Do Large Language Models Understand Double Mismatches? Evidence from Farsi

Maryam Mohammadi
Bielefeld University
maryam.mohammadi@uni-bielefeld.de

Abstract

Large language models (LLMs) are increasingly used for communication in many languages, therefore, understanding their limitations with respect to culture-specific pragmatics is important. While LLMs perform well on statistically frequent structures, their shortcomings are most evident in rare pragmatic phenomena. This study investigates whether LLMs can generate a (rare) complex honorific mismatch in Farsi. The pattern arises at two levels: (i) a plural pronoun disagrees with a singular referent for the sake of honorification, and (ii) the related components violate the Polite Plural Generalization due to intimacy implication. This *double mismatch* pattern is attested in everyday speech, though it is statistically sparse. We tested GPT-4 across multiple scenarios. The results reveal that the model successfully employs the first mismatch to indicate honorific, but fails to adopt the second mismatch that simultaneously conveys intimacy. The model thus deviates from human-like behavior at the syntax–pragmatics interface. These findings suggest that, while machine models demonstrate partial success in generating honorifics, they rely primarily on statistical patterns and lack the deeper pragmatic understanding necessary for contextual competence.

1 Introduction

In many languages, plural pronouns are used to express politeness when addressing a single honorific individual (Brown and Levinson, 1987), resulting in the (first) mismatch at the syntax–semantics interface between the plural form and its singular referent (e.g., using *Sie* 'you.2.PL' instead of *du* 'you.2.SG' in German to address a single individual politely). Comrie (1975) investigates the behavior of honorific pronouns and shows that while in some languages all related components (e.g., possessives and verb conjugations) symmetrically agree with polite plurals, in other languages the honorific pronoun controls plural agreement only on certain components. Notably, when the target components exhibit plural agreement with polite pronouns, they tend to do so consistently.

Such mismatch patterns are well studied across languages and are frequently used by speakers to systematically signal politeness. Consequently, large language models can readily learn these patterns, adopt their pragmatic functions, and deploy them in appropriate contexts (Noh et al., 2024). Farsi, as a positive face–saving language, employs many types of politeness (see Gohari Sadr et al., 2025, for the study on *Taarof*). Interestingly, Farsi exhibits a second type of mismatch that goes beyond politeness and is used to express intimacy toward an honorific addressee. This double–mismatch is statistically less frequent than the purely polite form, however, it is pragmatically effective in appropriate contexts.

This study investigates number (dis)agreement patterns in Farsi, focusing on their interaction at the syntax–pragmatic interface. The pronoun *šoma* serves both as the second-person plural and as the honorific second-person singular. In line with the *Polite Plural Generalization* (Wechsler, 2011), Farsi typically requires the polite plural controller to determine plural number agreement on all components marked for person and number features, i.e., verbs and related pronouns. Our data show that speakers employ singular agreement with the honorific plural pronoun in order to convey intimacy alongside politeness, but this pragmatic emerges only under specific settings. The present study explores how LLMs process double mismatches in Farsi. Our initial results reveal that LLMs successfully employ the first mismatch, exhibiting plural pronouns in honorific contexts. However, they fail to capture the second mismatch, expressing intimacy with the honorific addressee.

Proceedings of the First Workshop on NLP and LLMs for the Iranian Language Family (SilkRoadNLP), pages 24–28
March 29, 2026 ©2026 Association for Computational Linguistics

2 Polite Plural

Plural pronouns are frequently used to express politeness when addressing a single honorific individual (Brown and Levinson, 1987). According to the Polite Plural Generalization (Wechsler, 2011), a polite plural pronoun functions as an agreement controller, imposing plural number on any related components marked for person and number features. In many languages, second-person plural pronouns employed honorifically for a single addressee systematically trigger plural agreement on all eligible targets (Wechsler, 2011).

Comrie (1975) examines the syntax–semantics behavior of polite/honorific pronouns, showing that while in some languages (e.g., Croatian) all agreement targets (e.g., verbs, possessives, reflexives pronouns) uniformly agree with polite plurals, in others (e.g., French) the honorific pronoun triggers plural agreement only on certain targets (e.g., verbal predicates but not adjectival predicates). This contrast motivates a typology distinguishing *uniform* and *mixed* agreement systems. Generally, languages tend to adhere to their (mis)matching patterns consistently, whether uniform or mixed. In other words, while such patterns vary across languages, they are fixed within individual languages.

There are two main approaches in the literature; in the first syntactic approach, honorificity is encoded on a functional projection (e.g., CP) in the clause–periphery, via a feature. This projection is responsible for licensing honorificity on all 2.person components in the clause, via binding on 2.person pronouns (Portner et al., 2019; Alok, 2021). In the second semantic–pragmatic approach, the expressive honorific content on the lexical entries of the two items is comparable in semantics/pragmatics (Potts, 2007; McCready, 2019). That is, if a speaker uses an item which honors the addressee, and also an item that dishonors the addressee within the same sentence, the combination would be conceptually strange and language would block the combination.

We will see that the double mismatch in Farsi does indeed occur, pragmatically conveying both honorificity and intimacy simultaneously (see also Puškar-Gallien, 2019, for discussions on verb–politeness conflicts).

3 Core Data

In this section, we present data from Farsi; particularly the colloquial form form spoken in Tehran, Iran. Farsi is an SOV language and exhibits subject–verb agreement in person and number. Like many other languages, Farsi employs polite plural: the pronoun *šoma* functions both as the canonical second person plural form and as an honorific singular form. In line with the Polite Plural Generalization (Wechsler, 2011), Farsi typically requires the polite plural controller to determine plural number agreement on all components marked for number feature, including verbs and related pronouns (e.g., possessives and reflexives).

As illustrated in example (1), the speaker addresses her grandmother in polite plural pronoun *šoma* you.2.PL (the $\rightsquigarrow$ indicates the added implication). The polite pronoun controls the number agreement on the verb and the possessive. Note that in Farsi, subject (A1) and object (A2) pronouns have the same form.

(1) **Context:** A is talking with her grandma.

A1: *šoma* *čâe-tun* *ro* *xord-id?*
you.2.PL tea-2.PL.POSS ACC eat-2.PL

'Did you have your tea?' $\rightsquigarrow$ Honorific implication

A2: *man* *lebâse* *šoma* *ro* *behe-tun* *dâdam.*
I clothes you.2.PL ACC to-2.PL gave

'I gave your clothes to you.' $\rightsquigarrow$ Honorific implication

Colloquial data shows that number disagreement can be strategically employed to convey an *Honorific* yet *Intimacy* relationship with the addressee. In the same context (1), alternative (A1′) features a plural subject pronoun, while the corresponding possessive and the verb appear in singular form, resulting in a number mismatch with the plural pronoun controller (Ferguson, 1991).

A1′: *šoma* *čâe-t* *ro* *xord-i?*
you.2.PL tea-2.SG.POSS ACC eat-2.SG

'Did you have your tea?' $\rightsquigarrow$ Honorific and Intimacy.

Crucially, the pattern is pragmatically infelicitous in the absence of intimacy between the interlocutors, for instance, when addressing a socially superior individual (e.g., in a manager–employee relationship). This deliberate mismatch reflects a nuanced role of discourse on syntactic structures.

By contrast, in (A1″), the use of plural object possessive pronouns, agreeing with the plural subject, alongside a singular verb results in infelicity. This suggests that, the number feature must be symmetrically underspecified across all target components in the sentence.

A1″:# *šoma* *čâe-tun* *ro* *xord-i?*
you.2.PL tea-2.PL.POSS ACC eat-2.SG

It is worth to note that since Farsi is a subject-drop language, the subject pronoun *šoma* can optionally be omitted. While removing *šoma* from (A1′) eliminates the honorific implication, omitting it in (A1″) does not resolve the infelicity caused by the number mismatch, though the form still includes a plural pronoun, in clitic form, and a singular verb.

The same pattern holds when *šoma* appears as an object pronoun. Object pronouns in Farsi can occur in either free form or clitic form that attaches to objects or verbs (Rasekh Mahand, 2014). In example (2), using the free object pronoun *šoma* with number disagreement with the reflexive in (A1) is pragmatically felicitous, however, in (A2), the object clitic form *-tun* cannot achieve the same effect and results in infelicity. Thus, this underspecification is possible only with the free pronouns *šoma*, but not with the clitic counterpart *-tun*.

(2) **Context:** A is talking with her grandma.

A1: *qorse* *šoma* *ro* *emruz* *bara-t* *mixram.*
tablet you.2.PL ACC today for-2.SG buy

'I will buy your tablet today.'

A2:# *qorse-tun* *ro* *emruz* *bara-t* *mixaram.*
tablet-your.PL ACC today for-2.SG buy

4 LLMs and Honorific Mismatches

Recent advances in large language models have prompted linguists to explore how these artificial systems handle the complex structures characteristic of natural languages. While some theoretical linguists remain skeptical about the relevance of LLMs for studying natural language, their utility extends beyond serving as sophisticated computational artifacts. They can function as valuable tools for testing linguistic hypotheses and refining theories of human linguistic competence.

Given the complexity of double mismatch pattern in Farsi, the present study aims to evaluate whether LLMs can recognize and adopt these less frequent yet highly functional forms. Although honorific (mis)matches have been extensively investigated in English and other European languages, the behavior of LLMs with respect to Farsi remains largely unexplored (see Gohari Sadr et al., 2025, among all). In this paper, we aim to control the performance of LLMs on Farsi double mismatches.

It is worth acknowledging that this study primarily focused on the linguistic analysis of double mismatch phenomena. However, we extended our investigation as a pilot study on LLM behavior in double mismatches. Here, we report only the preliminary results of this first attempt. Regarding the intriguing initial observations, in which the tested model failed to capture the second mismatch in intimacy, a more comprehensive evaluation is reserved for our future work. Our next steps will involve a strengthened experimental design, testing more contexts across multiple LLMs, including Persian models (e.g., ParsBERT[1]), and comparing model performance with a larger baseline of native speakers. Although our current investigation is limited in both the number of trials and the number of models tested, the findings underscore the importance of examining how LLMs process complex pragmatic patterns.

We tested GPT-4 on the double mismatch pattern using six carefully designed contexts. Each context involved two interlocutors with an age difference (e.g., grandchild–grandparent, niece–elder uncle) to trigger honorific marking, as well as a friendly relation to trigger intimacy. The experimental settings, illustrated in example (3), positioned the model as the speaker, who asks a question from the honorific addressee. In all trials, we included a possessive or reflexive pronoun that, in addition to the verb, should agree with the subject. If the LLM only adopts intimacy, all referring forms should appear in the singular, as in (A1). If it adopts the single mismatch (honorific), pronouns and verbs should agree with the polite plural, as in (A2). If it adopts the double mismatch (honorific + intimacy), the agreement components should appear in the singular form, regardless of the polite plural pronoun, as in (A3). Note that the prompts were originally presented in Farsi, they are provided here in English for brevity.

(3) **Prompt:** You have a close relationship with your grandma. You are sitting with your grandparents in the living-room. Address your grandma and ask if she has already taken her tablets.

A1: *to* *qorshâ-t* *ro* *xord–i?*
you.2.SG tablets-2.SG ACC ate-2.SG

A2: *šoma* *qorshâ-tun* *ro* *xord–id?*
you.2.PL tablets-2.PL ACC ate-2.PL

A3: *šoma* *qorshâ-t* *ro* *xord–i?*
you.2.PL tablets-2.SG ACC ate-2.SG

'Did you have your tablets?'

[1]Thanks to the EACL reviewer for the helpful suggestion.

Across all six contexts, GPT consistently adopted the intimacy interpretation over politeness and produced singular forms, as illustrated in (A1). We then tested the same contexts with the explicit close-relationship cue replaced by an implicit formulation. For instance, in example 3, we replaced '*You have a close relationship with your grandma*' with '*You grew up with your grandma*'. This time, the model constantly produced the honorific mismatch pattern, as in (A2), ignoring the intimacy.[2]

Notably, we validated the trial contexts with four native Farsi speakers. The results indicate that they readily interpreted both experimental settings, those with explicit and implicit cues, and predominantly produced the double mismatch forms shown in (A3), while occasionally producing the other two forms.

Although the number of trials is admittedly limited, the results underscore the importance of investigating complex pragmatic patterns in LLMs. Double mismatches require the language user, whether human or machine, to consider two layers of social relations, encompassing both honorific marking and intimacy. While the honorific and intimacy distinctions are each frequent and readily captured by LLMs in isolation, their combination at the second layer appears less accessible, likely due to its relative infrequency.

It is worth noting that we observe double mismatch patterns in two conditions: (i) when the addressee is older than the speaker, as shown in (3), and (ii) when the addressee is (significantly) younger than the speaker. Following the politeness principles (Brown and Levinson, 1987), in the second condition the speaker typically does not use the honorific plural.

However, in our investigation, we observed several instances of the second condition in which double mismatches were employed. For example, some parents use the polite plural when addressing their children to emphasize the child's developing character, signaling that, despite their young age, they should behave respectfully. In this regard, some parents use the double mismatch to maintain intimacy with their children.

As illustrated in example (4), the mother addresses her three-year-old son with the polite plural *šoma* to practically teach respectful behavior, while using the singular verb to preserve intimacy.

(4) **Context:** A mother is talking to her three-year-old son.
A: šoma pofak mixor–i?
you.2.PL crunchy eat-2.SG

'Do you want a crunchy?' ⇝ Honorific and Intimacy.

We evaluated this condition using GPT-4 across four scenarios. In every case, the model consistently used the second-person singular pronoun '*to*' (corresponding to pattern (A1) in example (3) above), failing to produce the double mismatch option. Notably, our pilot native informants also showed less frequent use of double mismatches in this setting. This can be explained by the fact that using honorifics for children reflects emerging psychological and behavioral practices among younger parents and is less familiar to older parents. Consequently, the double mismatch in this context is currently less prevalent. We expect it to be used more frequent among younger speakers and, consequently, to increase in the near future by (at least human) speakers.

5 Discussion

This study examines syntax–pragmatic mismatches in Farsi, focusing on the interaction between the polite plural and its number-agreement targets. The data show that the number feature is strategically underspecified to signal intimacy alongside politeness. Specifically, the polite plural pronoun *šoma* mismatches with its singular referent due to honorific implication, while other components, such as the verb and associated pronouns, remain in the singular form to convey intimacy. Crucially, this underspecification operates symmetrically across all relevant components. Thus, although the polite plural appears to mismatch its targets, Farsi nonetheless exhibits *uniform* (dis)agreement in number at the syntax–pragmatics interface.

We tested GPT-4 on double mismatch constructions to investigate how a representative large language model handles such complex pragmatic function. The initial results suggest that, although the model exhibits partially human-like behavior in processing honorifics, it fails to produce the double mismatch form and instead appears to rely primarily on statistical cues in the training data.

Although the model was prompted with only a small number of trials, it consistently failed

[2] As an anonymous reviewer rightly noted, changing the explicit close-relationship cue to an implicit one may make it unclear what exactly the model is sensitive to (honorifics, kinship, register, or lexical triggers). We will consider this in our future work.

to generate double mismatches and instead produced the plural honorific in a single-mismatch form. This outcome highlights the importance of linguistic patterns that, while statistically sparse, are pragmatically functional. Such patterns are often overshadowed by highly frequent constructions in LLMs, which can lead to their omission in downstream LLM-based frameworks or applications.

Given the increasing role of LLMs in everyday communication, it is crucial that these models are trained to recognize and produce complex pragmatic patterns, ensuring that low-frequency yet functionally significant forms are not lost in future language technologies. Such rare sociopragmatic constructions are precisely where LLMs often struggle, and our data provide clear evidence of this limitation. Understanding these shortcomings in handling culture-specific pragmatic phenomena is therefore essential. We emphasize that the statistical rarity of certain syntactic exceptions encoding complex pragmatic meanings warrants further investigation, as such research may inform improvements in LLM architectures and training strategies.

Limitations

This study was conducted under a tight timeline as a pilot extension of a broader linguistic investigation of mismatch patterns at the syntax–pragmatics interface. As such, our novel human-based findings were not systematically tested in large language models. Consequently, the results should be interpreted as preliminary rather than robust. Our investigation is limited in several respects. First, the number of prompt contexts was small. Second, we tested only a single model (GPT-4), which restricts the generalizability of the findings. Third, no statistical analysis was conducted, and the evaluation is therefore underpowered.

While the findings underscore the importance of examining LLMs with rare pragmatic patterns, the limitations prevent strong empirical claims regarding model behavior. Furthermore, the present study focuses exclusively on the generation of double mismatch patterns. The perception and interpretation of such mismatches by LLMs were not examined and remain for future research. As this is an ongoing project, we plan to address these limitations by expanding the range of models and contexts tested, incorporating systematic statistical analyses, and investigating both production and perception of mismatch patterns in future works.

Acknowledgments

This research has been funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation): CRC 1646/1 2024 – 512393437, Project INF.

References

Deepak Alok. 2021. The morphosyntax of magahi addressee agreement. *Syntax*, 24:263–296.

Penelope Brown and Stephen C. Levinson. 1987. *Politeness: Some Universals in Language Usage*. Cambridge University Press, Cambridge, UK.

Bernard Comrie. 1975. Polite plurals and predicate agreement. *Language*, 51(2):406–418.

Charles A. Ferguson. 1991. Individual and social in language change: Diachronic changes in politeness agreement in forms of address. In Robert L. Cooper and Bernard Spolsky, editors, *The Influence of Language on Culture and Thought: Essays in Honor of Joshua A. Fishman's Sixty-Fifth Birthday*, pages 183–198. De Gruyter Mouton, Berlin and Boston.

Nikta Gohari Sadr, Sahar Heidariasl, Karine Megerdoomian, Laleh Seyyed-Kalantari, and Ali Emami. 2025. We politely insist: Your llm must learn the persian art of taarof. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 1819–1838. Association for Computational Linguistics.

Elin McCready. 2019. *The Semantics and Pragmatics of Honorification*. Oxford University Press, Oxford.

Kangsan Noh, Sanghoun Song, and Eunjeong Oh. 2024. How language models understand honorific mismatches in korean. *Language Research*, 60(3):303–322.

Paul Portner, Miok Pak, and Raffaella Zanuttini. 2019. The speaker–addressee relation at the syntax–semantics interface. *Language*, 95(1):1–36.

Christopher Potts. 2007. The expressive dimension. *Theoretical Linguistics*, 33(2):165–198.

Zrinka Puškar-Gallien. 2019. Resolving polite conflicts in predicate agreement. *Glossa: A Journal of General Linguistics*, 4(1):33.

Mohammad Rasekh Mahand. 2014. Persian clitics: Doubling and agreement. *Modern Language Journal*, 24:16–33.

Stephen Wechsler. 2011. Mixed agreement, the person feature, and the index/concord distinction. *Natural Language & Linguistic Theory*, 29(3):999–1031.