

## An Analytical Study of Grammatical Agreement Errors in Gender (Masculine and Feminine): A Comparative Linguistic Perspective between Arabic and Malay

Kujaimah Binti Haji Abdul Kahar<sup>1</sup>, Achmad Yani<sup>2</sup>, Hajah Rafizah binti Haji Abdullah<sup>3</sup>, Munifah<sup>4</sup>, Nurkhamimi Zainuddin<sup>5</sup>.

<sup>1, 2 & 3</sup>Universiti Islam Sultan Sharif Ali (UNISSA), Brunei Darussalam.

<sup>4</sup>Universitas Islam Negeri (UIN) Syekh Wasil, Indonesia.

<sup>5</sup>Universiti Sains Islam Malaysia (USIM), Malaysia.

### Abstract

This study investigates grammatical gender agreement errors among Malay-speaking learners of Arabic from a comparative linguistic perspective, focusing on the typological contrast between a gender-marked language (Arabic) and a gender-neutral language (Malay). Employing a descriptive-analytical design grounded in Error Analysis theory, the research examines written data produced by thirty undergraduate students enrolled in the Arabic Language and Translation Programme at Sultan Sharif Ali Islamic University (UNISSA), Brunei Darussalam. Data were elicited through a structured written test consisting of 100 items targeting gender agreement across major syntactic domains. The analysis follows an integrated Error Analysis framework encompassing error identification, classification, description, and interpretation. The findings reveal 404 gender agreement errors, with the highest frequency occurring in verb-subject agreement, followed by subject-predicate and passive constructions. Two dominant error patterns emerge: over-masculinization and over-feminization, reflecting learners' reliance on default forms and incomplete mastery of Arabic gender morphology. These patterns are attributed primarily to negative cross-linguistic interference from Malay, which lacks grammatical gender, as well as to rule simplification within the Arabic agreement system. The study confirms that typological distance significantly shapes interlanguage development and error distribution. Pedagogically, the findings underscore the necessity of explicit contrastive instruction and focused practice in high-frequency agreement structures to enhance grammatical accuracy in Arabic as a foreign language.

**Keywords:** Grammatical, Errors, Arabic and Malay

### 1. Introduction

Grammatical agreement represents a core organizing mechanism in natural language, playing a decisive role in syntactic cohesion, semantic interpretation, and communicative accuracy. Among agreement systems, grammatical gender agreement is particularly complex in languages where it is fully grammaticalized and morphologically expressed. Errors in gender agreement therefore constitute more than surface-level inaccuracies; they reflect deeper challenges in grammatical competence and linguistic processing.

Arabic exemplifies a gender-marked language in which masculine and feminine distinctions permeate the entire morpho-syntactic system. Gender agreement is obligatorily encoded in nouns, verbs, adjectives, predicates, and circumstantial constructions, rendering it a fundamental component of grammatical well-formedness. Classical Arabic grammarians consistently regarded gender accuracy as a hallmark of linguistic mastery, classifying gender mismatch alongside violations of case marking and syntactic governance (al-Mubarrad, 1996, pp. 45–47). Within this tradition, grammatical gender is not a peripheral feature but a structural principle that sustains syntactic harmony.

Arabic gender operates through a binary opposition between masculine and feminine, each encompassing real (biological) and metaphorical (conventional) categories. While real gender corresponds to natural sex distinctions, metaphorical gender applies to inanimate and abstract nouns through historically established linguistic convention rather than semantic motivation. Classical scholarship emphasizes that metaphorical gender in Arabic is not rule-governed in a productive sense but must be acquired through exposure to attested usage (Hasan, n.d., pp. 112–118; al-Daqar, 1988, pp. 73–81). This characteristic significantly increases the learning burden for non-native speakers.

Traditional grammatical theory further conceptualizes the masculine as the unmarked default, whereas the feminine is treated as a marked category requiring explicit differentiation. Ibn Ya'ish highlights this asymmetry by noting that masculine forms can function generically, while feminine forms necessitate additional morphological or contextual marking (Ibn Ya'ish, n.d., Vol. 3, pp. 9–12). This structural hierarchy has important implications for learner behavior, particularly in contexts of default form selection.

Gender agreement in Arabic is obligatorily realized across several syntactic domains, most notably verb-subject agreement, subject-predicate agreement, adjective-noun concord, and circumstantial constructions (ḥāl). Verbs inflect for gender through systematic prefixation and suffixation, adjectives must match their head nouns in gender alongside number and definiteness, and circumstantial expressions must agree with their controllers despite their fixed accusative and indefinite status (Ryding, 2005, pp. 139–176). These layered agreement requirements render Arabic gender morphology both salient and cognitively demanding.

In sharp typological contrast, Malay, an Austronesian language, lacks grammatical gender as a morpho-syntactic category. Malay nouns, verbs, and adjectives do not inflect for gender, and agreement relations are entirely independent of masculine-feminine distinctions (Mees, 1969, pp. 21–25; Za'ba, 1965, pp. 42–46). Gender is expressed only when semantically relevant through optional lexical modifiers such as *lelaki* and *perempuan*, while inanimate entities remain gender-neutral (Haji Omar, 1986, pp. 67–70). This structural absence positions Malay at the opposite end of the gender-marking continuum from Arabic.

From a second language acquisition perspective, such typological disparity is known to generate persistent learning difficulties. Contrastive and transfer-based models predict heightened error rates when a grammatical feature is obligatory in the target language but absent in the learner's first language (Lado, 1957, pp. 2–5; Odlin, 1989, pp. 27–32). For Malay-speaking learners of Arabic, grammatical gender agreement represents precisely such a feature.

Empirical learner data corroborate these predictions. Gender agreement errors occur most frequently in verb-centered constructions, followed by subject-predicate agreement and passive structures, while adjectival and circumstantial constructions exhibit lower but still notable error rates. Qualitative analysis reveals two dominant patterns: over-masculinization, whereby learners default to masculine forms in feminine contexts, and over-feminization, whereby feminine markers are extended inappropriately. The prevalence of masculine defaults reflects its unmarked grammatical status, while compound mismatches in gender and number suggest incomplete integration of agreement features.

These error patterns arise from the interaction of intralingual simplification strategies and negative cross-linguistic interference. Learners tend to reduce the complexity of the Arabic agreement system by relying on default forms, while simultaneously transferring L1 processing strategies that treat gender as semantically optional rather than grammatically obligatory. This interaction aligns with interlanguage theory (Selinker, 1972, pp. 214–216) and contemporary models of cross-linguistic influence (Jarvis & Pavlenko, 2008, pp. 21–25).

Accordingly, this study undertakes an analytical and comparative investigation of grammatical gender agreement errors among Malay-speaking learners of Arabic. By situating learner errors within a contrastive linguistic framework, the study seeks to identify the structural sources of persistent agreement difficulties and to contribute to both theoretical understanding and pedagogical practice. The findings underscore the need for explicit contrastive instruction and focused practice in high-frequency agreement domains to enhance grammatical accuracy and communicative competence in Arabic as a foreign language.

## 2. Theoretical Framework and Previous Studies

The present study is anchored in an integrative theoretical framework drawing on contrastive linguistics, cross-linguistic influence, interlanguage theory, and input-processing models within second language acquisition (SLA). This composite framework is particularly appropriate for examining grammatical gender agreement errors among Malay-speaking learners of Arabic, given the profound typological divergence between the two languages with respect to gender encoding.

From a contrastive linguistic perspective, Arabic and Malay occupy opposing typological positions. Arabic employs grammatical gender as a fully grammaticalized morpho-syntactic category, while Malay lacks grammatical gender altogether. According to classical contrastive analysis, such non-corresponding features are expected to generate persistent acquisition difficulties, as learners cannot rely on direct L1–L2 structural mapping (Lado, 1957; Comrie, 1989). In Arabic, gender agreement is obligatory across verbs, predicates, modifiers, and circumstantial constructions, whereas Malay expresses gender only lexically and optionally. This structural asymmetry renders Arabic gender agreement a cognitively demanding feature for Malay learners.

The concept of cross-linguistic influence further explains how L1 structures shape L2 processing. Odlin (1989) argues that negative transfer is most pronounced when a grammatical feature is obligatory in the target language but absent in the learner's first language. For Malay speakers, whose L1 treats gender as semantically driven rather than grammatically required, Arabic gender marking may initially be perceived as optional. Jarvis and Pavlenko (2008) extend this view by demonstrating that learners transfer not only forms but also underlying assumptions about grammatical relevance, leading to systematic omission or defaulting strategies in L2 production.

Within interlanguage theory, learner errors are understood as manifestations of a developing, rule-governed linguistic system rather than random performance failures (Selinker, 1972). Over-masculinization and inconsistent feminization in Arabic gender agreement reflect interim grammatical rules that simplify the target system while maintaining communicative efficiency. The persistence of such errors, even at advanced proficiency levels, can be attributed to fossilization, particularly when communicative success is attainable despite grammatical inaccuracy.

Input-processing theory provides additional explanatory power by highlighting the role of attentional constraints in grammatical acquisition. VanPatten (2015) posits that learners prioritize meaning-bearing elements over low-salience grammatical markers during processing. In Arabic, gender agreement markers—especially verbal affixes—carry limited semantic weight but high grammatical significance. Consequently, Malay-speaking learners may under-process these markers, resulting in high error rates in morphologically dense constructions such as verb–subject agreement.

Empirical studies consistently support these theoretical claims. Research on Arabic as a foreign language reports that gender agreement is among the most persistent sources of learner error, particularly in verb-centered and sentential agreement domains (Ryding, 2005; Al-Sā'idī, 2021). Broader SLA research likewise demonstrates that grammatical features absent in the L1 are acquired later and with greater variability (Ellis, 1997; White, 2003). Studies involving Malay learners further indicate a tendency to conceptualize gender as a semantic attribute rather than a grammatical constraint, leading to systematic agreement errors (Haji Omar, 1986).

Despite these findings, existing research rarely integrates contrastive linguistic analysis with quantitative and qualitative error data in a unified framework. This study addresses that gap by examining grammatical gender agreement errors through a comparative Arabic–Malay lens, thereby contributing to a more precise understanding of cross-linguistic interference and grammatical development in Arabic as a foreign language.

## 3. Methodology

This study adopts a descriptive–analytical research design grounded in the theoretical and methodological principles of Error Analysis. The descriptive approach is employed to construct the conceptual framework of grammatical agreement through a systematic review of relevant linguistic literature, including foundational works in second language acquisition and contrastive linguistics (Corder, 1981, pp. 9–36; Ellis, 1987, pp. 45–68). Complementing this approach, Error Analysis is utilized to identify, classify, and interpret grammatical agreement errors produced by learners in written Arabic, with the aim of uncovering their underlying causes and pedagogical implications.

The research population consists of thirty (30) undergraduate students enrolled in the first and second years of the Arabic Language and Translation Programme at the Faculty of Arabic Language, Sultan Sharif Ali Islamic University (UNISSA), Brunei Darussalam, during the 2021/2022 academic year. All participants are Bruneian Malay native speakers learning Arabic as a foreign language, aged between 18 and 21, which ensures relative homogeneity in linguistic background and learning context.

Data were collected through a structured written test, in line with the methodological preference for written production in Error Analysis research (Fisiak, 1981, pp. 227–241). The test comprised 100 items completed within a two-hour period and included multiple-choice questions, gap-filling tasks, and sentence reformulation exercises. Such guided written tasks are widely recognized as effective tools for eliciting systematic grammatical errors (Sīnī & al-Amīn, 1982, pp. 33–57).

Data analysis followed an integrated Error Analysis framework synthesizing the models proposed by Corder (1981, pp. 51–78), Ellis (1987, pp. 296–322), and Sridhar (1985, pp. 207–241). The analytical procedure proceeded through four stages: data collection, error identification, error classification and description, and error interpretation and correction. This systematic process enabled a principled examination of grammatical agreement errors within the learners' interlanguage system and allowed for explanations informed by cross-linguistic influence from Malay to Arabic.

## 4. Grammatical Gender Agreement in Arabic and Malay: Patterns, Learner Errors, and Cross-Linguistic Interference

Arabic is a morphologically rich language that encodes grammatical gender as a fundamental organizing principle across its syntactic system. Masculine and feminine distinctions permeate nominal structures, verbal inflection, adjectival modification, and circumstantial constructions, rendering gender agreement an essential component of grammatical competence and communicative precision. Classical Arabic grammarians consistently regarded accurate gender distinction as a core indicator of linguistic mastery, equating gender mismatch with serious grammatical deviation comparable to errors in case marking or syntactic governance (al-Mubarrad, 1996).

From a semantic and grammatical perspective, Arabic gender classification operates through a well-established dichotomy. The masculine category is typically divided into real masculine, referring to nouns denoting male beings with corresponding female counterparts, and metaphorical masculine, encompassing inanimate or abstract nouns conventionally treated as masculine. The feminine category mirrors this division, comprising real feminine nouns denoting female beings and metaphorical feminine nouns whose gender assignment derives from linguistic convention rather than biological sex. Crucially, metaphorical gender in Arabic is not governed by productive rules but is determined through attested usage and native-speaker intuition, underscoring the language-specific nature of grammatical gender (Ḥasan, n.d.; al-Daqar, 1988).

Arabic grammatical theory traditionally conceptualizes the masculine as the unmarked default and the feminine as the marked derivative, typically requiring overt or implicit morphological markers. Ibn Ya'ish argues that this asymmetry is evidenced by the existence of masculine generic nouns capable of encompassing both genders and by the structural necessity for feminine markers to differentiate the feminine category from the default masculine (Ibn Ya'ish, n.d.). Consequently, gender in Arabic is not merely a lexical feature but a fully grammaticalized system embedded in morpho-syntactic agreement relations.

Gender agreement in Arabic is obligatory across several major syntactic domains. First, subject–predicate (mubtada’–khabar) agreement requires gender matching when the predicate is a derived adjective directly predicated of the subject. Agreement is neutralized only in specific contexts, such as non-derived predicates or gender-invariant forms. Second, verb–subject agreement represents one of the most salient manifestations of gender marking, with verbs obligatorily reflecting the gender of their subjects through suffixation in the perfect tense and prefixation in the imperfect. Third, modifier–head agreement, particularly in adjective–noun structures, requires strict matching in gender alongside number, definiteness, and case. Fourth, circumstantial constructions (hāl) require the circumstantial expression to agree with its controller in gender and number, despite its fixed accusative and indefinite status. Additional agreement-sensitive constructions include emphasis and apposition, which rely on precise pronominal and morphological alignment.

In stark contrast, Malay belongs to the Austronesian (Malayo-Polynesian) language family and is typologically characterized by the absence of grammatical gender. Malay nouns do not encode masculine or feminine distinctions morphologically, nor do verbs or adjectives inflect for gender agreement (Mees, 1969; Za’ba, 1965). Gender distinctions are expressed only when semantically necessary through optional lexical modifiers such as *lelaki* “male (human)”, *perempuan* “female (human)”, *jantan* “male (animal)”, and *betina* “female (animal)”. Certain kinship terms inherently encode sex, while inanimate entities remain gender-neutral except in rare metaphorical extensions (Haji Omar, 1986). This typological divergence establishes a sharp contrast between Arabic as a gender-marked language and Malay as a gender-neutral one.

This contrast has significant implications for second-language acquisition. Analysis of learner data revealed 404 grammatical gender agreement errors, distributed unevenly across syntactic domains. The most frequent errors occurred in verb–subject agreement (44%), followed by subject–predicate agreement (24%), verb–passive subject agreement (18%), adjective–noun agreement (7%), circumstantial agreement (4%), and emphatic constructions (3%). The predominance of verb-related errors suggests that morphologically complex agreement domains pose the greatest challenge to learners.

Qualitative analysis indicates two recurring error patterns: (i) over-masculinization, whereby learners default to masculine forms in contexts requiring feminization, and (ii) over-feminization, whereby feminine forms are applied in masculine contexts. The first pattern is particularly prevalent in verb–subject constructions, reflecting reliance on the unmarked masculine form as a processing default. Errors in circumstantial constructions often involve simultaneous mismatches in gender and number, indicating incomplete integration of multiple agreement features. Emphatic constructions further reveal instability in learners’ control of gender-sensitive pronominal elements.

These patterns can be attributed to two interrelated factors. The first is rule overgeneralization and simplification within the Arabic agreement system, whereby learners reduce a complex system to a limited set of defaults. The second, and more decisive, factor is negative cross-linguistic interference from Malay. Because Malay does not grammatically encode gender agreement, learners may fail to perceive gender marking as obligatory in Arabic. Instead, they may treat gender as optional or lexically determined, transferring L1 processing strategies into L2 production. Such interference aligns with established models of cross-linguistic influence, which predict heightened error rates when structural features present in the target language are absent in the learner’s first language (Odlin, 1989; Jarvis & Pavlenko, 2008).

Grammatical gender agreement constitutes a major source of difficulty for Malay-speaking learners of Arabic, particularly in verb-centered and sentential agreement domains. These difficulties stem from both the intrinsic complexity of Arabic gender morphology and the profound typological gap between Arabic and Malay. Pedagogically, the findings underscore the necessity of explicit contrastive instruction, systematic practice in high-frequency agreement structures, and heightened learner awareness of gender as a grammatical—not merely semantic—feature. Such strategies may mitigate persistent agreement errors and support more accurate and fluent Arabic language acquisition.

## 5. Conclusion

This study has shown that grammatical gender agreement represents a major source of difficulty for Malay-speaking learners of Arabic, particularly in verb–subject and sentential agreement structures. The error patterns observed are systematic and strongly motivated by the absence of grammatical gender in Malay, resulting in negative cross-linguistic transfer and reliance on unmarked masculine forms. These findings support established models of interlanguage development and cross-linguistic influence, confirming that features absent from the first language pose heightened acquisitional challenges when fully grammaticalized in the target language. The study highlights the need for pedagogical approaches that explicitly address typological differences through contrastive explanation, focused practice, and feedback on high-frequency agreement domains. By integrating insights from Arabic grammatical theory and second language acquisition research, this study contributes to a more nuanced understanding of how typological distance shapes the acquisition of core morpho-syntactic features.

## References

- Al-Daqar, ‘Abd al-Qādir. (1988). *Dirāsāt fī al-naḥw al-‘Arabī* [Studies in Arabic grammar]. Damascus: Dār al-Qalam.
- Al-Mubarrad, Muḥammad ibn Yazīd. (1996). *Al-Muqtaḍab* (Vol. 1). Edited by Muḥammad ‘Abd al-Khāliq ‘Aḏīmah. Cairo: ‘Ālam al-Kutub.
- Al-Šā’idī, Muḥammad. (2021). Gender agreement errors in Arabic as a foreign language: An interlanguage perspective. *Journal of Arabic Linguistics*, 14(2), 55–78.
- Comrie, Bernard. (1989). *Language universals and linguistic typology: Syntax and morphology* (2nd ed.). Chicago: University of Chicago Press.
- Corder, S. Pit. (1981). *Error analysis and interlanguage*. Oxford: Oxford University Press.
- Ellis, Rod. (1987). *Second language acquisition*. Oxford: Oxford University Press.
- Ellis, Rod. (1997). *Second language acquisition*. Oxford: Oxford University Press.
- Fisiak, Jacek (Ed.). (1981). *Contrastive linguistics and the language teacher*. Oxford: Pergamon Press.
- Ḥasan, ‘Abbās. (n.d.). *Al-Naḥw al-wāfi* (Vols. 1–2). Cairo: Dār al-Ma’ārif.
- Haji Omar, Asmah. (1986). *Bahasa dan alam pemikiran Melayu* [Language and Malay worldview]. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Ibn Ya’īsh, Ya’īsh ibn ‘Alī. (n.d.). *Sharḥ al-Mufaṣṣal* (Vol. 3). Beirut: Dār al-Kutub al-‘Ilmiyyah.
- Jarvis, Scott, & Pavlenko, Aneta. (2008). *Crosslinguistic influence in language and cognition*. New York: Routledge.
- Lado, Robert. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor: University of Michigan Press.
- Mees, C. A. (1969). *Tatabahasa Melayu* [Malay grammar]. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Odlin, Terence. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge: Cambridge University Press.
- Ryding, Karin C. (2005). *A reference grammar of Modern Standard Arabic*. Cambridge: Cambridge University Press.
- Sinī, Maḥmūd Ismā’īl, & al-Amīn, Aḥmad. (1982). *Taḥlīl al-akḥṭā’ fī ta’līm al-lughāt* [Error analysis in language teaching]. Riyadh: King Saud University Press.
- Selinker, Larry. (1972). Interlanguage. *International Review of Applied Linguistics*, 10(3), 209–231.
- Sridhar, S. N. (1985). Contrastive analysis, error analysis, and interlanguage. In J. C. Richards (Ed.), *Error analysis: Perspectives on second language acquisition* (pp. 207–241). London: Longman.
- VanPatten, Bill. (2015). *Input processing in second language acquisition*. New York: Routledge.
- White, Lydia. (2003). *Second language acquisition and Universal Grammar*. Cambridge: Cambridge University Press.
- Za’ba (Zainal Abidin Ahmad). (1965). *Pelita Bahasa Melayu* [The lamp of Malay language] (Vol. 1). Kuala Lumpur: Dewan Bahasa dan Pustaka. York, NY: Routledge.