental design like the current one, there may not be any moderate correlations or variance accounting between nonverbal IQ and the aorist preference, as the aorist in the current setting does not have to be conjugated for tense or aspect.

Future research studies on individual differences in the representation of the Turkish aorist should consider replicating such findings with more participants, more stimuli, participants from varying backgrounds of print exposure, and illiterate participants. Similarly, the current study did not investigate rhyme patterns or other similarity indexes and the correlation between them and print exposure. It might be useful to uncover the relationship between these variables. It may also be worth investigating the correlation between other cognitive measures such as phonological short-term memory and language aptitude in addition to print exposure and the performance on the nonce-verb conjugation task. Finally, replicating a similar study with a Turkish author recognition task (or other more valid print exposure measures) might also be useful to obtain more reliable correlations between the aorist and print exposure scores.

5. Conclusion

This study investigated print exposure related individual differences in eliciting the Turkish aorist in L1 Turkish speakers using a nonce-verb conjugation task. It provides further evidence for a long-standing assumption in usage-based linguistics, that individual differences in L1 learning (i.e., input) would lead to individual differences in the representation of various

constructions in the mind of L1 speakers and suggests that the convergence hypothesis needs to be revised. Previously when speakers are forced to pick between -Ir and -Ar, it was postulated that the -Ar takes the attained symbolic representation for monosyllabic sonorant ending Turkish verbs (i.e., Nakipoğlu and Michon, 2020). However, this study shows that adult L1 Turkish speakers with various levels of print exposure show different extractions of the Turkish aorist. Our results suggest that people who read more extract a statistically more sensitive generalization as to how the Turkish aorist can be used with monosyllabic sonorant ending nonce-verbs. People who read less tended to produce more -Ar, and people who read more produced more -Ir on sonorant ending monosyllabic nonce-verbs. Print exposure varied for about 12% of the individual differences in the elicited endings. Our study shows that print exposure related differences can be captured even in relatively smaller linguistic units, i.e., suffixes, in a highly literate group (e.g., university students) and our results are problematic for modular approaches to language acquisition.

CRediT authorship contribution statement

Tan Arda Gedik: Writing – review & editing, Writing – original draft, Project administration, Methodology, Data curation.

Data availability

The link to the data is available in Appendix

Acknowledgments

I would like to sincerely thank all of the participants that partook in this study, the lab members at the Language and Cognition Lab for their help with an earlier version of this manuscript, and the two reviewers for their valuable feedback. All remaining errors are mine.

Appendix

Appendix A

The nonce-verb matrix without the verbs that coincide with real words. Instead, control stimuli were inserted. Asterisks indicate control stimuli.

CVS	a	e	ı	I	o	ö	u	ü
ç*n	gör*	çen	çın	çin	çon	çön	çun	çün
p*n	pan	pen	pın	pin	pon	pön	pun	pün
r*n	ran	ren	rın	rin	ron	rön	run	rün
y*r	çal*	ver*	yır	yir	yor	yör	yur	sür*
z*r	kal*	zer	kır*	zir	sor*	zör	zur	zür
m*r	mar	mer	sız*	mir	koy*	mör	mur	mür
ş*l	şal	şel	şıl	şil	şol	şöl	şul	şül
c*l	cal	cel	cıl	cil	col	cöl	cul	cül
h*l	çat*	hel	hıl	hil	hol	höl	hul	hül

Appendix B

Okuma Alışkanlıkları (Print exposure questionnaire)

1) Aşağıdaki sorular, okuma alışkanlıklarınızı ölçmektedir. “Okuma” dediğimizde özellikle anlama, yazılı metinleri kullanma ve düşünmeyi kast ediyoruz. Bu beceri, kişinin hedeflerine erişmesi, bilgisini ve kendisini geliştirmesi, ve toplumda yer alabilmesi için gereklidir.

Önemli not: Aksi belirtilmedikçe, “okuma” dediğimizde hem basılı hem dijital **Türkçe** kaynaklardan bahsediyoruz.
Lütfen her soruyu dikkatlice okuyunuz ve sorulara olabildiğince dürüstçe cevap veriniz. Bu ankette doğru veya yanlış cevaplar yoktur. Cevaplar sadece sizin için doğru olmalıdır.

A) Ne sıklıkla aşağıda belirtilen okuma aktivitelerinde bulunursunuz?

	Asla veya çok nadir	Yılda birkaç kere	Ayda bir kere	Ayda birkaç kere	Haftada birkaç kere	Günlük veya neredeyse günlük
Romanlar						
Kurgu olmayan kitaplar						
Karikatürler						
Magazinler						
Gazeteler						
İnternet sayfaları						
Sosyal medya (Facebook, Twitter, Instagram gibi)						
Mesajlar						
E-postalar						
Çevrimiçi gruplar / forumlar						

Yorum(lar):

B) Ortalama ne kadar zamanınızı iş veya okul için **Türkçe** okuyarak geçiriyorsunuz?
- Hiç okumuyorum.
- Günlük 30 dakika veya daha az
- Günlük 30 ve 60 dakika arası
- Günlük 1–2 saat
- Günlük 2–3 saat
- Günlük 3 saatten fazla

C) Hobi olarak **Türkçe** okumak için genelde ne kadar vakit geçiriyorsunuz?
- Hobi olarak okumuyorum.
- Günlük 30 dakika veya daha az
- Günlük 30 ve 60 dakika arası
- Günlük 1–2 saat
- Günlük 2–3 saat
- Günlük 3 saatten fazla

D) Ne sıklıkla aşağıdaki materyalleri **siz istediğiniz** için **Türkçe** okuyorsunuz?

	Hiç veya neredeyse hiç	Yılda birkaç defa	Ayda bir kere	Ayda birkaç kere	Haftada birkaç kere	Her gün veya neredeyse her gün
Romanlar						
Kurgu olmayan kitaplar						
Karikatürler						
Magazinler						
Gazeteler						
İnternet siteleri						
Sosyal medya (Facebook, Twitter, Instagram gibi)						
Mesajlar						
E-postalar						
Çevrimiçi gruplar / forumlar						

E) Okumakla ilgili aşağıdaki ifadeler sizin için ne kadar doğru? (0 = hiç katılmıyorum, 6 = tamamen katılıyorum) (Likert scale)
- Sadece zorunluluktan okurum.
- Okumak en sevdiğim hobilerdendir.
- Diğer insanlarla kitaplar hakkında konuşmayı severim.
- Kitapları bitirmek bence zordur.
- Eğer bana hediye olarak kitap verilirse mutlu olurum.
- Bence okumak vakit kaybı.
- Kütüphaneye veya kitapçıya gitmekten keyif alırım.

The English version of the questionnaire can be accessed here: https://osf.io/uxr8e/files/osfstorage/654b91e0d45f5a066ae2fc07.

Appendix C

The code and the data are available at https://osf.io/uxr8e/?view_only=afb374d0351244c08fcd3891c50a349a, https://osf.io/uxr8e/?view_only=b15a50557f8d4ec584656f0b5ef0413

References

Acheson, D.J., Wells, J.B., MacDonald, M.C., 2008. New and updated tests of print exposure and reading abilities in college students. Behav. Res. Methods 40 (1), 278–289. https://doi.org/10.3758/BRM.40.1.278.

Aksan, Y., Aksan, M., Koltuksuz, A., Sezer, T., Mersinli, Ü., Demirhan, U.U., Yılmazer, H., Atasoy, G., Öz, S., Yıldız, İ., Kurtoğlu, Ö., 2012. Construction of the Turkish National corpus (TNC). In: Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12), pp. 3223–3227. http://www.lrec-conf.org/proceedings/lrec2012/pdf/991_Paper.pdf.

Bamyacı, E., 2016. Competing Structures in the Bilingual Mind: A Psycholinguistic Investigation of Optional Verb Number Agreement, first ed. 2016. Springer International Publishing: Imprint: Springer. https://doi.org/10.1007/978-3-319-22991-1.

Biber, Douglas., 1995. Dimensions of register variation: a cross-linguistic comparison. 1. publ. Cambridge University Press, Cambridge.

Biber, D., 2009. Dimensions of Register Variation. Cambridge University Press.

Bley-Vroman, R., 2009. The Evolving context of the Fundamental difference hypothesis. Stud. Sec. Lang. Acquis. 31 (2), 175–198. https://doi.org/10.1017/S0272263109090275.

Bybee, J., 2010. Language, Usage and Cognition. Cambridge University Press. https://doi.org/10.1017/CBO9780511750526.

Chomsky, N., 1965. Aspects of the Theory of Syntax. MIT press.

Chomsky, N., 1975. Reflections on Language. Temple Smith London.

Crain, S., Lillo-Martin, D.C., 1999. An Introduction to Linguistic Theory and Language Acquisition. Blackwell Publishers.

Crain, S., Thornton, R., Murasugi, K., 2009. Capturing the Evasive passive. Lang. Acquis. 16 (2), 123–133. https://doi.org/10.1080/10489220902769234.

Crawley, M., 2010. R a Language and Environment for Statistical Computing: Reference Index. R Foundation for Statistical Computing.

Cunningham, A.E., Stanovich, K.E., 1998. What reading does for the mind. Am. Educat. 22, 8–17.

Dąbrowska, E., 2008. The effects of frequency and neighbourhood density on adult speakers' productivity with Polish case inflections: an empirical test of usage-based approaches to morphology. J. Mem. Lang. 58 (4), 931–951. https://doi.org/10.1016/j.jml.2007.11.005.

Dąbrowska, E., 2012. Different speakers, different grammars: individual differences in native language attainment. Linguistic Approaches to Bilingualism 2 (3), 219–253. https://doi.org/10.1075/lab.2.3.01dab.

Dąbrowska, E., 2013. Functional constraints, usage, and mental grammars: a study of speakers' intuitions about questions with long-distance dependencies. Cognit. Ling. 24 (4), 633–665. https://doi.org/10.1515/cog-2013-0022.

Dąbrowska, E., 2014. Words that go together: Measuring individual differences in native speakers' knowledge of collocations. Ment. Lexicon 9 (3), 401–418. https://doi.org/10.1075/ml.9.3.02dab.

Dabrowska, E., 2015. Individual differences in grammatical knowledge. In: Dabrowska, E., Divjak, D. (Eds.), Handbook of Cognitive Linguistics. De Gruyter Mouton, pp. 650–668. https://doi.org/10.1515/9783110292022-033.

Dąbrowska, E., 2018. Experience, aptitude and individual differences in native language ultimate attainment. Cognition 178, 222–235. https://doi.org/10.1016/j.cognition.2018.05.018.

Dąbrowska, E., 2019. Experience, aptitude, and individual differences in linguistic attainment: a comparison of native and nonnative speakers: experience, aptitude, and individual differences. Lang. Learn. 69, 72–100. https://doi.org/10.1111/lang.12323.

Dąbrowska, E., 2020. Language as a phenomenon of the third kind. Cognit. Ling. 31 (2), 213–229. https://doi.org/10.1515/cog-2019-0029.

Dąbrowska, E., Pascual, E., Macías Gómez-Estern, B., 2022. Literacy improves the comprehension of object relatives. Cognition 224, 104958. https://doi.org/10.1016/j.cognition.2021.104958.

Dąbrowska, E., Pascual, E., Macías Gómez-Estern, B., Llompart, M., 2023. Literacy-related differences in morphological knowledge: a nonce-word study. Front. Psychol. 14.

Dionne, G., Dale, P.S., Boivin, M., Plomin, R., 2003. Genetic evidence for bidirectional effects of early lexical and grammatical development. Child Dev. 74 (2), 394–412. https://doi.org/10.1111/1467-8624.7402005.

Divjak, D., 2019. Frequency in Language: Memory, Attention and Learning. Cambridge University Press.

Divjak, Dagmar, Catherine Caldwell-Harris, L., 2015. Frequency and Entrenchment. In: Dabrowska, Ewa, Divjak, Dagmar (Eds.), Cognitive Linguistics. Foundation of Language, pp. 61–86. https://doi.org/10.1515/9783110626476-004.

Divjak, D., Milin, P., Medimorec, S., Borowski, M., 2022. Behavioral signatures of memory resources for language: looking beyond the lexicon/grammar divide. Cognit. Sci. 46 (11). https://doi.org/10.1111/cogs.13206.

Ellis, Nick C., 1996. Sequencing in SLA: phonological memory, chunking, and points of order. In: Studies in second language acquisition, 18. Cambridge University Press, pp. 91–126.

Feldman, J., Kerr, B., Streissguth, A.P., 1995. Correlational analyses of procedural and declarative learning performance. Intelligence 20 (1), 87–114. https://doi.org/10.1016/0160-2896(95)90007-1.

Georgiou, G.P., 2023. Nonverbal intelligence does matter for the perception of second language sounds. Psych 5 (1), 123–132. https://doi.org/10.3390/psych5010012.

Georgiou, G.P., Giannakou, A., 2024. Discrimination of second language vowel contrasts and the role of phonological short-term memory and nonverbal intelligence. J. Psycholinguist. Res. 53 (1), 9. https://doi.org/10.1007/s10936-024-10038-z.

Goldberg, A.E., 2006. Constructions at Work: The Nature of Generalization in Language. Oxford University Press.

Goldberg, A.E., 2019. Explain Me This: Creativity, Competition, and the Partial Productivity of Constructions. Princeton University Press.

Grömping, U., 2007. Estimators of relative importance in linear regression based on variance decomposition. Am. Statistician 61 (2), 139–147. https://doi.org/10.1198/000313007X188252.

Gruszka, A., Matthews, G., Szymura, B. (Eds.), 2010. Handbook of Individual Differences in Cognition: Attention, Memory, and Executive Control. Springer, New York. https://doi.org/10.1007/978-1-4419-1210-7.

Hartshorne, J.K., Tenenbaum, J.B., Pinker, S., 2018. A critical period for second language acquisition: evidence from 2/3 million English speakers. Cognition 177, 263–277.

Herschensohn, J., 2009. Fundamental and gradient differences in language development. Stud. Sec. Lang. Acquis. 31 (2), 259–289. https://doi.org/10.1017/S0272263109090305.

Hirsh-Pasek, K., Golinkoff, R.M., 1999. The Origins of Grammar: Evidence from Early Language Comprehension. MIT press.

Hoff, E., 2009. Language development at an early age: learning mechanisms and outcomes from birth to five years. Encyclopedia on Early Childhood Development, 1–5.

Huettig, F., Pickering, M.J., 2019. Literacy advantages beyond reading: prediction of spoken language. Trends Cognit. Sci. 23 (6), 464–475. https://doi.org/10.1016/j.tics.2019.03.008.
Huttenlocher, J., 1998. Language input and Language growth. Prev. Med. 27 (2), 195–199. https://doi.org/10.1006/pmed.1998.0301.
Kaplan, D., Berman, R., 2015. Developing linguistic flexibility across the school years. First Lang. 35 (1), 27–53. https://doi.org/10.1177/0142723714566335.
Kidd, E., Donnelly, S., 2020. Individual differences in first language acquisition. Annual Review of Linguistics 6 (1), 319–340. https://doi.org/10.1146/annurev-linguistics-011619-030326.
Kidd, E., Donnelly, S., Christiansen, M.H., 2018. Individual differences in language acquisition and processing. Trends Cognit. Sci. 22 (2), 154–169. https://doi.org/10.1016/j.tics.2017.11.006.
Larson-Hall, J., 2010. A Guide to Doing Statistics in Second Language Research Using SPSS and R, second ed. Routledge.
Lervåg, A., Dolean, D., Tincas, I., Melby-Lervåg, M., 2019. Socioeconomic background, nonverbal IQ and school absence affects the development of vocabulary and reading comprehension in children living in severe poverty. Dev. Sci. 22 (5). https://doi.org/10.1111/desc.12858.
Li, L., Wu, X., 2015. Effects of metalinguistic awareness on reading comprehension and the mediator role of reading fluency from grades 2 to 4. PLoS One 10 (3), e0114417. https://doi.org/10.1371/journal.pone.0114417.
Lidz, J., Williams, A., 2009. Constructions on holiday. Cognit. Ling. 20 (1). https://doi.org/10.1515/COGL.2009.011.
Llompart, M., Dąbrowska, E., 2020. Explicit but not implicit memory predicts ultimate attainment in the native language. Front. Psychol. 11, 569586.
Montag, J.L., MacDonald, M.C., 2015. Text exposure predicts spoken production of complex sentences in 8-and 12-year-old children and adults. J. Exp. Psychol. Gen. 144 (2), 447. https://doi.org/10.3389/fpsyg.2020.569586.
Montrul, S., 2009. Incomplete Acquisition in Bilingualism: Re-examining the Age Factor. Benjamins.
Nakipoğlu, M., Ketrez, N., 2006. Children's overregularizations and irregularizations of the Turkish aorist. Proc. Boston Univ. Conf. Lang. Dev. 30, 399–410.
Nakipoğlu, M., Michon, E., 2020. Abstraction vs. Analogy in the Turkish aorist. In: Gürer, A., Uygun Gökmen, D., Öztürk, B. (Eds.), Morphological Complexity within and across Boundaries: In Honour of Asli Göksel. John Benjamins, pp. 13–38.
Nakipoğlu, M., Üntak, A., 2008. A complete verb lexicon of Turkish based on morphemic analysis. Turk. Lang. 12 (2), 221–280.
Nippold, M.A., Hesketh, L.J., Duthie, J.K., Mansfield, T.C., 2005. Conversational versus expository discourse: a study of syntactic development in children, adolescents, and adults. J. Speech Lang. Hear. Res. 48 (5), 1048–1064. https://doi.org/10.1044/1092-4388(2005/073.
Nowak, M.A., Komarova, N.L., Niyogi, P., 2001. Evolution of universal grammar. Science 291 (5501), 114–118. https://doi.org/10.1126/science.291.5501.114.
Özel, S.A., Bektaş, Y., Yılmazer, H., 2016. Formulaicity in Turkish: evidence from the Turkish national corpus. Mersin University Journal of Linguistics & Literature/Mersin Üniversitesi Dil ve Edebiyat Dergisi 13 (2).
Paradis, J., 2009. Maturation: for better or for worse? Z. Sprachwiss. 28 (1), 41–45 https://doi.org/10.1515/ZFSW.2009.004.
Pinker, S., 1995. Language acquisition. In: Gleitman, L.R., Liberman, M. (Eds.), An Invitation to Cognitive Science, vol. 1: LanguageMIT Press, pp. 135–182.
Pinker, S., 1997. Words and rules in the human brain. Nature 387 (6633), 547–548. https://doi.org/10.1038/42347.
Plunkett, K., Marchman, V., 1993. From rote learning to system building: acquiring verb morphology in children and connectionist nets. Cognition 48 (1), 21–69. https://doi.org/10.1016/0010-0277(93)90057-3.
Prasada, S., Pinker, S., 1993. Generalisation of regular and irregular morphological patterns. Lang. Cognit. Process. 8 (1), 1–56. https://doi.org/10.1080/01690969308406948.
R Core Team, 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. https://www.r-project.org/.
Reber, A.S., Walkenfeld, F.F., Hernstadt, R., 1991. Implicit and explicit learning: individual differences and IQ. J. Exp. Psychol. Learn. Mem. Cognit. 17 (5), 888–896. https://doi.org/10.1037/0278-7393.17.5.888.
Ritchie, S.J., Tucker-Drob, E.M., 2018. How much does education improve intelligence? A meta-analysis. Psychol. Sci. 29 (8), 1358–1369. https://doi.org/10.1177/0956797618774253.
Roland, D., Dick, F., Elman, J.L., 2007. Frequency of basic English grammatical structures: a corpus analysis. J. Mem. Lang. 57 (3), 348–379. https://doi.org/10.1016/j.jml.2007.03.002.
Seidenberg, M.S., 1997. Language acquisition and use: learning and applying probabilistic constraints. Science 275 (5306), 1599–1603. https://doi.org/10.1126/science.275.5306.1599.
Skehan, P., Ducroquet, L., 1988. A Comparison of First and Foreign Language Ability. ESOL Department, London.
Stanovich, K.E., Cunningham, A.E., 1992. Studying the consequences of literacy within a literate society: the cognitive correlates of print exposure. Mem. Cognit. 20 (1), 51–68. https://doi.org/10.3758/BF03208254.
Stanovich, K.E., West, R.F., 1989. Exposure to print and orthographic processing. Read. Res. Q. 24 (4), 402. https://doi.org/10.2307/747605.
Street, J.A., 2017. This is the native speaker that the non-native speaker outperformed: individual, education-related differences in the processing and interpretation of Object Relative Clauses by native and non-native speakers of English. Lang. Sci. 59, 192–203. https://doi.org/10.1016/j.langsci.2016.10.004.
Street, J.A., 2020. More lexically-specific knowledge and individual differences in adult native speakers' processing of the English passive. Lang. Sci. 78, 101254. https://doi.org/10.1016/j.langsci.2019.101254.
Street, J.A., Dąbrowska, E., 2010. More individual differences in language attainment: how much do adult native speakers of English know about passives and quantifiers? Lingua 120 (8), 2080–2094.
Street, J.A., Dąbrowska, E., 2014. Lexically specific knowledge and individual differences in adult native speakers' processing of the English passive. Appl. Psycholinguist. 35 (1), 97–118. https://doi.org/10.1017/S0142716412000367.
Tomasello, M., 2003. Constructing a Language: A Usage-Based Theory of Language Acquisition. Harvard University Press.
Ullman, M.T., 2006. The declarative/procedural model and the shallow structure hypothesis. Appl. Psycholinguist. 27 (1), 97–105. https://doi.org/10.1017/S0142716406370038.
Warren-Leubecker, A., Carter, B.W., 1988. Reading and growth in metalinguistic awareness: relations to socioeconomic status and reading readiness skills. Child Dev. 59 (3), 728. https://doi.org/10.2307/1130572.
Yıldız, E., Nakipoğlu, M., 2012. The negative of the Turkish aorist and the challenges it poses during acquisition. Studia Uralo-Altaica 49, 581–591.