Journal of Research in Language & Translation
Issue No. 2 Vol. 5 (2025)

Verb–subject gender agreement in Standard Arabic: L2 acquisition by Arabic heritage speakers and French-speaking learners of Arabic

Kholoud A. Al-Thubaiti

Umm Al-Qura University,
Makkah, Saudi Arabia
kathubaiti@uqu.edu.sa
https://orcid.org/0000-0000

Received: 3/11/2024; Revised: 6/4/2025; Accepted: 21/4/2025

الملخص

تتناول هذه الدراسة المطابقة في الجندر (التذكير والتأنيث) بين الفعل والفاعل المفرد الغائب في سياق اكتساب اللغة العربية الفصحى، وذلك من خلال مقارنة الأداء اللغوي للناطقين بالعربية من أصول تراثية (ثنائيي اللغة في العربية العامية والفرنسية) ومتعلمي العربية الناطقين بالفرنسية. كما تبحث في التأثيرات الدلالية لفئة الاسم — وخاصة الأسماء البشرية والحيوانية، وكلاهما يعبّر عن الجندر الطبيعي — على اكتساب المطابقة في الجندر على الأفعال. وقد استُخدمت أداة الحُكم النحوي ضمن منهج كمي تجريبي لتقييم قدرة المتعلمين على التمييز بين حالات المطابقة الصحيحة وغير الصحيحة بين الفعل والفاعل. شملت عينة الدراسة 12 متحدثًا تراثيًا بالعربية، و15 متعلمًا فرنسياً للغة العربية الفصحى كلغة ثانية، و25 مشاركًا من الناطقين الأصليين بالعربية (أحادي اللغة) كمجموعة ضابطة. أظهرت النتائج تفوق المتحدثين التراثيين على متعلمي العربية من الناطقين بالفرنسية. وقد فُسّر هذا التفوق بالرجوع إلى خصائص اللغة الأم، إذ إن الفعل يتطلب المطابقة في الجندر مع الفاعل المفرد في كلٍّ من العربية العامية والفصحى، بخلاف اللغة الفرنسية التي لا تُلزم بهذه المطابقة. بالتالي، يستفيد المتحدثون التراثيون من النقل الإيجابي من العامية إلى الفصحى، فجاءت أحكامهم النحوية مماثلة لأحكام المجموعة الضابطة. كما أظهرت كلٌّ من المجموعتين — التراثية والضابطة — دقة أعلى في الجمل التي تحتوي على فاعل بشري مقارنةً بالجمل التي تحتوي على فاعل حيواني. في المقابل، لم يُظهر المتعلمون الفرنسيون للفصحى نفس التأثير المرتبط بفئة الاسم، بل أظهروا تأثيرًا عامًا للجندر تفاعل مع السلامة النحوية، حيث تم قبول الصيغة المؤنثة للفعل مع الفاعل المذكر والمؤنث على حد سواءٍ. تشير هذه النتائج إلى أن اكتساب الفصحى قد يشبه اكتساب لغة جديدة، إلا أن وجود خصائص لغوية مشتركة بين الفصحى والعامية يسهم في تسهيل عملية الاكتساب. كما تُظهر الدراسة أن الأسماء البشرية والحيوانية، رغم اشتراكها في التعبير عن الجندر الطبيعي، تختلف في تأثيرها الدلالي على المطابقة بين الفعل والفاعل، مما يشير إلى دور محتمل لعوامل إضافية، مثل التشابه الصرفي بين الصيغ المذكرة والمؤنثة، وهو ما يستدعي مزيدًا من البحث.

Verb–subject gender agreement in Standard Arabic: L2 acquisition by Arabic heritage speakers and French-speaking learners of Arabic

Kholoud A. Al-Thubaiti

Umm Al-Qura University,
Makkah, Saudi Arabia
kathubaiti@uqu.edu.sa
https://orcid.org/0000-0000

Received: 3/11/2024; Revised: 6/4/2025; Accepted: 21/4/2025

Abstract

This study investigates second language (L2) acquisition of gender agreement on verbs with third-person singular subjects in Standard Arabic by comparing the linguistic performance of Arabic heritage speakers (bilingual in Colloquial Arabic and French) and French-speaking L2 learners. It also examines the semantic effects of noun class —specifically, human and animal nouns, both of which encode natural gender—on the acquisition of gender agreement. A quantitative, experimental design using a grammaticality judgment task was employed to assess the participants' sensitivity to verb–subject gender agreement violations. The sample included 12 Arabic heritage speakers, 15 L2 learners, and 25 monolingual native speakers of Arabic as the control group. The results showed that heritage speakers outperformed L2 learners. Unlike in French, gender agreement on verbs with third-person singular subjects is similarly inflected in both Colloquial and Standard Arabic, allowing heritage speakers to benefit from positive transfer. Their judgments were comparable to those of native speakers. Both heritage and native speakers demonstrated significantly higher accuracy for sentences with human subjects compared to those with animal subjects. In contrast, L2 learners did not show a noun class effect, but instead exhibited a gender effect that interacted with grammaticality: the feminine verbal form was accepted in both masculine and feminine contexts. These findings suggest that acquiring a standard variety resembles acquiring a new language. However, when linguistic properties overlap between the standard and colloquial varieties, acquisition is facilitated. The study also reveals that, although both human and animal nouns encode natural gender, they differ in their semantic effects on verb–subject gender agreement. This indicates the role of additional factors, such as the morphological similarity between masculine and feminine forms, which merits further investigation.

Keywords: *Arabic heritage; verb–subject gender agreement; noun class; French, Standard Arabic*

Introduction

Arabic is one of the most widely studied second languages (L2)[1] globally. Its rich and complex[2] morphological system presents considerable challenges to both heritage speakers and L2 learners. This study examines Arabic heritage speakers (bilingual in Colloquial Arabic and French) and French-speaking L2 learners. Colloquial Arabic refers to regional spoken varieties that are distinct from Standard Arabic, the formal variety used in writing and official contexts. Arabic heritage speakers are bilingual individuals who have limited proficiency in their first language (L1) due to being raised in a foreign country (France), where exposure to the heritage language is typically confined to the home environment. In contrast, French-speaking L2 learners are typically students of Arabic as a foreign language. Both groups commonly make up the student population in classrooms where Standard Arabic is taught as an additional language or distinct variety. This study investigates whether acquiring a standard variety is comparable to acquiring a new language when there is a partial overlap with a colloquial variety.

Arabic varieties, both standard and colloquial, exhibit a complex morphological system of gender agreement across nominal and verbal domains. Studies on Arabic heritage speakers (Albirini et al., 2013) and L2 learners of Standard Arabic (Al-Hamad, 2003; Alamry, 2019; Alhawary, 2019) have consistently reported that gender agreement is a challenging grammatical feature to acquire. These studies have shown that gender agreement tends to be acquired earlier in verb–subject constructions than in noun–adjective constructions, and that the masculine form is acquired before the feminine form, likely due to its simpler morphology. The complexity of gender agreement increases when inanimate and plural nouns are used.

Given these challenges, this study examines L2 acquisition of verb–subject gender agreement in Standard Arabic. Gender marking on verbs is determined by both the morphological form (imperfective or perfective) and the subject's agreement features (person, number, and gender). For instance, with third-person singular subjects, gender is marked in the imperfective through prefixes (*j-* for masculine and *t-* for feminine), as in (1a–b), and in the perfective through suffixes (*-a* for masculine and *-at* for feminine), as in (2a–b):[3]

1. Imperfective

(a) ***ja-****ktub* *ʔatˁ-tˁa:lib*

3MS-write the-student.MS

'The (male) student writes.'

(b) ***ta-****ktub* *ʔatˁ-tˁa:lib-a*

3FS-write the-student-FS

---

[1] The term second language (L2) is used generically to refer to any language acquired after the first language (L1), regardless of the context of exposure (e.g., classroom instruction or naturalistic setting) (Ortega, 2013). In second language acquisition (SLA) research, L2 is typically preferred as a cover term unless the learning context itself is the focus of investigation, which is not the case in this study. Here, Arabic heritage speakers are acquiring Standard Arabic as a third language or variety, while French-speaking learners are acquiring it as a foreign language. Consistent with this approach, contemporary SLA does not distinguish between *learning* and *acquisition* (Ortega, 2013).

[2] The term 'complex' is used in its technical linguistic sense, referring to the structural richness and interaction of morphological features (e.g., gender, number, person), without implying difficulty or deficiency.

[3] Arabic examples are transliterated using IPA, except in cited examples, where the original transliteration from the source is retained. The following abbreviations are used: F = feminine, M = masculine, S = singular, P = plural, 3 = 3rd person, and AUX = auxiliary.

‘The (female) student writes.’

2. Perfective

(a) *katab-**a*** *ʔatˁ-tˁa:lib*

wrote-3MS the-student.MS

‘The (male) student wrote.’

(b) *katab-**at*** *ʔatˁ-tˁa:lib-a*

wrote-3FS the-student-FS

‘The (female) student wrote.’

When all combinations of person (first, second, third), number (singular, dual, plural), and gender (masculine and feminine) are considered, the verb–subject agreement paradigm in Standard Arabic includes 13 distinct affixal forms for the imperfective and another set for the perfective (see Alhawary, 2011; Ryding, 2005, for full paradigms). Mastery of these forms is essential, as gender agreement on verbs is obligatorily marked regardless of word order (subject–verb or verb–subject). The only cases in which the verb does not exhibit gender agreement occur with first-person singular and plural subjects, and second-person dual subjects (Aoun et al., 2010).

This study investigates whether Arabic heritage speakers and L2 learners differ in their sensitivity to violations of verb–subject gender agreement in Standard Arabic. Despite extensive research on Arabic morphology (Albirini et al., 2011; Albirini et al., 2013; Albirini, 2015; Alhawary, 2002, 2005, 2009; Benmamoun & Albirini, 2018), there is a lack of comparative studies that directly examine Arabic heritage speakers and L2 learners in their acquisition of grammatical gender in Standard Arabic. To address this gap, we conduct a comparative study of the acquisition of verb–subject gender agreement in Standard Arabic among two learner groups: Arabic heritage speakers and French-speaking L2 learners of Arabic. The heritage speakers are bilingual individuals who speak a local Arabic variety (Moroccan, Tunisian, or Algerian Arabic) as their home language and use French as their dominant language. The L2 learners are monolingual native speakers of French. Both groups study Standard Arabic as an additional language—or a distinct variety—at a language institute in Paris. A grammaticality judgment (GJ) task was used to test their sensitivity to agreement violations on imperfective verbs with third-person singular subjects, as shown earlier in (1a–b), while examining the semantic effects of two types of nouns (human and animal) that bear natural gender and the gender categories (masculine and feminine).

The remainder of this paper is organized as follows. First, we describe the gender agreement systems in Arabic (standard and colloquial) and French. Next, we review prior research on gender acquisition in Arabic with particular attention to nominal properties (gender and animacy) and L1 effects. We then present the research questions, describe the method, and report the results. Finally, we conclude with a discussion of the findings in light of the research questions.

The gender agreement systems in Arabic and French

Arabic varieties require gender agreement in both nominal and verbal domains, while French requires gender agreement in the nominal domain and, to a limited extent, in the verbal domain. This section provides an overview of the gender systems in Arabic and French.

Gender agreement in Arabic

Arabic varieties, including Standard Arabic, distinguish between two gender categories for nouns: masculine and feminine. The gender of a noun is determined by either biological sex or linguistic conventions. Inanimate nouns that refer to objects are assigned grammatical gender arbitrarily, while nouns referring to humans and animals are typically assigned a natural gender corresponding to the biological sex of the referent.

Markers of nominal gender are illustrated using Standard Arabic, the target variety in this study. Masculine nouns typically lack a gender affix, as shown in (3), while feminine nouns are marked by one of the feminine suffixes (*-a*, *-aa*, *-aʔ*), as shown in (4)–(6).

3. *maktab*
   office.MS (Alhawary, 2011, p. 38)
4. *maktab-**a***
   library-FS (Alhawary, 2011, p. 38)
5. *ħumm-**aa***
   fever-FS (Alhawary, 2011, p. 40)
6. *sˁaħr-**aʔ***
   desert-FS (Alhawary, 2011, p. 40)

Although most nouns in Standard Arabic follow this affixal pattern, some exceptions exist. A few masculine nouns may carry a feminine suffix, as in (7), while some feminine nouns may lack it, as in (8). In other cases, certain nouns can be treated as either masculine or feminine, as in (9).

7. *χali:f-**a***
   caliph.MS (Ryding, 2005, p. 120)
8. *ʃams*
   sun.FS (Ryding, 2005, p. 124)
9. *suug*
   market.F/MS (Ryding, 2005, p. 125)

Despite these exceptions, most feminine nouns in Standard Arabic are marked with (*-a*) in the singular (Alhawary, 2011; Ryding, 2005).

Most nouns referring to humans or animals have both masculine and feminine forms, derived from the same root. The masculine form, which refers to males, typically serves as the base form, while the feminine form is derived by adding the feminine suffix, as illustrated in (10) and (11).

10. (a) *malik* (b) *malik-**a***

king.MS queen-FS (Ryding, 2005, p. 124)

11. (a) *qitˤtˤ* (b) *qitˤtˤ*-***a***
cat.MS cat.FS (Ryding, 2005, p. 125)

As shown in (10)–(11), masculine and feminine nouns share the same roots (e.g., *mlk* and *qtˤtˤ*). However, not all feminine animal names are morphologically derived from a masculine base. In some cases, masculine and feminine forms are distinct lexical items, as in (12) and (13). The feminine form still carries the suffix (*-a*), while the corresponding masculine form lacks gender marking.

12. (a) *θawr* (b) *baqar*-***a***
bull.MS cow-FS

13. (a) *ʒamal* (b) *na:q*-***a***
camel.MS camel-FS (Alhawary, 2011, p. 39)

For inanimate object nouns, gender is assigned by convention, as shown in (14) and (15).

14. tˤa:wil-***a***
table-FS

15. kursiyy
chair-MS (Alhawary, 2011, p. 39)

In Standard Arabic, grammatical gender is marked on demonstratives, relative pronouns, adjectives, and verbs (Alhawary, 2011; Ryding, 2005). In the nominal domain, demonstratives and adjectives agree in gender with the head noun, as shown in (16) and (17). Feminine agreement is marked on the demonstrative by the suffix (*-hi*) and on the adjective by (*-a*).

16. ha:ða: *maktab* *kabi:r*
this.MS office.MS big.MS
'This is a big office.'

*17. ha:ði:**hi*** *maktab*-***a*** *kabi:r*-***a***
this.FS library.FS big.FS
'This is a big library.'

In the verbal domain, Arabic nouns typically require gender agreement with verbs at the Tense-Phrase (TP) level regardless of word order (Aoun et al., 2010; Benmamoun, 2000). Table 1 presents the full subject–verb agreement paradigm for the imperfective form in Standard Arabic and Moroccan Arabic, based on Aoun et al. (2010, pp. 74-75). Gender agreement on the imperfective verbal form is realized in either number or person affixes (Benmamoun, 2000). In Standard Arabic, verbal gender agreement appears in the prefix with third-person singular and dual subjects (*j-* for masculine, *t-* for feminine) and in the suffix for second- and third-person plural subjects (*-u:n* for masculine, *-na* for feminine) and second-person singular feminine subjects (*-i:n*). Moroccan Arabic diverges from Standard Arabic "with respect to the absence of the dual and gender distinctions in the plural" (Benmamoun, 2000, p. 23). However, it maintains gender agreement with third-person singular subjects (*j-* for masculine, *t-* for feminine), as highlighted in the shaded rows of Table 1.

Table 1

Agreement marking on Standard Arabic (SA) and Moroccan Arabic (MA) imperfective verbs.

Person	Number	Gender	SA	MA	Translation
1	Singular	-	ʔa-ktub	n-əktəb	I write.
	Plural	-	na-ktub	n-kətb-u	We write.
2	Singular	M	ta-ktub	t-əktəb	You write.
	Singular	F	ta-ktub-**i:n**	t-kətb-i	You write.
	Dual	-	ta-ktub-a:n	-	You both write.
	Plural	M	ta-ktub-**u:n**	t-kətb-u	You all write.
	Plural	F	ta-ktub-**na**	t-kətb-u	You all write.
3	Singular	M	**ja**-ktub 3MS-write	**j**-əktəb 3MS-write	He writes.
	Singular	F	**ta**-ktub 3FS-write	**t**-əktəb 3FS-write	She writes.
	Dual	M	**ja**-ktub-a:n	-	They both write.
	Dual	F	**ta**-ktub-a:n	-	They both write.
	Plural	M	ja-ktub-**u:n**	j-kətb-u	They all write.
	Plural	F	ja-ktub-**na**	j-kətb-u	They all write.

Note. For simplicity, inflectional endings were omitted. Affixes in bold mark gender, in addition to person or number. Shaded rows represent the forms tested in this study.

Gender agreement in French

Similar to Arabic, French categorizes nouns as either masculine or feminine, with gender determined either by biological sex or linguistic convention. According to Ayoun (2007), certain animate nouns are naturally assigned a gender that corresponds to the biological sex of the referent, as demonstrated in (18) and (19). However, most nouns, whether animate or inanimate, are arbitrarily assigned gender, as shown in (20) and (21). The cited examples are from Ayoun (2007, p. 131).

18. neveu
nephew.MS

19. nièce
niece.FS

20. vélo
bicycle.MS

21. coccinelle
ladybug.FS

Predicting the gender of French nouns based on form or meaning is generally not possible due to numerous exceptions to morphological or semantic patterns (Ayoun, 2007, 2018; Hawkins & Towell, 2015).

The assignment of gender to animal nouns in French differs from that in Standard Arabic. French typically uses a single grammatical gender to refer to both male and female animals. To specify the biological sex, the modifiers *mâle* ‘male’ or *femelle* ‘female’ are added. For example, the noun *abeille* ‘bee’ is feminine; to indicate a male bee, the word *mâle* ‘male’ is

appended, as shown in (22). Conversely, *singe* ‘monkey’ is masculine, and a female monkey is specified by adding *femelle* ‘female,’ as illustrated in (23). The cited examples are from Ayoun (2018, p. 119).

22. *une* *abeille* *mâle*
a.FS bee.FS male
‘a male bee’

23. *un* *singe* *femelle*
a.MS monkey.MS female
‘a female monkey’

French marks gender agreement on adjectives, determiners (such as articles and demonstratives), and possessives (Hawkins & Towell, 2015). As illustrated in (24) and (25), French uses distinct masculine and feminine forms for definite (*le*$_{(MS)}$ vs. *la*$_{(FS)}$) and indefinite articles (*un*$_{(MS)}$ vs. *une*$_{(FS)}$) with singular nouns. Adjective forms also vary depending on the gender of the noun or pronoun with which they agree. Masculine adjectives usually lack a gender suffix, while feminine adjectives often end in (*-e*) in both singular and plural forms (Hawkins & Towell, 2015).

24. *le / un* *petit* *garcon.*
the.MS / a.MS little.MS boy.MS
‘The/A little boy’

25. ***la / une*** *petit**e*** *fille.*
the.FS / a.FS little.FS girl.FS
‘The/A little girl’

Unlike Standard Arabic, French typically marks subject–verb agreement only for person and number and not for gender. This is illustrated in (26) and (27), adapted from Hawkins and Towell (2015, p. 222), with gender-paired subjects added. In both examples, the imperfect tense form of *vouloir* ‘to want’ agrees with the subject in person and number but does not change according to gender, whether singular (*elle* ‘she’ vs. *il* ‘he’) or plural (*garçons* ‘boys’ vs. *filles* ‘girls’).

26. ***elle / il*** *voul**ait*** *partir* *en* *vacances.*
She / he wanted to.3S go on holiday
‘She/He wanted to go on holiday.’

27. ***les garçons / les filles*** *voul**aient*** *tous* *participer* *au match.*
The.P boys / The.P girls wanted to.3P all take part in match
‘The boys/The girls all wanted to take part in the match.’

However, French shows limited cases of subject–verb gender agreement, specifically with past participles following *être* ‘to be’ in compound tenses and in passive constructions. This is illustrated in (28)–(31), drawn from Hawkins and Towell (2015, pp. 226-227), with the addition of gender-paired subjects. In these cases, feminine agreement is marked on the verb via the suffix (*-e*), while the masculine form lacks overt gender marking.

28. Suzanne *est* *sorti**e**.*
Suzanne.F is.AUX went out.FS

'Suzanne went out.'

29. *John* *est* *sorti.*
John.M is.AUX went out.MS

'John went out.'

30. la ***guerre*** *a été déclenché**e*** *par un malentendu.*
the.FS war.FS has been started.FS by a.MS misunderstanding.MS

'The war was started by a misunderstanding.'

31. *le combat a été déclenché par un malentendu.*
the.MS fight.MS has been started.MS by a.MS misunderstanding.MS

'The fight was started by a misunderstanding.'

Unlike Standard Arabic, French does not consistently require gender agreement on verbs. As a result, French-speaking L2 learners of Standard Arabic encounter a new grammatical pattern: third-person singular subjects that require gender agreement on the verb, a feature not marked in their L1 grammar. Conversely, Arabic heritage speakers acquiring Standard Arabic are already familiar with this type of verb–subject gender agreement from their colloquial variety, particularly with third-person singular subjects.

Previous studies

Gender and animacy in Arabic acquisition

Arabic-speaking children generally acquire gender morphology by the age of three, though some errors in gender agreement may persist until around age four. Initially, children tend to use the singular masculine form as the default (Aljenaie, 2009; Omar, 2007). In contrast, Arabic heritage speakers and L2 learners often struggle to correctly apply gender agreement (Al-Hamad, 2003; Alamry, 2019; Albirini et al., 2013; Alhawary, 2019).

Research consistently shows that gender agreement is acquired earlier in subject–verb constructions than in noun–adjective constructions, and that masculine forms are acquired before feminine forms (Albirini et al., 2013; Alhawary, 2002, 2005, 2009). In a study of Arabic heritage speakers who spoke Palestinian and Egyptian varieties, Albirini et al. (2013) found that participants were more accurate in subject–verb agreement (82.78%) than in noun–adjective agreement (63.92%). Notably, the singular masculine form was overused in 80.14% of the verb agreement errors and 74.62% of the adjective agreement errors. According to Albirini et al. (2013), this reflects the relative robustness of the singular masculine form in the grammar of Arabic heritage speakers.

This pattern of overgeneralizing a singular masculine morphology is not unique to Arabic. Similar tendencies have been observed among heritage speakers of other languages with rich agreement systems, including Spanish, Armenian, Lithuanian, Polish, and Russian (Montrul, 2008; Polinsky, 1995). As Albirini et al. (2013, p. 9) note, these forms "are less marked in the acquisition of inflectional morphology." Supporting this finding, Alhawary (2002, 2005, 2009, 2019) reported that L2 learners of Standard Arabic from diverse L1 backgrounds

(English, French, Japanese, Mandarin Chinese, and Russian) were consistently more accurate in producing masculine than feminine singular forms across both nominal and verbal agreement contexts. These learners also tended to overgeneralize the masculine form with third-person singular feminine subjects. Taken together, the oral production data suggest that the singular masculine form functions as a default in both nominal and verbal domains for Arabic heritage speakers and L2 learners.

Another important factor in the acquisition of gender agreement is the effect of noun animacy. Studies show that both Arabic heritage speakers and L2 learners perform better in gender-agreement tasks when the noun's gender is biologically determined (Alamry, 2019; Albirini et al., 2013; Alhawary, 2005, 2009, 2019). For instance, Albirini et al. (2013) found that Arabic heritage speakers had particular difficulty assigning correct verbal gender agreement when the subject was a non-human noun. The researchers attributed this to the absence of a natural gender in such nouns. However, the classification of non-human nouns is complex as it includes both inanimate objects and animate animals. Since animals have biological sex, the human/non-human classification presents challenges for this explanation.

Nouns with a biologically based gender are often grouped as 'animate,' encompassing both animal and human referents. Alamry (2019) examined this distinction by testing L2 learners' sensitivity to verb–subject gender agreement violations, using both self-paced reading and grammaticality judgment tasks. In his stimuli, human and animal nouns were grouped as animate, while non-living objects were classified as inanimate. His results revealed a clear animacy effect: learners, regardless of L1 background, were more accurate in their judgments and faster in their reading times when processing sentences with animate subjects compared to inanimate ones. Examples (32) and (33), adapted from Alamry's stimuli (2019, p. 121), illustrate this difference. The asterisk (*) indicates ungrammatical gender agreement due to a mismatched verb–subject gender.

32. *	sˁanaʕ-at	ʔan-naʒʒaru	ʔbwaaban	ʒamiilah	(animate)
	make.past-3FS	the-carpenter.MS	doors	beautiful	

'The carpenter made beautiful doors.'

33. *	ʔisˁtˁadam-a	ʔas-sayyar-tu	bir-rasˁiif	(inanimate)
	collide.past.3MS	the-car-FS	with-the-sidewalk	

'The car collided with the sidewalk.'

Alamry (2019) explains this effect by suggesting that L2 learners rely more heavily on the semantics of a noun, particularly its natural gender, than on its morphological form when determining grammatical gender. As such, animate nouns with semantic transparency are easier to process than inanimate nouns whose gender is assigned arbitrarily.

However, Almary's (2019) study did not separate animal and human nouns as independent categories, nor did it balance the number of each type of stimuli. For example, in the self-paced reading task, eight animate items were animal nouns, while thirty-two were human nouns (names and professions). Thus, the reported animacy effect may have been driven by the semantic salience of the human referents. At present, there is insufficient evidence to determine whether animal nouns, which also carry a natural gender, facilitate gender agreement processing in the same way as human nouns.

If semantic gender plays a key role in learners' judgments of noun gender, then both human and animal nouns should yield similar effects on the detection of verb–subject gender agreement violations. The current study investigates this prediction.

L1 transfer effects in the L2 acquisition of Arabic gender

In addition to nominal effects (gender and animacy), L1 transfer offers another crucial perspective on the acquisition of verb–subject gender agreement. The effect of L1 transfer (positive or negative) is more evident in the accurate use of gender agreement on adjectives than on verbs in Standard Arabic (Alhawary, 2005, 2009, 2019). Previous research has tested L2 learners from different L1 backgrounds for their knowledge of grammatical gender. These participants spoke L1s that either marked grammatical gender (e.g., French and Russian) or did not (e.g., English, Japanese, and Mandarin Chinese).

Alhawary (2009, p. 80) found that, after three years of instruction, French speakers achieved a significantly higher accuracy rate (93%) in feminine marking on adjectives than English (78%) and Japanese (69%) speakers at the same level of proficiency. Conversely, the same French speakers did not demonstrate significantly higher accuracy in feminine verbal marking (93%) compared to English (86%) or Japanese (79%) speakers (p. 78). Alhawary (2019) also found that the oral production of Russian and Mandarin Chinese speakers showed comparable rates of correct gender agreement on verbs across past and present tenses, indicating no L1 transfer effects in the verbal domain.

Furthermore, Alamry (2019) found no L1 effects on L2 learners' accuracy or processing of ungrammatical sentences involving mismatched subject–verb gender agreements. However, his findings showed that gender effects varied depending on the task type and L1 gender system. The participants in his study spoke various L1s (e.g., Filipino, Chinese, English, Tajik, Urdu, French), which he grouped into gender-marked and non-gender-marked L1 groups. In the GJ task, learners whose L1s marked gender performed more accurately with masculine nouns than with feminine nouns. In contrast, learners whose L1s did not mark gender performed similarly with both masculine and feminine nouns. Interestingly, the results from the self-paced reading task revealed the opposite pattern: learners with gender-marked L1s showed comparable reading times for masculine and feminine nouns, while those whose L1s lacked grammatical gender showed slower reading times for feminine nouns than for masculine nouns.

Al-Thubaiti (2024a, 2024b) reported similar findings, showing no L1 effect on L2 acquisition of gender agreement on imperfective verbs in Standard Arabic. One study (2024a) examined L1 speakers of English, Urdu, Filipino, and Romance languages, whereas the other (2024b) examined L1 speakers of French and English. Both studies consistently found no significant performance differences between the L1 groups, suggesting a lack of transfer effects from the L1 gender system.

Research on Arabic has also examined L1 transfer effects at the level of informal varieties, though this area remains underexplored. Two studies (Albirini, 2014; Benmamoun & Albirini, 2018) focused on how heritage speakers and L2 learners use negation in Standard Arabic. The rules governing sentential negation differ between Standard Arabic and colloquial varieties such as Egyptian and Palestinian Arabic. Based on oral assessments, these studies found no significant performance differences between heritage speakers and L2 learners in terms of the use of negation. Because negation differs between the source and target language variety, no positive transfer was expected, which may explain the lack of observed differences.

To date, no studies have directly compared Arabic heritage speakers and L2 learners to examine the effect of Colloquial Arabic on sensitivity to gender agreement violations on verbs in Standard Arabic. Existing research has focused either on Arabic heritage speakers (e.g., Albirini et al., 2013) or on L2 learners from different L1 backgrounds (e.g., Alamry, 2019; Alhawary, 2005, 2009, 2019), but without direct comparison between the two groups. In contrast, an extensive body of research has been conducted on Spanish heritage speakers. Studies by Montrul and colleagues (Montrul, 2008; Montrul et al., 2008; Montrul, 2010) found that heritage speakers often outperform L2 learners in core linguistic properties, particularly in oral rather than written performance.

Building on this body of research, the present study compares Arabic heritage speakers and French-speaking L2 learners in their sensitivity to verb–subject gender agreement violations in Standard Arabic, with a particular focus on the role of animacy and natural gender.

Research aims and questions

Previous literature on the L2 acquisition of gender agreement on verbs reveals two main gaps: (a) the absence of comparative studies that directly examine the linguistic performance of Arabic heritage speakers and L2 learners, and (b) a lack of empirical evidence regarding the semantic effects of noun type, particularly the distinction between human and animal nouns, both of which carry natural gender. Based on these gaps and the cross-linguistic differences between French and Standard Arabic, the present study investigates the following research questions:

RQ1. Do Arabic heritage speakers show an advantage over native French speakers in detecting gender agreement violations in verb–subject constructions with imperfective verbs?

RQ2. Do human and animal nouns—both of which carry natural gender—exhibit similar semantic effects on the detection of mismatched verb–subject gender agreement?

To address these questions, this study employed a GJ task designed to test participants' sensitivity to gender agreement violations in verb–subject constructions. The sample, materials, procedures, and data analysis are described in the Method section.

Method

Participants

The study included three groups: two experimental and one control. The experimental groups consisted of 12 Arabic heritage speakers (10 females, 2 males) and 15 French-speaking L2 learners (8 females, 7 males). The control group comprised 25 L1 Arabic speakers (21 females, 4 males). The average age at the time of testing was 24 years for the heritage speakers, 30 years for the L2 learners, and 27 years for the L1 controls.

The Arabic heritage speakers were second-generation immigrants in France. They spoke a Western Arabic variety (Moroccan, Algerian, or Tunisian) and used it to communicate with their parents, with French being the dominant language. They were all identified as Arabic–French bilinguals based on their language background. In contrast, the L2 learners were monolingual French speakers residing in France, none of whom had Arab ancestry. At the time of testing, both heritage speakers and L2 learners were enrolled in Standard Arabic classes at a language institute in Paris. Their motivation for studying Arabic was either career-related or religious. The learning experiences of both experimental groups are summarized in Table 2.

Table 2

Learning experience of Standard Arabic (SA) among experimental groups

	Arabic heritage speakers (*n* = 12)		**French-speaking L2 learners (*n* = 15)**	
	Mean (SD)	**Min-Max**	**Mean (SD)**	**Min-Max**
Age at testing	24.4 (5.0)	20–39	30.1 (6.1)	21–42
Age of Arabic learning (years)	14.7 (9.9)	5–37	22.1 (5.9)	8–34
Arabic program (weeks)	86 (32)	44–180	63 (27)	16–100
Self-rating of SA (general)	2.6 (0.6)	1.5–3.25	2.2 (0.5)	1–3
Self-rating of SA (speaking)	2.5 (0.6)	1–3	2.3 (0.6)	1–3
Self-rating of SA (listening)	2.9 (0.8)	2–4	2.1 (0.5)	1–3
Self-rating of SA (reading)	2.7 (0.6)	2–4	2.3 (0.6)	1–3
Self-rating of SA (writing)	2.3 (0.6)	1–3	2.2 (0.7)	1–3

Note. Self-ratings were on a 4-point ordinal scale (1= Novice, 2= Intermediate, 3= Advanced, 4= Superior).

The Arabic L1 control group was composed of monolingual speakers recruited from an Arabic-speaking population in Saudi Arabia. All participants held a bachelor's degree and had studied Standard Arabic continuously since childhood, as part of their formal schooling and university education.

Material

The GJ task was adapted from the version used in a previous study (Al-Thubaiti, 2024a). It included 16 sets of target items and 32 filler sentences. All sentences were six words in length to maintain uniformity across conditions. The target items were designed to assess participants' sensitivity to correct and incorrect verb–subject gender agreement.

Each target item was constructed as part of four minimal pairs using 16 masculine-feminine subject noun forms (eight human and eight animal nouns) and 16 imperfective verb forms (see the Appendix for the full list of nouns and verbs). The design manipulated three factors: grammaticality (grammatical vs. ungrammatical), gender (masculine vs. feminine), and noun class (human vs. animal). Each item appeared in four versions, yielding 64 target sentences (32 with human nouns and 32 with animal nouns). Two representative sets of target items are shown in Examples (34) and (35). An asterisk (*) marks verb–subject gender mismatches.

34. **Human nouns** (8 items x 4 versions)

(a) grammtical, masculine

ja[3MS]-ʃtari: az-zawʒ[MS] al-χudˁrawa:t atˁ-tˁa:ziʒa min as-su:q

buys the-husband the-vegetables the-fresh from the-market

يشتري الزوج الخضروات الطازجة من السوق.

(b) ungrammtical, masculine

*ta[3F3]-ʃtari: az-zawʒ[MS] al-χudˁrawa:t atˁ-tˁa:ziʒa min as-su:q

buys the-husband the-vegetables the-fresh from the-market

***تشتري الزوج الخضروات الطازجة من السوق.**

(c) grammatical, feminine

ta[3FS]-ʃtari: az-zawʒ-a[FS] al-χudˁrawa:t atˁ-tˁa:ziʒa min as-su:q

buys the-wife the-vegetables the-fresh from the-market

تشتري الزوجة الخضروات الطازجة من السوق.

(d) ungrammatical, feminine

*ja[3MS]-ʃtari: az-zawʒ-a[FS] al-χudˁrawa:t atˁ-tˁa:ziʒa min as-su:q

buys the-wife the-vegetables the-fresh from the-market

***يشتري الزوجة الخضروات الطازجة من السوق.**

35. **Animal nouns** (8 items x 4 versions)

(a) grammtical, masculine

ja[3MS]-hrub al-qitˁ[MS] min ʕala: su:r al-manzil

escapes the-cat from top-of fence the-house

يهرب القط من على سور المنزل.

(b) ungrammtical, masculine

*ta[3FS]-hrub al-qitˁ[MS] min ʕala: su:r al-manzil

escapes the-cat from top-of fence the-house

***تهرب القط من على سور المنزل.**

(c) grammatical, feminine

ta[3FS]-hrub al-qitˁ-a[FS] min ʕala: su:r al-manzil

escapes the-cat from top-of fence the-house

تهرب القطة من على سور المنزل.

(d) ungrammatical, feminine

*ja[3MS]-hrub al-qitˁ-a[FS] min ʕala: su:r al-manzil

escapes the-cat from top-of fence the-house

***يهرب القطة من على سور المنزل.**

All subject nouns were singular and took third-person verbal agreement. Feminine subjects consistently ended in the suffix (*-a*), whereas masculine subjects carried no overt gender marker. Verbs appeared with either feminine or masculine third-person singular prefixes (*t-*, *j-*), which had two phonological realizations (*ta-*, *tu-* and *ja-*, *ju-*), depending on the lexical root. All verbs in the stimuli followed the (*ta-*, *ja-*) pattern, except for two: *ju-ha:ʒim* 'he attacks' and *ju-sa:ʕid* 'he helps.'Lexical items were selected from the Arabic learners' dictionary for core vocabulary by Buckwalter and Parkinson (2011). All verbs and nouns were drawn from the 5,000 most frequently used Arabic lemmas.

The GJ materials were divided into four lists. Each list contained 16 target and 32 filler items. The target items appeared in one of the four versions (a, b, c, or d). The items were equally distributed across three categories: grammaticality (grammatical vs. ungrammatical), gender

(masculine vs. feminine), and noun class (human vs. animal). The order of the items within each list was pseudorandomized to prevent consecutive occurrences of the same condition.

The 32 filler items (16 grammatical and 16 ungrammatical) were unrelated to gender agreement. Because the gender mismatch in the target items occurred on the first word (the verb), the ungrammatical element in the filler sentences was placed from the second word onward, in order to balance the position of ungrammaticality across the full set of materials.

Procedure

The GJ task was administered using the Gorilla Experiment Builder (www.gorilla.sc). After providing initial consent, participants received a link via email that included the informed consent form, a language background questionnaire, and the GJ task. Participation was voluntary, and the participants were informed of their right to withdraw at any time. As compensation, they were offered one hour of online Arabic practice. All necessary measures were taken to ensure the confidentiality of the participant information, and no personal data were stored in the dataset. Each participant was assigned a unique identifier code to ensure complete anonymity.

The language background questionnaire included 33 items addressing Arabic language exposure and learning history. Participants were asked to rate their overall proficiency in Standard Arabic, as well as their proficiency by skill (speaking, listening, reading, and writing), using a 4-point ordinal scale (1= Novice, 2= Intermediate, 3= Advanced, 4= Superior). Prior to beginning the GJ task, the participants received detailed instructions and completed a brief practice session. The researcher was available to answer any questions, and participants were allowed to take breaks during the task as needed.

The participants were instructed to judge the grammatical accuracy of each sentence as quickly and accurately as possible. They responded using three clickable options: صحيحة (correct), غير صحيحة (incorrect), and غير متأكد (not sure), which appeared below each sentence on the screen. Sentences were presented one at a time, following a fixation cross displayed for 250 ms. Each sentence remained on the screen until the participant responded. No time pressure was imposed. Upon completing the task, participants were shown the total number of correct responses on a final ‘Thank you’ screen.

Data Analysis

Reliability analyses of the GJ items were conducted using the *psych* package (Version 2.5.3) in R (Revelle, 2025). Internal consistency was assessed using Cronbach’s alpha, which was calculated separately for the grammatical and ungrammatical conditions. One version of each item was randomly selected from each participant to ensure independent responses. The ungrammatical condition showed good internal consistency ($\alpha = 0.76$), whereas the grammatical condition showed moderate reliability ($\alpha = 0.52$). These findings indicate that participants were more consistent in detecting mismatched verb–subject gender agreement than matched constructions. This pattern is commonly observed in GJ tasks.

In subsequent mixed-effects modeling, all items were retained to preserve the experimental structure, including the paired design of grammatical and ungrammatical items. Although some grammatical items contributed less to internal consistency, random effects modeling accounted for this variability at the item level.

GJ responses (correct, incorrect, and not sure) were coded as binary accuracy judgments (correct =1, incorrect =0). The proportion of 'not sure' responses was low across the dataset (5.65%), and even lower in L1 groups (Arabic controls = 1.32%, heritage speakers = 1.20%, L2 learners = 3.12%). Accordingly, 'not sure' responses were excluded from all analyses.

The analysis proceeded in two stages. First, between-group analyses were conducted on accuracy judgments, comparing performance by the L1 group and grammaticality (grammatical vs. ungrammatical). Second, separate within-group analyses were carried out for each L1 group, examining how the nouns' semantic properties—gender (masculine vs. feminine) and noun class (human vs. animal)—interacted with grammaticality.

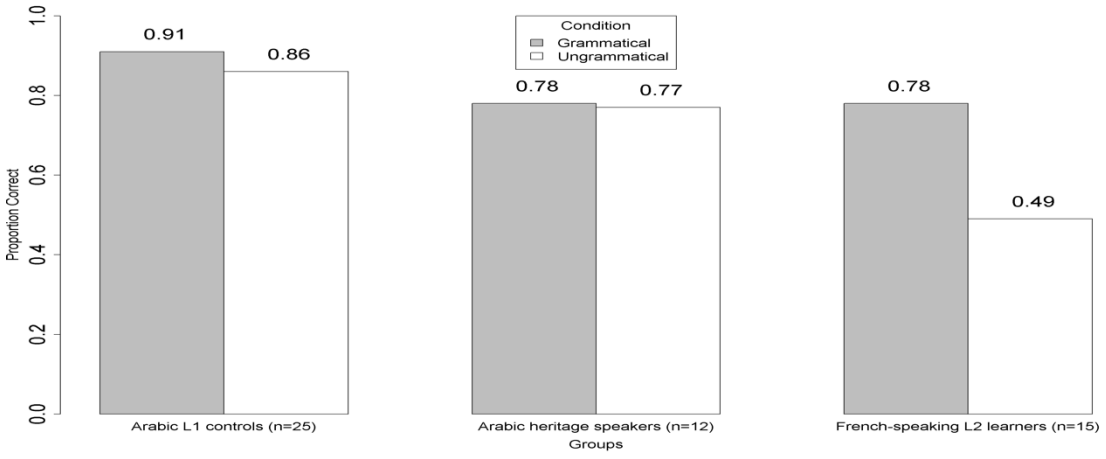
Results

Between-group accuracy judgments

As shown in Figure 1, descriptive statistics revealed differences in accuracy across the L1 groups and grammaticality conditions. To statistically assess these differences, a binomial generalized linear mixed-effects model (GLMM) was fitted using the *glmer()* function from the lme4 package (Version 1.1-31) (Bates et al., 2015). Post hoc pairwise comparisons were conducted using the function *emmeans()* from the emmeans package (Version 1.10.6) (Lenth, 2024). The model examined the proportion of correct judgments of grammatical and ungrammatical sentences across three groups: Arabic L1 controls (n = 25), Arabic heritage speakers (n = 12), and French-speaking L2 learners (n = 15).

Figure 1

Group accuracy judgments for grammatical and ungrammatical conditions.



The GLMM model included two fixed effects for grammaticality (grammatical vs. ungrammatical) and L1 group, as well as their interaction. Fixed effects were sum-coded to allow for an ANOVA-style interpretation. The L1 group factor was analyzed using two planned contrasts: Contrast 1 (C1) compared Arabic controls and Arabic heritage speakers, and Contrast 2 (C2) compared French-speaking L2 learners and Arabic heritage speakers. Random intercepts were included for participants (N = 52) and items (k = 16), along with a random slope for grammaticality by participant. The model output is presented in Table 3.

Table 3

GLMM: Accuracy in judging verb–subject gender agreement across L1 groups

Fixed effects	**Estimate**	**SE**	***z***	**OR**	***p***
(Intercept)	1.68	0.26	6.39	5.36	<0.001
L1 group.C1 (Controls vs. Heritage)	-0.96	0.24	-4.01	0.38	<0.001
L1 group.C2 (L2 vs. Heritage)	-0.88	0.25	-3.53	0.42	<0.001
Grammaticality (g vs. ug)	0.30	0.15	1.99	1.36	0.046
Grammaticality × L1 group.C1	0.11	0.19	0.58	1.12	0.561
Grammaticality × L1 group.C2	0.49	0.20	2.50	1.63	0.013
Random effects	**Variance**	**SD**			
Participant	0.90	0.95			
Grammaticality	0.36	0.60			
Item	0.51	0.71			

Note. Model formula: Correct ~ L1 * Grammaticality + (1 + Grammaticality | Participant) + (1 | Item). Number of observations = 832; Participants = 52; Items = 16. OR= odds ratio.

The model revealed a significant main effect of grammaticality, indicating that participants were more accurate in the grammatical than in the ungrammatical condition. Significant between-group effects showed that Arabic heritage speakers were less accurate than Arabic controls, but more accurate than French-speaking L2 learners. Crucially, there was a significant interaction between grammaticality and L1 group for the L2 vs. heritage contrast (C2), suggesting that French-speaking L2 learners were particularly affected by ungrammatical conditions. In contrast, no significant interaction was found between heritage speakers and Arabic controls (C1), suggesting similar accuracy patterns across the conditions (as illustrated in Figure 1). Pairwise comparisons using the *emmeans()* function support these findings. A significant difference between grammatical and ungrammatical conditions was found only for French-speaking L2 learners (*estimate* = 1.59, *SE* = 0.48, *z*–ratio = 3.32, *p* = 0.0009), but not for Arabic controls (*estimate* = 0.39, *SE* = 0.49, *z*–ratio = 0.80, *p* = 0.424) or Arabic heritage speakers (*estimate* = -0.146, *SE* = 0.54, *z*–ratio = -0.27, *p* = 0.788).

Within-group accuracy judgments

Separate GLMMs were fitted for each L1 group to examine the effects of noun class (human vs. animal) and gender (masculine vs. feminine) on grammaticality judgment accuracy. All models included three fixed effects (grammaticality, gender, and noun class) using contrast coding, along with all possible interactions among these factors. The models also incorporated two random intercepts (by participant and item) and a random slope for grammaticality by participant. The results for each group (GLMM-1, GLMM-2, and GLMM-3) are presented in Tables (4)–(6), respectively.

Arabic L1 controls

As shown in Table 4, the GLMM-1 model for Arabic L1 controls revealed no significant effects of grammaticality or gender but showed a significant main effect of noun class. Participants gave statistically comparable judgments for grammatical and ungrammatical sentences across both the masculine and feminine conditions. However, they were significantly more accurate in judging sentences with human subjects compared to animal subjects. The odds of making a correct judgment for human nouns were 0.56 times higher than for animal nouns,

suggesting that native Arabic speakers may be less familiar with certain animal names, which may have impacted their performance.

Table 4

GLMM-1: Arabic L1 controls' accuracy in judging verb–subject gender agreement

Fixed effects	**Estimate**	**SE**	***z***	**OR**	***p***
(Intercept)	2.95	0.48	6.16	19.08	<0.001
Grammaticality (g vs.ug)	0.62	0.34	1.81	1.87	0.070
Gender (masc. vs. fem.)	0.04	0.20	0.20	1.04	0.845
Noun Class (human vs. animal)	-0.57	0.29	-1.99	0.56	0.047
Grammaticality × Gender	0.10	0.20	0.50	1.11	0.620
Grammaticality × Noun Class	-0.12	0.20	-0.59	0.89	0.557
Gender × Noun Class	0.03	0.20	0.14	1.03	0.890
Grammaticality × Gender × Noun Class	-0.16	0.20	-0.81	0.85	0.421
Random effects	**Variance**	**SD**			
Participant	1.27	1.13			
Grammaticality	0.43	0.65			
Item	0.61	0.78			

Note. Model formula: Correct ~ Grammaticality * Gender * NounClass + (1 + Grammaticality | Participant) + (1 | Item). Number of observations = 400; Participants = 25; Items = 16. OR= odds ratio.

Arabic heritage speakers

As shown in Table 5, the GLMM-2 model for Arabic heritage speakers revealed a significant main effect of noun class, and a significant three-way interaction among grammaticality, gender, and noun class. While the main effect of noun class indicated reduced accuracy for animal nouns, this effect was qualified by the interaction, which showed that accuracy patterns varied across gender and grammatical conditions (see Figure 2). To unpack this interaction, pairwise comparisons were performed using the *emmeans()* function. For grammatical sentences, heritage speakers were more accurate with human subjects than animal subjects across both gender categories. The accuracy gap between human and animal subjects was larger in feminine conditions than masculine ones; however, this difference did not reach significance in either case (feminine: *estimate* = -1.53, *SE* = 0.87, *z*–ratio = -1.75, *p* = 0.08; masculine: *estimate* = -0.48, *SE* = 0.88, *z*–ratio = -0.55, *p* = 0.58). For ungrammatical sentences, participants showed significantly lower accuracy in the masculine condition for animal subjects compared to human subjects (*estimate* = -2.62, *SE* = 1.08, *z*–ratio = -2.42, *p* = 0.016). In contrast, accuracy in the feminine condition did not differ between animal and human subjects (*estimate* = 0.03, *SE* = 0.90, *z*–ratio = 0.04, *p* = 0.971).

These findings suggest that Arabic heritage speakers experience specific difficulties with masculine animal nouns, particularly in identifying ungrammatical sentences. They were more likely to incorrectly accept feminine verbal forms with masculine animal names (42% of the time) than masculine verbal forms with feminine animal names (21% of the time).

Table 5

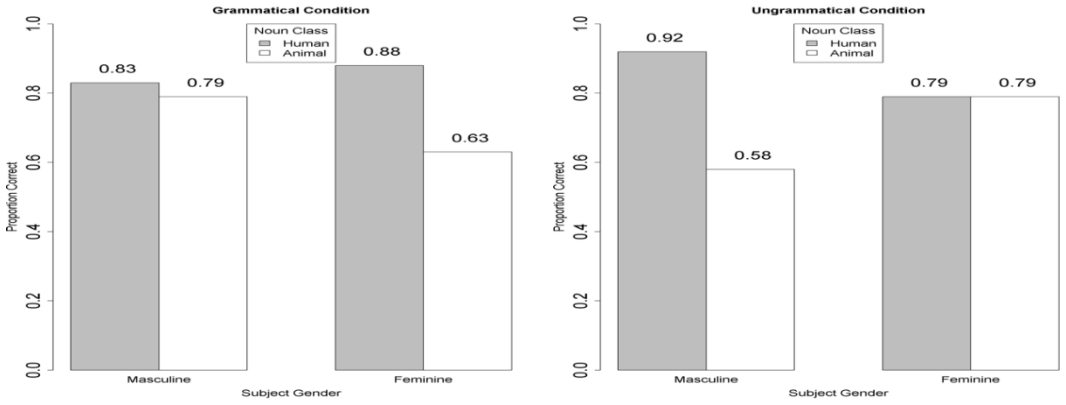
GLMM-2: Arabic heritage speakers' accuracy in judging verb–subject gender agreement

Fixed effects	**Estimate**	**SE**	***z***	**OR**	***p***
(Intercept)	1.63	0.37	4.37	5.13	<0.001
Grammaticality (g vs.ug)	-0.21	0.26	-0.80	0.81	0.426
Gender (masc. vs. fem.)	-0.10	0.22	-0.45	0.91	0.656
Noun Class (human vs. animal)	-0.57	0.28	-2.05	0.56	0.041
Grammaticality × Gender	-0.04	0.21	-0.20	0.96	0.843
Grammaticality × Noun Class	0.07	0.22	0.34	1.07	0.738
Gender × Noun Class	0.20	0.22	0.92	1.22	0.357
Grammaticality × Gender × Noun Class	-0.46	0.22	-2.14	0.63	0.033
Random effects	**Variance**	**SD**			
Participant	0.51	0.72			
Grammaticality	0.15	0.38			
Item	0.49	0.70			

Note. Model formula: Correct ~ Grammaticality * Gender * NounClass + (1 + Grammaticality | Participant) + (1 | Item). Number of observations =192; Participants=12; Items=16. OR= odds ratio.

Figure 2

Three-way interaction of grammaticality, gender, noun class in Arabic heritage speakers



French-speaking L2 learners

Table 6 shows that the GLMM-3 model for French-speaking L2 learners revealed a significant two-way interaction between grammaticality and gender, with no significant main or interaction effects for noun class. Pairwise comparisons were conducted using *emmeans()* to explore this interaction. The results showed that in the masculine condition, accuracy was significantly higher for grammatical than for ungrammatical sentences (*estimate* = 2.564, *SE* = 0.86, *z*–ratio = 2.99, *p* = 0.003). However, in the feminine condition, there was no significant difference between grammatical and ungrammatical sentences (*estimate* = 0.086, *SE* = 0.83, *z*–ratio = 0.103, *p* = 0.918). Additionally, in the ungrammatical condition, L2 learners showed higher accuracy when masculine verbal forms were incorrectly used with feminine subjects (0.58) compared to the reverse pattern (feminine verbal forms with masculine subjects (0.4) (see Figure 3). This difference was marginally significant (*estimate* = 1.08, *SE* = 0.59, *z*–ratio =

1.834, $p = 0.067$). These findings suggest that L2 learners were more likely to incorrectly accept feminine verbal forms with masculine subjects (regardless of noun class), indicating a possible asymmetry in sensitivity to gender agreement violations.

Table 6

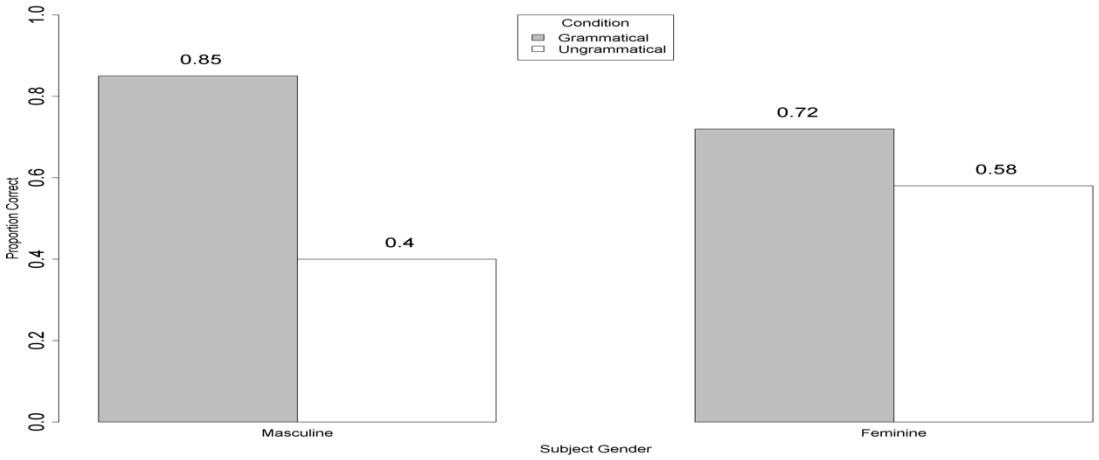
GLMM-3: French-speaking L2 learners' accuracy in judging verb–subject gender agreement

Fixed effects	**Estimate**	**SE**	***z***	**OR**	***p***
(Intercept)	0.92	0.50	1.83	2.50	0.068
Grammaticality (g vs.ug)	0.66	0.37	1.78	1.94	0.076
Gender (masc. vs. fem.)	-0.08	0.22	-0.37	0.92	0.713
Noun Class (human vs. animal)	-0.34	0.32	-1.06	0.71	0.290
Grammaticality × Gender	-0.62	0.20	-3.10	0.54	0.002
Grammaticality × Noun Class	-0.12	0.20	-0.60	0.88	0.547
Gender × Noun Class	-0.11	0.21	-0.51	0.90	0.611
Grammaticality × Gender × Noun Class	0.01	0.19	0.05	1.01	0.957
Random effects	**Variance**	**SD**			
Participant	2.01	1.42			
Grammaticality	1.24	1.12			
Item	1.04	1.02			

Note. Model formula: Correct ~ Grammaticality * Gender * NounClass + (1+ Grammaticality | Participant) + (1 | Item). Number of observations = 240; Participants =15; Items =16. OR= odds ratio.

Figure 3

Two-way interaction between grammaticality and gender for French-speaking L2 learners.



Overall summary

The results revealed distinct patterns among the three groups. Arabic L1 controls showed consistently high accuracy in judging grammaticality, regardless of the subject's gender. However, they showed a noun class effect, with higher accuracy for human nouns. In contrast, Arabic heritage speakers exhibited lower overall accuracy, particularly in the ungrammatical condition, where performance was affected by gender and noun class. Accuracy was notably reduced when dealing with masculine animal subjects. French-speaking L2 learners, by comparison, did not exhibit a noun-class effect. Instead, their performance was influenced by gender, which interacted with grammaticality. In the ungrammatical condition, they showed reduced accuracy when feminine verbal forms were used with masculine subjects, suggesting difficulty in detecting certain gender agreement violations.

Discussion

The findings of this study contribute to the ongoing debate on whether acquiring a standard variety is comparable to acquiring a foreign language (addressing RQ1). The study examined sensitivity to grammatical violations of verb–subject gender agreement in two groups (French-speaking L2 learners and Arabic heritage speakers), both of whom were acquiring Standard Arabic as an additional language or variety. In the GJ task, Arabic heritage speakers demonstrated greater sensitivity to gender agreement violations than L2 learners. Similar to Arabic L1 controls, heritage speakers accurately rejected sentences with incorrect gender agreement and accepted sentences with correct agreement.

The superior performance of heritage speakers can be attributed to the structural similarities between their colloquial Arabic dialects (Moroccan, Tunisian, or Algerian Arabic) and Standard Arabic, particularly in the marking of third-person singular verb–subject gender agreement. This alignment likely facilitated positive L1 transfer. In contrast, French (the L1 of L2 learners) does not mark verbal gender agreement in the same way, resulting in negative transfer effects. For French-speaking L2 learners, gender agreement in the verbal domain is a novel grammatical feature that requires substantial exposure and practice to be fully acquired. According to their self-ratings, these learners are, on average, at the novice-to-intermediate proficiency level. In future research, proficiency should be measured objectively to better assess its influence on grammatical sensitivity.

These findings align with those of Montrul (2008, 2010) and Montrul et al. (2008), who found that Spanish heritage speakers outperformed L2 learners on grammatical gender tasks, particularly in oral rather than written performance. Montrul and colleagues attributed their findings to task effects, early language experience, and the nature of input. They argued that heritage speakers receive naturalistic input in the home from an early age, giving them an advantage over L2 learners, who are typically exposed to classroom-based input during adolescence or adulthood. This explanation is also applicable to Arabic heritage speakers, who acquire their native colloquial variety in early childhood at home. However, unlike Spanish heritage speakers, Arabic heritage speakers in this study outperformed L2 learners in a written task. The GJ stimuli contained short, high-frequency vocabulary and simple sentences common to both colloquial and standard varieties, which likely facilitated positive L1 transfer. Their performance on the GJ task suggests a robust metalinguistic awareness of a core grammatical feature. At the same time, these results diverge from those reported by Albirini (2014) and Benmamoun and Albirini (2018), who examined knowledge of sentential negation in Standard Arabic, a linguistic property that differs more substantially between dialects and the standard variety. Their results, based on oral assessments, showed no advantage for heritage speakers who performed at the same level as the L2 learners. Together, these contrasting results suggest that acquiring a standard variety resembles the process of acquiring an additional language, with both positive and negative effects of L1 transfer depending on the specific linguistic domain under investigation.

Turning to RQ2, this study also examined the effects of nominal properties, specifically gender and noun class, on learners' sensitivity to correct and incorrect verb–subject gender agreement. Unlike Arabic controls, who exhibited no gender effects, Arabic heritage speakers showed effects only with animal nouns, whereas L2 learners showed effects with both human and animal nouns. Notably, ungrammatical sentences with masculine subjects were judged less accurately than those with feminine subjects, indicating a tendency to overgeneralize the

feminine verbal form. This finding contrasts with previous research by Alhawary (2005, 2009, 2019) and Albirini et al. (2013), who identified the masculine form as the default in Arabic. It is important to note, however, that these studies focused on oral production, where the omission of feminine inflection is more likely due to real-time processing constraints.

The noun class results revealed that, although both human and animal nouns encode natural gender, they have different effects on the performance of Arabic heritage speakers and L1 controls, but not on that of L2 learners. The L1 controls and heritage speakers performed better on human subjects than on animal subjects, whereas L2 learners performed similarly across both noun categories. It is noteworthy that a noun class effect emerged among Arabic speakers (both controls and heritage), but not among L2 learners. One possible explanation for this group difference is that L2 learners are less proficient than Arabic speakers. As they are still in the process of acquiring gender agreement across all noun categories, L2 learners exhibit a general gender effect without sensitivity to noun-class distinctions. In contrast, L1 controls and heritage speakers may face challenges in identifying the correct gender for certain animal names (e.g., *female camel* and *female lion*), which are less frequently encountered. The divergent findings for human and animal nouns indicate that the semantic effects of natural gender alone are not sufficient to guide learners' judgments of gender agreement. If semantic gender were the primary factor in gender assignment, both human and animal nouns would be expected to yield similar effects on sensitivity to verb–subject gender agreement violations. This suggests that other factors, such as lexical familiarity, interact with the animacy dimension of the noun-class effects. Although the tested nouns are among the 5000 most frequent Arabic lemmas, it is possible that some, especially animal names, may occur less frequently in the input available to L1 controls and heritage speakers. According to the SAMER-leveled readability lexicon for Standard Arabic (Al Khalil et al., 2020), the masculine and feminine nouns used in the GJ stimuli are generally appropriate for first-grade Arabic readers. However, only two feminine animal names (*female camel* and *female lion*) were rated as suitable for readers at the fourth- and fifth-grade levels, suggesting reduced early exposure and lexical familiarity.

Another contributing factor could be the gender form of human and animal nouns. Most animal noun pairs (e.g., na:q-**a** 'female camel' vs. ʒamal 'male camel') consist of distinct lexical items, whereas most human nouns share a common root (e.g., ʒadd-**a** 'grandma' vs. ʒadd 'grandpa'). It could be argued that nouns with shared roots are generally easier to acquire because they require the addition of gender inflection to the same lexical base. In contrast, nouns with different lexical items are more complex, as learners must acquire both forms separately. In such cases, masculine and feminine forms are stored separately in the lexicon, with the feminine typically marked by the suffix (*-a*). This explanation, however, requires empirical validation in future studies. A further general factor that may influence these results is the number of irregularities and exceptions in the gender-marking system of Standard Arabic, which can slow acquisition. The specific factors contributing to this distinction between human and animal nouns still require further investigation, using stimuli designed to systematically measure the effects of root-sharing, familiarity, and morphological regularity. This study affirms that human and animal nouns should be tested as distinct constructs, laying the groundwork for future studies.

Conclusion and future research

In conclusion, this study provides new insights into the acquisition of grammatical gender in the verbal domain of Standard Arabic by Arabic heritage speakers and French-

speaking L2 learners. The results demonstrated superior linguistic performance by heritage speakers compared to L2 learners. Unlike French, verbal gender agreement with singular subjects is similarly inflected in both Colloquial and Standard Arabic. However, before it can be claimed that acquiring the standard variety is similar to acquiring an additional language, empirical evidence is needed from heritage speakers' performance on linguistic properties that differ substantially between the two varieties. Future research should investigate verbal gender agreement with plural subjects, as this is an area where notable differences exist between Colloquial and Standard Arabic.

The findings also raise important questions regarding the animacy effects of noun class (human and animal) and the morphological relatedness of gendered nominal forms (i.e., distinct forms vs. shared roots). The existing literature on grammatical gender acquisition in the verbal domain remains limited and would benefit from further research that systematically examines the nominal properties of the subject, particularly in terms of animacy, number, and gender features.

Acknowledgments

I would like to express my gratitude to my colleagues, Professor Uri Horesh and Professor Bruno Herin, whose support was instrumental in facilitating my connection with students at the Paris Language Institute (INALCO). I am also grateful to the editor and two anonymous reviewers for their valuable comments.

Funding

The authors did not receive funding from any organization for this study.

Bio

Kholoud A. Al-Thubaiti is an Associate Professor of Applied Linguistics in the Department of English, College of Social Sciences, Umm Al-Qura University, Saudi Arabia. Her research focuses on the acquisition and processing of morphosyntactic features in second language development, with particular interest in the influence of the first language and age of acquisition. Her recent work examines the processing of English articles and the acquisition of grammatical gender in Arabic as a second language. She has published in peer-reviewed journals indexed in the Web of Science and serves on the editorial board of the journal *Arabic Linguistics* (John Benjamins).

References

Alamry, A. (2019). *Grammatical gender processing in Standard Arabic as a first and a second language* [Doctoral dissertation, University of Ottawa].

Albirini, A., Benmamoun, E., & Saadah, E. (2011). Grammatical features of Egyptian and Palestinian Arabic Heritage speakers' oral production. *Studies in Second Language Acquisition, 33*(2), 273-303. https://doi.org/10.1017/S0272263110000768

Albirini, A., Benmamoun, E., & Chakrani, B. (2013). Gender and number agreement in the oral production of Arabic Heritage speakers. *Bilingualism: Language and Cognition, 16*(1), 1-18. https://doi.org/10.1017/S1366728912000132

Albirini, A. (2014). The role of the colloquial varieties in the acquisition of the Standard variety: The case of Arabic heritage speakers. *Foreign Language Annals, 47*(3), 447-463. https://doi.org/10.1111/flan.12087

Albirini, A. (2015). Factors affecting the acquisition of plural morphology in Jordanian Arabic. *Journal of Child Language, 42*(4), 734-762. https://doi.org/10.1017/S0305000914000270

Al-Hamad, M. M. (2003). *Morphological and syntactic properties in the acquisition of Arabic as a second language: Implications for the theory of SLA and for language teaching* [Doctoral dissertation, University of Essex].

Alhawary, M. (2002). Role of L1 transfer in L2 acquisition of inflectional morphology. In D. B. Parkinson & E. Benmamoun (Eds.), *Perspectives on Arabic Linguistics XIII-XIV* (pp. 219-248). John Benjamins.

Alhawary, M. (2005). L2 acquisition of Arabic morphosyntactic features: Temporary or permanent impairment? In M. Alhawary & E. Benmamoun (Eds.), *Perspectives on Arabic Linguistics XVII–XVIII* (pp. 273–311). John Benjamins.

Alhawary, M. (2009). *Arabic second language acquisition of morphosyntax*. Yale University Press.

Alhawary, M. (2011). *Modern Standard Arabic: A learner's guide*. Wiley-Blackwell.

Alhawary, M. (2019). *Arabic second language learning and effects of input, transfer, and typology*. Georgetown University Press.

Aljenaie, K. (2009). *The emergence of tense and agreement in Kuwaiti Arabic children* [Doctoral dissertation, University of Reading].

Al Khalil, M., Habash, N., & Jiang, Z. (2020). A large-scale leveled readability lexicon for Standard Arabic. In N. Calzolari, F. Béchet, P. Blache, K. Choukri, C. Cieri, T. Declerck, S. Goggi, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, & S. Piperidis (Eds.), *Proceedings of the 12th Language Resources and Evaluation Conference* (pp. 3053-3062). European Language Resources Association. https://aclanthology.org/2020.lrec-1.373

Al-Thubaiti, K. A. (2024a). Verbal gender agreement in L2 Standard Arabic: Proficiency or L1 effects? *The Arabic Linguistics Journal, 18*, 117–137.

Al-Thubaiti, K. A. (2024b). Verbal gender agreement in L2 Standard Arabic by L1 English and L1 French speakers. In H. AlThagafi & J. Ray (Eds.), *Proceedings of the 48th Annual Boston University Conference on Language Development* (pp. 31–44). Cascadilla Press.

Aoun, J., Benmamoun, E., & Choueiri, L. (2010). *The syntax of Arabic*. Cambridge University Press.

Ayoun, D. (2007). The second language acquisition of grammatical gender and agreement. In D. Ayoun (Ed.), *French applied linguistics* (Vol. 16, pp. 130-170). John Benjamins.

Ayoun, D. (2018). Grammatical gender assignment in French: Dispelling the native speaker myth. *Journal of French Language Studies, 28*(1), 113-148. https://doi.org/10.1017/S095926951700014X

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software, 67*(1), 1-48.

Benmamoun, E. (2000). *The feature structure of functional categories: A comparative study of Arabic dialects*. Oxford University Press.

Benmamoun, E., & Albirini, A. (2018). Is learning a Standard variety similar to learning a new language?: Evidence from heritage speakers of Arabic. *Studies in Second Language Acquisition, 40*(1), 31-61. https://doi.org/10.1017/S0272263116000383

Buckwalter, T., & Parkinson, D. (2011). *A frequency dictionary of Arabic: Core vocabulary for learners*. Routledge.

Hawkins, R., & Towell, R. (2015). *French grammar and usage* (4th ed.). Routledge.

Lenth, R. V. (2024). *emmeans: Estimated Marginal Means, aka Least-Squares Means* (version 1.10.6 ed.). https://CRAN.R-project.org/package=emmeans

Montrul, S. (2008). *Incomplete acquisition in bilingualism: Re-examining the age factor*. Johns Benjamins

Montrul, S., Foote, R., & Perpiñán, S. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning, 58*, 503-553. https://doi.org/10.1111/j.1467-9922.2008.00449.x

Montrul, S. (2010). Dominant language transfer in adult second language learners and heritage speakers. *Second Language Research, 26*(3), 293-327. https://doi.org/10.1177/0267658310365768

Omar, M. (2007). *The acquisition of Egyptian Arabic as a native language.* Georogetown University Press.

Ortega, L. (2013). *Understanding second language acquisition*. Routledge.

Polinsky, M. (1995). Cross-linguistic parallels in language loss. *Southwest Journal of Linguistics, 14*(1-2), 87-123.

Revelle, W. (2025). *psych: Procedures for Psychological, Psychometric, and Personality Research* (version 2.5.3 ed.). Northwestern University. https://CRAN.R-project.org/package=psych

Ryding, K. C. (2005). *A reference grammar of Modern Standard Arabic*. Cambridge University Press.

Appendix

Full list of Nouns and Verbs Used in the Stimuli

	Noun Class	Subject Animate nouns		Translation	Verb Present		Translation
		Masculine	Feminine		Masculine	Feminine	
1	Human (x8)	الجدّ	الجدّة	grandfather-grandmother	يحكي	تحكي	tell
2		الولد	الفتاة	boy-girl	يجلس	تجلس	sit
3		الزوج	الزوجة	husband-wife	يشتري	تشتري	buy
4		الرجل	المرأة	woman-man	يساعد	تساعد	help
5		الطالب	الطالبة	student	يحلّ	تحلّ	solve
6		المذيع	المذيعة	broadcaster	يقرأ	تقرأ	read
7		المعلّم	المعلّمة	instructor	يكتب	تكتب	write
8		الطبيب	الطبيبة	physician	يذهب	تذهب	go
1	Animal (x8)	الجمل	الناقة	camel	يمشي	تمشي	walk
2		الثور	البقرة	cow	يسحب	تسحب	pull
3		الديك	الدجاجة	chicken	يأكل	تأكل	eat
4		الأسد	اللبوة	lion	يهاجم	تهاجم	attack
5		الثعبان	الحية	snake	يزحف	تزحف	crawl
6		القط	القطة	cat	يهرب	تهرب	escape
7		العصفور	العصفورة	bird	يقف	تقف	stand
8		الغزال	الغزالة	dear	يجري	تجري	run