Check for updates

Original Article

QJEP

Quarterly Journal of Experimental Psychology
1–18
© Experimental Psychology Society 2025
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/17470218251348218
qjep.sagepub.com
Sage

Gender effects on non-gendered pronouns in Hindi and Mandarin Chinese

Kumiko Fukumura[1], Shi Zhang[1,2], Sakshi Bhatia[3] and Samar Husain[4]

Abstract

We investigated strategies to avoid referential ambiguity in pronoun use. The non-linguistic competition account suggests that speakers avoid pronouns when referential candidates share the same gender, as increased similarity between them triggers a need for more specific referential information. We tested this hypothesis in Hindi and Mandarin Chinese, both of which use non-gendered pronouns. In Hindi, gender similarity between referential candidates reduced pronoun usage, supporting the account. In spoken Mandarin, where null pronouns are common, the use of overt pronouns was unaffected by either gender-based or situation-based similarity (i.e., whether more than one entity in the visual display could serve as a potential referent), while situation-based competition reduced the use of null pronouns. In written Mandarin, overt pronouns were preferred over null pronouns, and both gender- and situation-based competition influenced their use, although the gender effect was marginal. Null pronouns were unaffected by gender-based competition regardless of modality. These findings suggest that gender-based competition depends on pronoun features (e.g., null vs. overt) and the communication mode that influences pronoun preferences and susceptibility to similarity-based competition.

Keywords

Ambiguity; pronoun; gender; Hindi; Mandarin Chinese

Received: 3 February 2025; revised: 16 April 2025; accepted: 7 May 2025

In many languages, pronouns such as *she*, *he*, and *they* (as in English) are frequently used as substitutes for more explicit referring expressions like full nouns and names. While pronouns may enhance production efficiency, their minimal semantic content can sometimes lead to referential ambiguity. For instance, the pronoun *he* can refer to any male, potentially creating confusion when multiple male referents are present in the same context. This raises an important question about speakers' communication strategies: How do speakers balance efficiency with clarity?

Theoretical accounts generally agree that speakers are more likely to use pronouns when their intended referents are highly accessible, prominent, or in focus in discourse (e.g., Ariel, 1990; Chafe, 1976; Givón, 1983; Gundel et al., 1993). For example, entities introduced as grammatical subjects are more likely to be pronominalized than those introduced in less prominent syntactic positions (e.g., Arnold, 2001; Brennan, 1995; Fukumura & Van Gompel, 2010; Kaiser & Trueswell, 2008; Prasad & Strube, 2000; Rohde & Kehler, 2014; Stevenson et al., 1994). This prominence constraint helps reduce ambiguity since listeners typically resolve pronouns by linking them to previously mentioned subjects (e.g., Crawley et al., 1990; Frederiksen, 1981; Fukumura & Van Gompel, 2015; Kaiser & Trueswell, 2008). However, pronoun interpretation is influenced by factors beyond grammatical roles (e.g., Caramazza et al., 1977; Gordon & Hendrick, 1998; Sanford & Garrod, 1981). In particular, evidence suggests that comprehenders encounter difficulty when a pronoun's gender feature is compatible with multiple referents (e.g., Arnold et al., 2000; Garnham et al., 1995). This raises the question of whether and how speakers adjust their pronoun use in response to such referential ambiguity in language production.

According to research, speakers often fail to avoid certain types of ambiguity (e.g., Ferreira & Dell, 2000; Ferreira

[1]Faculty of Natural Sciences, University of Stirling, Stirling, UK
[2]Chengdu University of Technology, Chengdu, China
[3]University of Delhi, Delhi, India
[4]Indian Institute of Technology, Delhi, India

Corresponding author:
Kumiko Fukumura, Psychology, Faculty of Natural Sciences, University of Stirling, Stirling FK9 4LA, UK.
Email: kumiko.fukumura@stir.ac.uk

et al., 2005; Fukumura et al., 2022). Investigating noun phrase production, Ferreira et al. (2005) showed that speakers frequently avoid ambiguous bare nouns (e.g., *bat* in contexts with multiple flying bats) but are less likely to do so when the referents differ categorically (e.g., a flying bat vs. a baseball bat), where the ambiguity arises from homophony. When referents share the same semantic category (e.g., both are flying bats), referential ambiguity can be detected non-linguistically, since the shared semantic category implies that a single noun can refer to either entity. In contrast, when referents differ categorically (e.g., a flying mammal vs. a baseball bat), the ambiguity cannot be detected non-linguistically; it can only be recognized—if at all—once the noun is linguistically retrieved. Thus, Ferreira et al. suggested that speakers are more likely to avoid *non-linguistic ambiguity* than *linguistic ambiguity.*

In the case of pronouns, one straightforward strategy for avoiding ambiguity is to avoid using pronouns altogether when multiple potential referents are present. Supporting this idea, Arnold and Griffin (2007) observed that English speakers use fewer pronouns and more names when a preceding sentence introduces two referential candidates (e.g., *Mickey went for a walk with Daisy in the hills*) compared to one (e.g., *Mickey went for a walk in the hills*; see also Clancy, 1980). Similarly, Fukumura et al. (2010) found that the presence of a referential competitor in the visual context reduces pronoun use, compared to cases with no referential alternative—even when the competitor's gender is incompatible with the pronoun. This effect was explained by suggesting that the presence of a referential alternative in the referential context reduces the target referent's accessibility, leading to fewer pronouns being used.

A more contentious issue is whether speakers avoid pronouns when the pronoun's features (e.g., gender) are compatible with multiple referents. Evidence from English and French—languages in which third-person singular pronouns encode the referent's gender—suggests that they do, at least when referring to human entities (e.g., Arnold et al., 2000; Fukumura et al., 2010, 2011; Fukumura et al., 2022; Karmiloff-Smith, 1985). For instance, English speakers are less likely to use *he* for a male referent when another male referent is also present in the context. According to the *ambiguity avoidance account*, this effect reflects speakers' tendency to avoid pronouns whose gender features may create referential ambiguity. However, recent findings suggest that this tendency may not extend to inanimate objects with grammatical gender. For example, Fukumura et al. (2022) found that participants used the masculine pronoun *il* for *le pain* ("the bread") equally often, regardless of whether the context included another masculine noun (*le tournesol*, "sunflower"), where *il* was gender-ambiguous, or a feminine noun (*la rose*), where *il* was gender-unambiguous.

Grammatical gender is typically semantically arbitrary (Corbett, 1991) and can only be computed linguistically from the gender feature of the noun introducing the referent. In contrast, the gender information of human referents can be determined non-linguistically through various sources; from the inherent meanings of nouns (e.g., *king* refers to a male), from gender stereotypes associated with certain professions (e.g., *sailor* is typically associated with males; Misersky et al., 2014), or from the referents' physical characteristics such as appearance and clothing. In Fukumura et al. (2022), French speakers reduced pronoun use when human referents shared the same gender, consistent with patterns observed in English. Therefore, the absence of a gender effect for inanimate referents with grammatical gender suggests that speakers are not sensitive to pronoun gender ambiguity when it can only be determined linguistically rather than non-linguistically.

In fact, research suggests that gender congruence between human referents can reduce pronoun use, even when the pronouns themselves lack gender marking. Fukumura et al. (2013) found that Finnish speakers use the Finnish pronoun *hän* less frequently when multiple referents share the same gender. This effect cannot be attributed to gender ambiguity avoidance, since *hän* is non-gendered and can refer to either a male or a female. Thus, the effect of human gender congruence on pronoun usage, found in English and French, may not be solely due to gender ambiguity, as the effect also arises with non-gendered pronouns. One explanation is that shared gender increases referent similarity, which in turn reduces referent accessibility and leads speakers to use fewer pronouns.

Indeed, studies have shown that pronoun use is sensitive to the similarity between potential referents. In Fukumura et al. (2011), speakers describing actions used fewer pronouns (e.g., *he got off the horse*) when a referential competitor was in the same situation as the intended referent (e.g., a pilot sitting on a horse), compared to when the competitor was in a different situation (e.g., a pilot standing on the ground). Similarly, animacy congruence has been shown to affect pronoun usage, with English speakers using fewer pronouns when the referential context contains two entities with the same animacy (e.g., both humans, as in *The rowers overtook the swimmers*) compared to contexts where the entities have different animacy (e.g., *The rowers overtook the boats*) (Fukumura & Van Gompel, 2011).

These results led to the proposal that pronoun usage is influenced by the referents' non-linguistic similarity (Fukumura et al., 2011, 2013; Fukumura et al., 2021; 2022). Specifically, according to the *non-linguistic competition account*, when speakers initiate reference production, they activate the referent's non-linguistic features, such as animacy, gender, or situational properties, to identify the referent. When referential alternatives share any of these features, the shared features activate both the intended referent and its competitor. This leads to increased competition between them, prompting speakers to use more explicit expressions, such as nouns, to resolve it. On this view, human gender congruence decreases the use of gendered pronouns because the gender of a human referent is a non-linguistic feature. This gender information may be activated regardless of whether the pronoun carries gender marking, potentially causing

competition when referential candidates share the same gender by increasing their non-linguistic similarity.

The aim of the current study was to examine the generalizability of this account. While the non-linguistic competition account explains pronoun usage in some languages such as English, French, and Finnish, different languages employ distinct pronoun systems. For example, languages such as Mandarin Chinese, Italian, and Spanish allow subjects or objects to be omitted entirely (commonly referred to as *null* pronouns) (Alonso-Ovalle et al., 2002; Carminati, 2002; Di Eugenio, 1998; Huang, 1984; C. N. Li & Thompson, 1979; Lozano, 2016; Neeleman & Szendrői, 2007; Tomioka, 2003). Fukumura et al. (2021) investigated Italian, where null pronouns are favored over overt pronouns for subject referents (e.g., Carminati, 2002; Contemori & Di Domenico, 2021, 2023; Filiaci et al., 2014). Their study found that the use of null pronouns (relative to nouns) is unaffected by whether the human referential candidates share the same gender. This held true regardless of whether gender congruence between referents rendered the verb's gender agreement ambiguous, challenging the view that null pronouns are licensed by the presence of agreement features that uniquely identify the referent (cf. Rizzi, 1986).

Furthermore, in spoken Chinese, the third-person pronoun *ta* is non-gendered, serving the role of *she*, *he*, or *it* in English. However, in written form, *ta* becomes gender-marked (她 meaning "she" vs. 他 meaning "he"). Hwang (2021) found that gender congruence between human referents reduced the use of the non-gendered pronoun *ta* in written Mandarin, but not in spoken Mandarin. The author argued that gender affects pronoun usage only through a gender ambiguity avoidance strategy, rather than through similarity-based competition. Hwang (2018) found similar results in Cantonese, where the non-gendered pronoun *keoi* was used more often when the referent had been introduced as the subject. Pronoun use was unaffected by the gender of other characters, showing no evidence of similarity-based competition. Although *keoi* was reported to have occurred less frequently in human-human contexts (e.g., "the actor kissed the actress") than in human-inanimate ones (e.g., "the actor kissed the ring"), this may not necessarily reflect an animacy-based competition effect (Fukumura & Van Gompel, 2011). This is because in their free sentence completion task, participants could refer to any entity. The researcher noted that responses were excluded when the referent of *keoi* could not be determined, and such ambiguity may have been more frequent in human-human contexts (e.g., *keoi zau faan nguk kei laa* "he/she went home").

While Hwang (2018, 2021) focused on overt pronouns and nouns in Cantonese and Mandarin Chinese, null pronouns (or argument omissions) are common in both languages (R. Li, 2016, 2020; Zhang & Kwon, 2022). Zhou (2021) and Zhou et al. (2023) examined how semantic similarity between two human referential candidates influenced the use of null pronouns, overt pronouns, and repeated nouns in Mandarin Chinese. In one experiment involving candidates who shared the same gender, null pronouns decreased and repeated nouns increased when the candidates were semantically more similar (e.g., killer vs. murderer rather than killer vs. proofreader). No reliable effect of semantic similarity was found on overt pronouns. In another experiment with gender-different candidates, similar patterns emerged. Importantly, overt pronouns were significantly less frequent in the experiment where candidates shared the same gender (Zhou, 2021).

Given these findings, the current study addressed two key issues. First, we investigated whether non-gendered overt pronouns can be influenced by gender-based competition in languages other than Finnish. Findings from Italian null pronouns (Fukumura et al., 2021) and overt pronouns in Mandarin (Hwang, 2021; but see Zhou, 2021) have challenged this possibility. We therefore examined pronoun usage in Hindi. In Hindi, pronouns used for third-person reference do not distinguish gender or animacy (Kachru, 2006). Specifically, the pronoun वह (*vah*) functions as a gender-neutral personal pronoun, equivalent to *he*, *she*, or *it* in English. Although null pronouns are possible in Hindi, their usage patterns may differ from those observed in languages such as Italian and Chinese (Prasad, 2000). Additionally, unlike Finnish and Mandarin, Hindi features a rich agreement system, similar to Italian. This allowed us to investigate whether gender ambiguity in verb agreement influences pronoun usage. Second, we aimed to reconcile the conflicting findings of Hwang (2021) and Zhou (2021). Although Hwang (2021) failed to find a gender-based competition effect, the author suggested that gender information might not have been salient enough in the study, given the absence of a visual context—unlike in the Finnish experiment by Fukumura et al. (2013). This explanation is supported by findings from Zhou (2021), who employed a visual context and observed evidence of gender-based competition.

Experiment 1

Corpus studies suggest that Hindi pronoun use is influenced more by antecedents' grammatical roles than by their order of mention (Prasad & Strube, 2000). However, to our knowledge, this observation has not been experimentally corroborated. Therefore, Experiment 1 manipulated both the grammatical roles of referential candidates and their gender congruence, using the same paradigm as Fukumura et al. (2021; 2022). Participants first read a context sentence (1), which introduced two referential candidates visually displayed on a computer screen. Example displays are shown in Figure 1. In Figure 1A, the candidates were of the same gender, featuring two female characters (policewoman and queen). In Figure 1B, the candidates had different genders, featuring one female (policewoman) and one male character (king). The grammatical role of each character was varied, with each introduced as either the subject (1a) or the postpositional object (1b). After reading sentence (1) aloud, participants saw a display change, in which the target

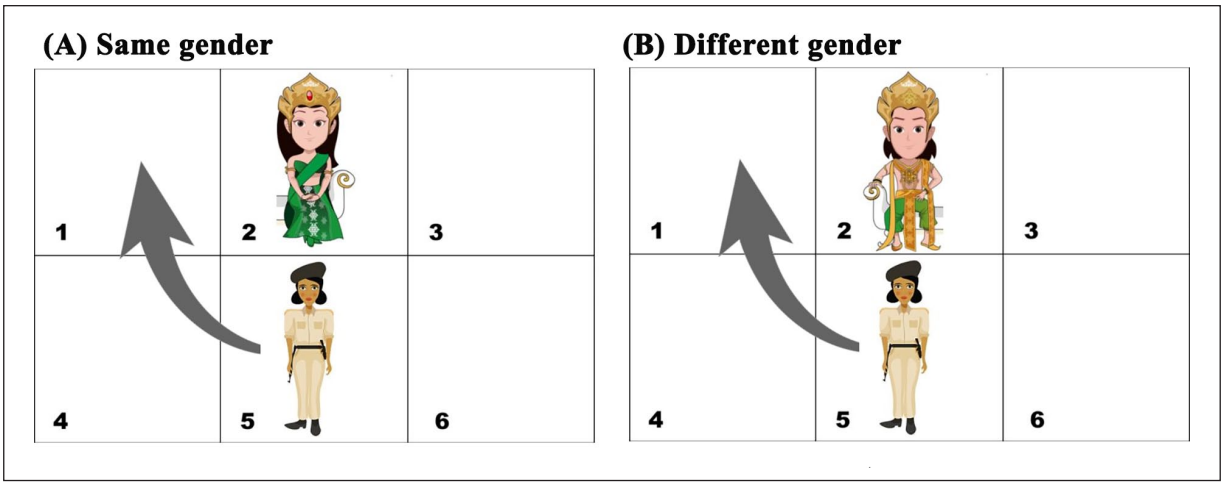

Figure 1. Example displays for Experiment 1.

character moved to a new location (e.g., the policewoman moving from number 5 to number 1, as indicated by the arrows in Figure 1). Participants were then prompted to produce a sentence describing the target character's movement (2). We analyzed their choice of referring expression for the target.

(1) Context sentence

a. Subject antecedent

रानी/राजा	के नीचे वाली	पुलिसवाली	नंबर	५	पर	है।
raanii/raajaa	ke neeche waalii	puliswaalii	nambar	5	par	hai
Queen/king	below	policewoman	number	5	on	is.

"The policewoman below the queen/king is on number 5."

b. Postpositional object antecedent

पुलिसवाली	के ऊपर वाली/वाला	रानी/राजा	नंबर	२	पर	है।
puliswaalii	ke uupar waalii/waalaa	raanii/raaja	nambar	2	par	hai.
Policewoman	above	queen/king	number	2	on	is.

"The queen/king above the policewoman is on number 2."

(2) Target response

a. Repeated noun

अब पुलिसवाली नंबर १ पर है।
ab puliswaalii nambar 1 par hai.
"Now policewoman is on number 1."

b. Demonstrative-Noun[1]

अब वह पुलिसवाली नंबर १ पर है।
ab vah puliswaalii nambar 1 par hai.
"Now that policewoman is on number 1."

c. Pronoun

अब वह नंबर १ पर है।
ab vah nambar 1 par hai.
"Now that is on number 1."

d. Null Pronoun

अब नंबर १ पर है।
ab nambar 1 par hai.
"Now Ø is on number 1."

Gender effect

If gender-based competition affects the use of the non-gendered Hindi pronoun *vah*, participants should produce fewer pronoun-only responses (2c) in the same-gender condition than in the different-gender condition. In contrast, more explicit referring expressions—such as repeated nouns (2a) and demonstrative-noun phrases (2b)—are expected to increase in the same-gender condition. In the latter (2b), *vah* precedes a noun and is often translated as *that*; we refer to these as *demonstrative-noun* responses.

Grammatical role

Additionally, if the use of Hindi overt pronouns is influenced by the grammatical role of the antecedent (Prasad & Strube, 2000), then pronoun responses should occur more frequently following subject antecedents (1a) than postpositional object antecedents (1b). If Hindi null pronouns are used similarly as in languages like Italian and Spanish, they should occur more often after subject antecedents than after postpositional object antecedents. While demonstrative-nouns in Hindi are often said to convey *definiteness* (Kachru, 2006)—that is, whether the referent is given or identifiable in context—how they are affected by antecedent grammatical role has remained unclear. Theories of reference generally treat demonstrative-nouns as expressions of intermediate accessibility or prominence (e.g., Ariel, 1990; Gundel et al., 1993). According to Gundel et al. (1993), demonstrative-nouns tend to be used when the referent is 'activated' or 'familiar', rather than merely 'uniquely identifiable'. If this account holds for Hindi, we would expect more demonstrative-noun responses following subject antecedents than following postpositional object referents, since subject referents are typically more prominent or accessible.

Method

Participants. Thirty-two participants were recruited from students at the Indian Institute of Technology, Delhi, the University of Delhi, and Jawaharlal Nehru University. All participants were native speakers of Hindi and had no visual impairments. Participants were between 18 and 30 years old and provided informed consent prior to the experiment.

Procedure and materials. We used the same procedure as in Fukumura et al. (2021; 2022), and informed consent was obtained before the experiment. The participant took the role of the speaker, while a confederate acted as the addressee. During the experiment, both the speaker and the addressee sat side by side at a table, each facing a computer monitor. A shade was attached to the speaker's screen to prevent the addressee from seeing it. In each trial, both participants were presented with a 2×3 visual display on their monitors. The speaker pressed the down-arrow key on the keyboard, which caused the context sentence (1) to appear on their screen but not on the addressee's screen. The speaker read this sentence aloud to the addressee and then pressed the down-arrow key again, triggering the presentation of a second display in which the target referent changed location on the speaker's screen. The speaker described this change by producing a new sentence starting with *ab* ("Now") (2). The addressee was instructed to act out the speaker's descriptions by pointing at the mentioned object and its new location. The speaker pressed the "yes" key on the keyboard if the addressee correctly understood the description or the "no" key if the speaker believed the addressee's response was incorrect. We used E-Prime to present stimuli and record participants' responses. Each session lasted approximately 30 min, with a 1 to 2-min break in the middle.

We constructed 40 experimental items, each comprising 2×3 visual displays and context sentences. Figure 1 shows example visual displays where gender congruence was manipulated. Figure 1A represents the *same-gender condition*, where the policewoman is paired with a queen. Figure 1B represents the *different-gender condition*, where the policewoman is paired with a king. The competitor pairs had comparable roles or professions (e.g., king and queen), with each character appearing as both a same-gender competitor and a different-gender competitor across different items, thereby counterbalancing the attributes of the competitors across conditions.

The context sentences were constructed in Hindi and followed a structure similar to that used by Fukumura et al. (2021; 2022), with one character appearing in the subject position, while the other was introduced in a modifier phrase.[2] In the *subject antecedent condition* (1), the target character (the policewoman) was introduced in the subject position, while the competitor character (the queen or king) was mentioned in the postpositional object position. In the *postpositional object antecedent condition* (2), the target character (the policewoman) was mentioned as the postpositional object in the prenominal modifier, while the competitor (the queen or the king) was mentioned in the subject position.

In each visual display, the target character was accompanied by a female (Figure 1A) or male competitor (Figure 1B). The competitor pairs were selected to be clearly distinguishable from the target referent (i.e., the policewoman in Figure 1 is dissimilar from the queen or the king) and to share comparable professions or roles (e.g., queen vs. king in Figure 1). In total, there were 40 target human characters; 20 male and 20 female. Each character also appeared once as a same-gender competitor and once as a different-gender competitor across two different items. This ensured that the lexical properties (e.g., frequency, length) of the competitor characters were controlled across conditions. For the prenominal modifiers, we used the postpositions *ke uupar* ("above"), *ke niiche* ("below"), or *ke saath meM* ("next to"). The conditions were distributed such that each participant saw each character once as a target and once as a competitor.

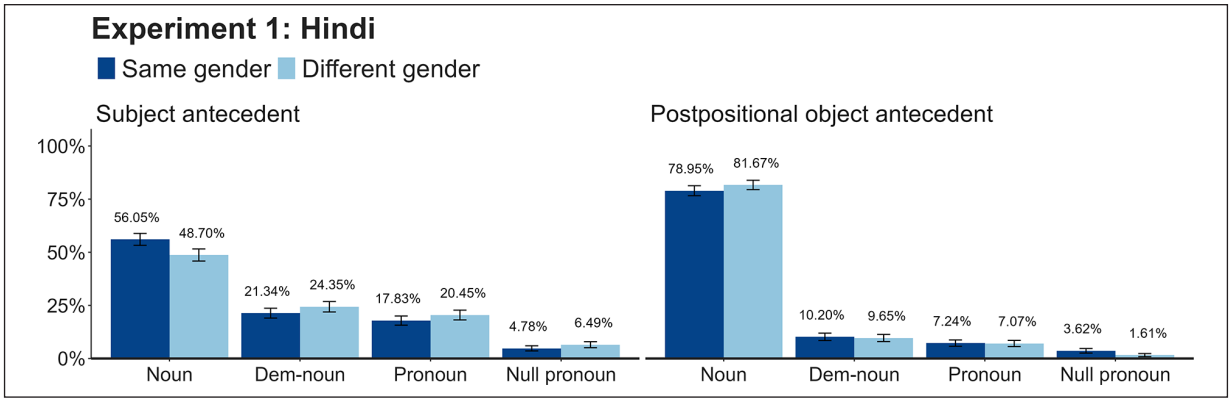

Figure 2. Percentages of referring expressions in Experiment 1.
Note. Dem = demonstrative.

Additionally, we constructed 80 filler items, closely modeled on those used in Fukumura et al. (2021; 2022). The context sentences in the filler items had structures different from those in the experimental items, with variation in the position and grammatical function of the target referent. About 50% of filler trials introduced the target referent as the subject, while 25% introduced it in a non-first-mentioned position. The number of referential candidates and their animacy were also varied. For example, some filler items (approximately 30%) included only one human character or object in the display (e.g., नंबर 6 पर एक तांत्रिक है ।, "On number 6, (there) is a pagan-priest," or सीडी नंबर 5 पर है ।, "The CD is on number 5"). Other filler items (approximately 25%) included more than two referential candidates (e.g., नंबर 6 पर एक तांत्रिक है दुकानदार नंबर 3 पर है, फुटबॉल खिलाड़ी और अमरूद के ऊपर ।, "The shopkeeper is on number 3, above the football player and the guava").

Design. We manipulated the antecedent's grammatical role (subject vs. object), the referents' gender congruence (same vs. different). In addition, we counterbalanced the positions of the target and competior characters. These were all manipulated as within-subject/within-item variables, resulting in the creation of eight lists. Together with 80 filler items, 40 experimental items were distributed in a fixed quasi-random order, subject to the constraint that the same target or competitor gender should not occur in more than three experimental trials consecutively and there should be at least one filler item between the experimental trials. Collapsing across the counterbalancing variable of the characters' position, each list contained 10 items in each condition, with only one version of each item. Four participants were randomly assigned to each list.

Scoring. Responses (n = 1,280) were scored based on whether the subject of the response sentence was realized using repeated nouns ($n = 820$), demonstrative-nouns ($n = 203$), pronouns ($n = 163$), or null pronouns ($n = 51$). We excluded cases where participants mistakenly referred to the competitor as the subject or referred to the competitor's location first ($n = 24$); used nouns other than those provided in the context sentences ($n = 8$); misidentified the character's gender ($n = 8$); altered their responses (e.g., from a pronoun to a demonstrative noun) ($n = 1$); or did not produce a new sentence ($n = 2$). In total, 43 trials (3.4% of responses) were excluded from analysis.

Results and discussion

Figure 2 presents the distribution of four response types: repeated nouns, demonstrative-nouns, overt pronouns, and null pronouns.

The analyses focused on how the grammatical role of the antecedent (subject vs. object) and the gender congruence of the referents (same vs. different gender) influenced the frequency of each expression. To this end, we analyzed the rate of each response category relative to all alternatives, following previous studies that examined the rates of more than two referring expressions (e.g., Zhou et al., 2023) or structures (e.g., Bernolet et al., 2009). Although this required multiple dichotomizations of the same dataset, each analysis addressed a distinct research question—specifically, how the predictors influenced the occurrence of each referring expression relative to the others. While some studies employed multinomial analyses (e.g., Hwang et al., 2022), such analyses rely on a reference level and would be uninformative in this context. These models would show how gender influences the rate of null pronouns compared to repeated nouns, or overt pronouns compared to repeated nouns—assuming repeated nouns are chosen as the reference level. Critically, however, the rate of null or overt pronouns depends on all available alternatives, not just repeated nouns—precisely because the language permits a non-dichotomous choice.

Thus, logit mixed-effects analyses were conducted on the choice of each referring expression (coded as 1) relative to all alternatives (coded as 0), using the lme4 package (Bates et al., 2015) implemented in R 4.2.1 (R Core Team, 2022). Unless stated otherwise, the variables were

sum-coded to reduce collinearity between the predictor variables (Baayen et al., 2008). Specifically, the analysis in Experiment 1 included gender congruence (same vs. different), which were coded as −0.5 and 0.5, respectively, and antecedent role (subject vs. postpositional object), coded as −0.5 and 0.5, respectively, and their interaction as fixed effects. Additionally, by-participant and by-item random intercepts, along with random slopes for all relevant fixed factors (Barr et al., 2013), were included. Correlations between random effects were suppressed to avoid overparameterization (Bates et al., 2015, 2018; Kliegl, 2014; Singmann & Kellen, 2020). In cases where the model failed to converge, the random slope parameter with zero or near-zero variance was removed to achieve convergence (Gelman & Hill, 2007).

Table 1 summarizes the results. First, repeated nouns occurred less frequently when the antecedent noun was the grammatical subject (52.4%, *SE* = 2.0%) than when it was a postpositional object (80.3%, *SE* = 1.6%). This effect interacted with gender congruence: When the antecedent was the subject, speakers used significantly more repeated nouns (relative to all alternatives involving pronouns or null pronouns) in the same-gender condition compared to the different-gender condition. No such effect was observed when the antecedent was a postpositional object. Second, the antecedent's grammatical role also influenced both demonstrative noun and (overt) pronoun responses. Demonstrative-noun responses were more frequent for subject antecedents (22.8%, *SE* = 1.7%) than for postpositional object antecedents (9.9%, *SE* = 1.2%). Similarly, pronoun responses were more frequent for subjects (19.1%, *SE* = 1.6%) than for postpositional object antecedents (7.2%, *SE* = 1.0%). Gender congruence did not significantly reduce the rates of either expression relative to all alternatives, despite numerical trends in that direction. Finally, a significant interaction between gender congruence and grammatical role was observed in the rate of null pronouns. This interaction reflected a non-significant tendency for gender congruence to slightly reduce null pronouns for subject antecedents while slightly increasing them for postpositional object antecedents.

In sum, when the antecedent was the grammatical subject, gender congruence between referential candidates led to a higher frequency of repeated nouns relative to pronoun-based alternatives. This finding provided support for the non-linguistic competition account, demonstrating the gender effect on the use of non-gendered Hindi pronouns. Additionally, pronoun responses and demonstrative-noun responses were influenced by the antecedent's grammatical role, occurring more frequently when the antecedent was the grammatical subject rather than a postpositional object. This suggests that speakers can use either expression to signal topic continuity in preceding discourse—specifically, the referent's continuation as the subject. To our knowledge, this is the first experimental evidence demonstrating the role of the antecedent's grammatical function in the use of both overt pronouns and demonstrative-noun responses in Hindi.

Table 1. Analyses of choice of expressions in Experiment 1.

Fixed effect	Estimate	*SE*	*z*	*p*
Repeated noun				
(Intercept)	1.50	0.54	2.80	.005
Gender congruence	−0.12	0.20	−0.59	.557
Grammatical role	2.76	0.38	7.24	<.001
Gender × Grammatical Role	1.01	0.46	2.18	.029
Subject antecedent				
(Intercept)	0.32	0.64	0.50	.617
Gender congruence	−0.64	0.27	−2.38	.018
Postpositional object antecedent				
(Intercept)	2.78	0.58	4.78	<.001
Gender congruence	0.35	0.34	1.05	.294
Demonstrative-noun				
(Intercept)	−3.14	0.47	−6.69	<.001
Gender congruence	0.10	0.22	0.44	.662
Grammatical role	−1.33	0.52	−2.57	.010
Gender × Grammatical Role	−0.35	0.44	−0.80	.421
Pronoun				
(Intercept)	−5.82	1.34	−4.34	<.001
Gender congruence	0.15	0.27	0.56	.577
Grammatical role	−2.70	0.49	−5.48	<.001
Gender × Grammatical Role	−0.36	0.56	−0.64	.522
Null pronoun				
(Intercept)	−9.65	1.85	−5.22	<.001
Gender congruence	−0.21	0.41	−0.53	.600
Grammatical role	−1.17	0.98	−1.19	.233
Gender × Grammatical Role	−1.88	0.81	−2.31	.021
Subject antecedent				
(Intercept)	−8.57	1.86	−4.61	<.001
Gender congruence	0.83	0.63	1.32	.186
Object antecedent				
(Intercept)	−4.29	0.54	−7.99	<.001
Gender congruence	−0.82	0.67	−1.23	.220

Experiment 2

Although Experiment 1 showed that gender congruence between referential candidates increased the use of repeated nouns, it did not significantly reduce the use of (overt) pronouns. One possible explanation is that other pronoun-based expressions also decreased in frequency under the same condition, albeit not significantly. Another possibility is that the lack of a significant reduction in pronoun use was because the referent was introduced as a postpositional object in half of the trials; pronouns were generally disfavored for object antecedents, regardless of gender congruence. Thus, in Experiment 2, the referent was always

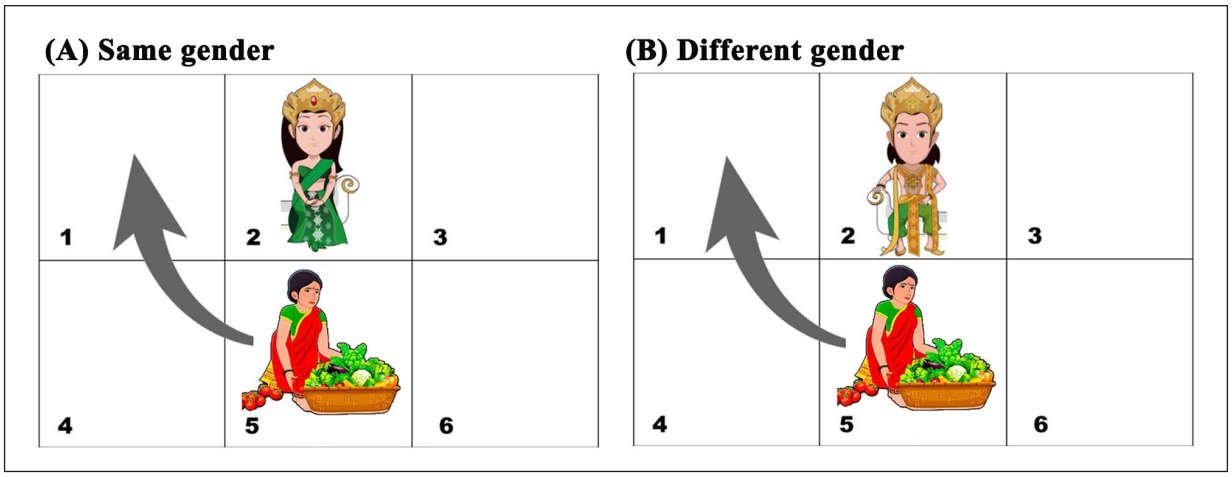

Figure 3. Example displays for Experiment 2.
Note. Image for Vegetable Seller: From Telugu Stories for Kids, by Koo Koo TV, 2019, YouTube. https://youtu.be/fqmBNVi79hY?si=-3DOcjbgB-tUsatCE.

introduced as the grammatical subject in the context sentence, aligning with prior studies that reported gender effects on pronoun usage. This also allowed us to examine whether Hindi pronoun usage is influenced by gender ambiguity in verb agreement.

In Hindi, verbs in certain constructions must agree in gender with the subject if the subject is in the nominative case. For example, the verb participle रही (*rahii*) in the present progressive tense, as in अब वह नंबर 1 पर बैठ रही है।, meaning "Now it (she) is sitting at number 1", explicitly marks the gender of the female vegetable seller in Figure 3. The gender information marked on the verb participle could uniquely identify her in Figure 3B (different-gender condition), where the competitor is a king, but not in Figure 3A (same-gender condition), where the competitor is a queen. However, the seller's movement in the display could also be described using a copular construction in the simple present tense, as in अब वह नंबर 1 पर है, meaning "Now it (she) is at number 1", which does not mark gender. In this case, the utterance remains referentially ambiguous, regardless of gender conditions. In Experiment 2, we investigated whether the effect of gender on pronoun usage varies depending on whether speakers use gender-marked verb forms since speakers select more explicit referring expressions when verb gender agreement fails to uniquely identify the referent.

If Hindi pronoun usage is influenced by both gender ambiguity avoidance in verb agreement and gender-based competition, then speakers should be more likely to avoid pronouns in the same-gender condition (compared to the different-gender condition) when verb forms are gender-marked than when they are not. When verbs are gender-marked, speakers may choose referring expressions that reflect both gender-based competition and ambiguity avoidance. When verbs are not gender-marked, any effect of gender congruence would primarily reflect similarity-based competition between referents. On the other hand, if similarity-based competition is the primary driver of gender effects, we would expect fewer pronouns in the same-gender condition than in the different-gender condition, regardless of verb agreement. To test this, we asked participants to describe the displays using either the present progressive tense (gender-marked verb condition) or the copular construction (non-gendered verb condition).

Method

Participants. A further 32 participants were recruited from the same population as before.

Procedure and materials. Unlike in Experiment 1, the referent in Experiment 2 was always introduced as the grammatical subject in the context sentence. In addition to the referents' gender congruence, we manipulated the presence or absence of gender agreement in participants' responses across blocks. In one block (verb-agreement condition), participants were instructed to produce sentences in the progressive form by describing the character's state or action, which included gender agreement on the verb participle. In the other block (non-agreement condition), participants were asked to produce responses ending with *par hai* ("is on"), resulting in the use of the present *be*-copula, which conveyed no gender information about the referent. While participants were asked to keep the sentence ending uniform, they were free to frame the rest of the sentence as they wished. The order of blocks was counterbalanced: Half of the participants completed the gender agreement block first, followed by the non-agreement block, while the other half completed the blocks in the reverse order.

For experimental items, most states or actions were described using verbs such as "standing", "sitting", or "lying". For filler items, different action verbs were used depending on the characters and their perceived actions

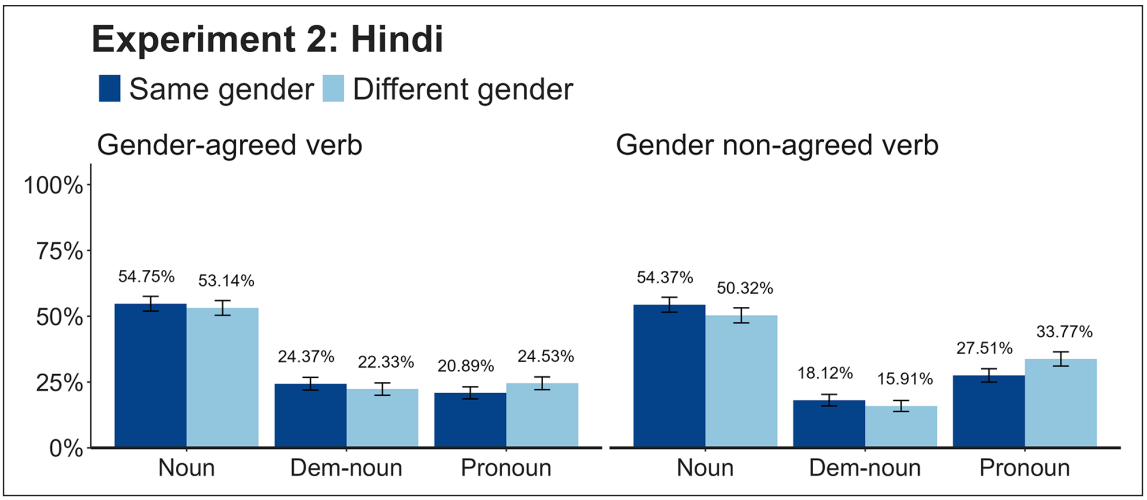

Figure 4. Percentages of referring expressions in Experiment 2.
Note. Dem = demonstrative.

(e.g., a cricket player playing, a bee flying, or inanimate objects like a CD or candy lying). To ensure that mentioning the character's state or action did not itself contribute to discrimination, the combination of characters in some experimental items—where the target character's state or action differed from its referential alternative—was adjusted so that the physical states of the target and competitor characters aligned. For example, the policewoman in Figure 1 was replaced with the vegetable seller in Figure 3, so that all characters were sitting, making the mention of "sitting" unable to distinguish the vegetable seller from either the queen or the king. The filler items were the same as in Experiment 1.

Design. We used a 2 (*Gender Congruence*: Same Gender vs. Different Gender) × 2 (*Agreement Presence*: Present vs. Absent) repeated measures design. In addition, we counterbalanced the order of blocks. This resulted in the creation of eight lists, and four participants were randomly assigned to each list.

Scoring. We scored responses as before, categorizing them into repeated nouns (n=665), demonstrative-nouns (n=253), pronouns (n=333), and null pronouns (n=17). Trials with null pronouns were excluded due to their low frequency and the fact that they were produced by only a few participants. Additionally, cases were excluded if participants: changed response types (n=5), mentioned the competitor instead of the target (n=5), produced nouns other than those given in the context sentence (n=2), or used copular constructions (i.e., failed to produce gender agreement on the verb) when instructed to use the present progressive tense (n=19). Responses using a different tense (e.g., present perfect) to achieve verb gender agreement were included (n=35). In total, 48 trials (3.8%) were excluded, leaving 654 noun responses, 331 pronoun responses, and 247 demonstrative-noun responses for analysis.

Table 2. Analyses of Hindi choice of expressions Experiment 2.

Fixed effect	Estimate	*SE*	*z*	*p*
Repeated nouns				
(Intercept)	0.47	0.49	0.96	.339
Gender congruence	−0.24	0.16	−1.52	.128
Agreement presence	−0.14	0.43	−0.34	.736
Gender × Agreement	−0.25	0.32	−0.79	.432
Demonstrative-nouns				
(Intercept)	−2.91	0.54	−5.38	<.001
Gender congruence	−0.23	0.19	−1.19	.235
Agreement presence	−0.70	0.42	−1.65	.099
Gender × Agreement	−0.07	0.39	−0.17	.865
Pronouns				
(Intercept)	−2.67	0.63	−4.25	<.001
Gender congruence	0.61	0.20	2.98	.003
Agreement presence	0.76	0.72	1.05	.293
Gender × Agreement	0.35	0.41	0.84	.403

Results and discussion

Figure 4 reports the frequencies of target referring expressions. As mentioned in the scoring section, the rate of null pronouns was too low (n=17) for meaningful analysis. The focus was instead on the rates of repeated nouns, demonstrative-nouns, and pronouns as a function of the referents' gender congruence and the presence of gender agreement (present vs. absent, coded as −0.5 and 0.5, respectively) on the verb. The results are summarized in Table 2.

A main effect of gender congruence was found in the rate of pronoun responses, showing fewer pronouns in the same-gender condition (24.4%, SE=1.7%) than in the different-gender condition (29.3%, SE=1.8%), with no interaction with verb agreement. If anything, the gender effect was numerically smaller when the verb was gender-marked (3.5%)—when referents' gender congruence affected verb

agreement ambiguity—than when the verb was not gender-marked (6.3%). This provided no evidence that the gender effect on Hindi non-gendered pronouns is affected by gender ambiguity avoidance in the verb agreement. Unlike in Experiment 1, gender congruence did not significantly increase the probability of repeated nouns relative to all alternatives. As shown in Figure 4, gender congruence in the current experiment numerically increased demonstrative-noun responses while reducing pronoun responses. This pattern differs from Experiment 1, where demonstrative-noun responses behaved more like pronoun responses (both decreasing following gender congruence for subject antecedents), which led to a reliable gender effect on repeated nouns relative to these alternatives.

Taken together, the results support the non-linguistic competition account, showing that speakers of Hindi reduce their use of non-gendered pronouns in favor of more explicit referring expressions when gender congruence increases the similarity between human characters. This effect occurred regardless of whether gender congruence introduced ambiguity in the verb's gender agreement.

Experiment 3

The effect of gender-based competition on the use of Hindi non-gendered pronouns indicates that these pronouns are sensitive to non-linguistic competition. Experiments 3 and 4 aimed to determine whether Mandarin Chinese overt pronouns show similar sensitivity. As discussed earlier, Hwang's (2021) sentence-completion study found that gender congruence between referents reduced the use of overt pronouns only in the written modality, where pronouns are gendered. Although Zhou's (2021) between-experiment comparison suggested a gender-based competition effect in spoken Mandarin, semantic similarity between referents (e.g., *killer* and *murderer*) reduced the rate of null pronouns, but not overt ones (Zhou et al., 2023). This suggests that Mandarin overt pronouns may not be sensitive to similarity-based competition. Experiment 3 thus investigated gender-based competition in Mandarin Chinese by manipulating both situational similarity and gender congruence between referents. If overt pronouns in Mandarin Chinese are insensitive to non-linguistic competition, then neither factor should influence their usage.

As previously discussed, the effect of situational similarity on pronoun usage was first observed in a storytelling task—where the addressee acted out the speaker's description—in both English and Finnish (Fukumura et al., 2011, 2013). More recently, Fukumura et al. (2021; 2022) showed an effect of situation-based competition using a different paradigm, which we adopted in the present study. See Figure 5 for sample displays. In the *two-box condition* (a), red boxes around both characters indicated that either could move. In the *one-box condition* (b), only the target (e.g., the sailor) could move. Referential candidates were either same-gender (e.g., sailor–king) or different-gender (e.g., sailor–queen). Although this gender manipulation had no effect on Italian null pronouns—unlike in English or French—speakers used fewer null pronouns and more repeated nouns in the two-box condition than in the one-box condition, similar to patterns in English and French. This suggests that while Italian speakers are sensitive to non-linguistic competition arising from situational similarity, gender congruence does not induce competition in null pronoun use.

(3) Context sentence
国王/王后上方的水手在二号位。
Guowang/wanghou shangfang de shuishou zai erhaowei.
King/queen above sailor at[3] position 2
"The sailor above the king/queen is on number 2."

(4) Target sentence
a. Noun
现在水手在四号位。
Xianzai shuishou zai sihaowei.
Now sailor at position 4
"Now the sailor is on number 4."

b. Pronoun
现在他在四号位。
Xianzai ta zai sihaowei.
Now he .at position 4
"Now he is on number 4."

c. Null Pronoun
现在在四号位。
Xianzai Ø zai sihaowei.
Now Ø at position 4
"Now Ø is on number 4."

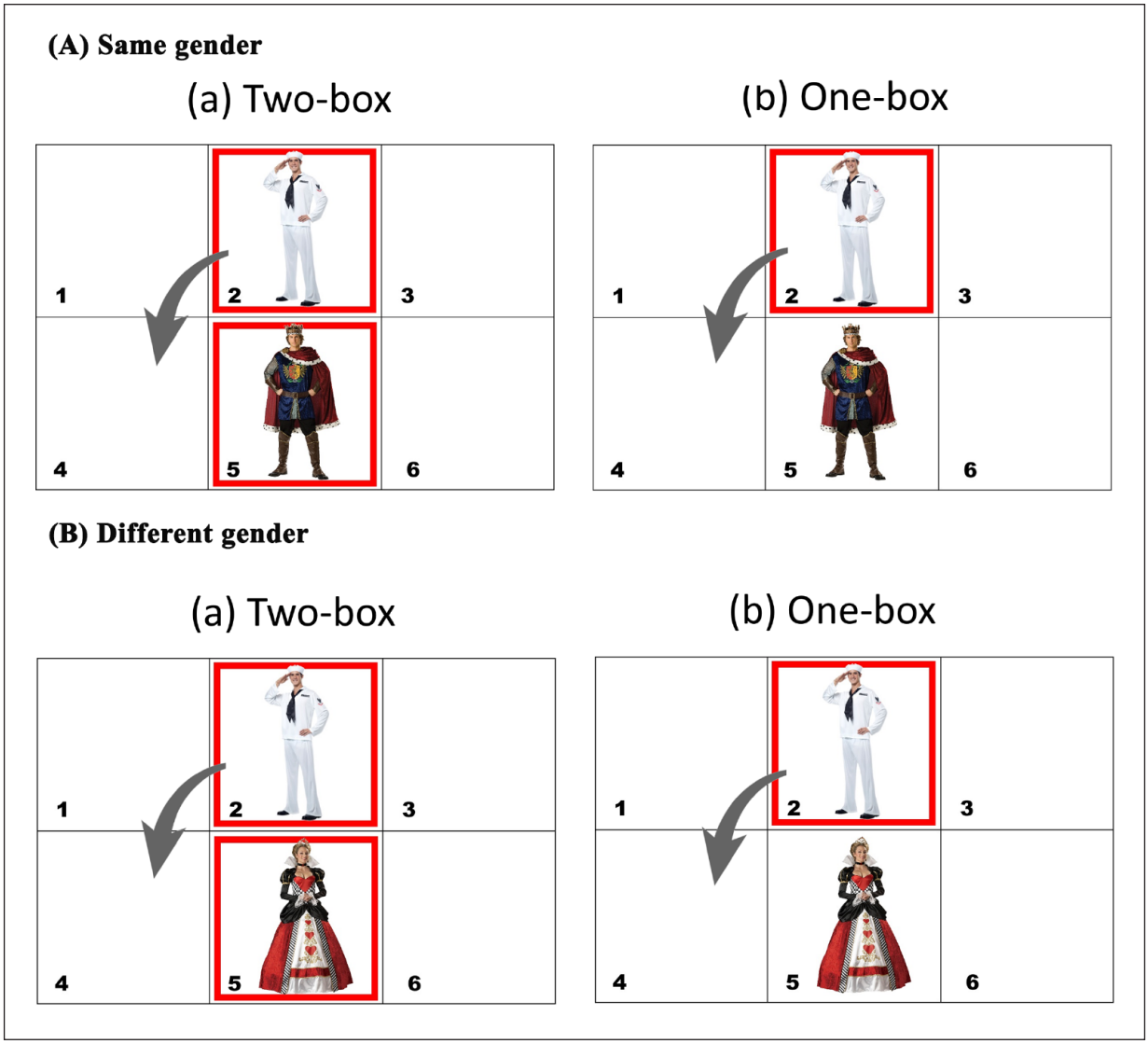

Figure 5. Example displays in Experiment 3.
Note. Sailor: © California Costume Collections, Inc. King & Queen: Costume images used by permission In Character Costumes. Division of Fun World Easter Unlimited Inc. All Rights Reserved.

If Mandarin overt pronouns are sensitive to non-linguistic competition but gender congruence does not affect this competition, we expect an effect of situational similarity, leading to reduced usage of overt pronouns in the two-box condition compared to the one-box condition, with no effect of gender congruence. If, instead, overt pronouns are sensitive to gender-based but not situation-based competition, we may observe the reverse pattern. Finally, if Mandarin overt pronouns are generally insensitive to similarity-based competition, neither factor should affect their use. Additionally, if Mandarin null pronouns are affected by similarity-based competion similarly to Italian null pronouns, we should find fewer null pronouns in the two-box than in the one-box condition, with no effect of gender congruence.

As in Experiments 1 and 2, the context sentence (3) linguistically introduced both the target referent (sailor) and the competitor (king or queen). Participants' responses were categorized as repeated nouns (4a), overt pronouns (4b), or null pronouns (4c), based on how they referenced the target while describing its change in location.

Method

Participants. We recruited 32 native speakers of Mandarin Chinese from the University of Stirling student community in exchange for cash. All the participants reported to be native Chinese speakers grown up in Mandarin-speaking families, aged between 18 and 30, with no dyslexia or visual impairments.

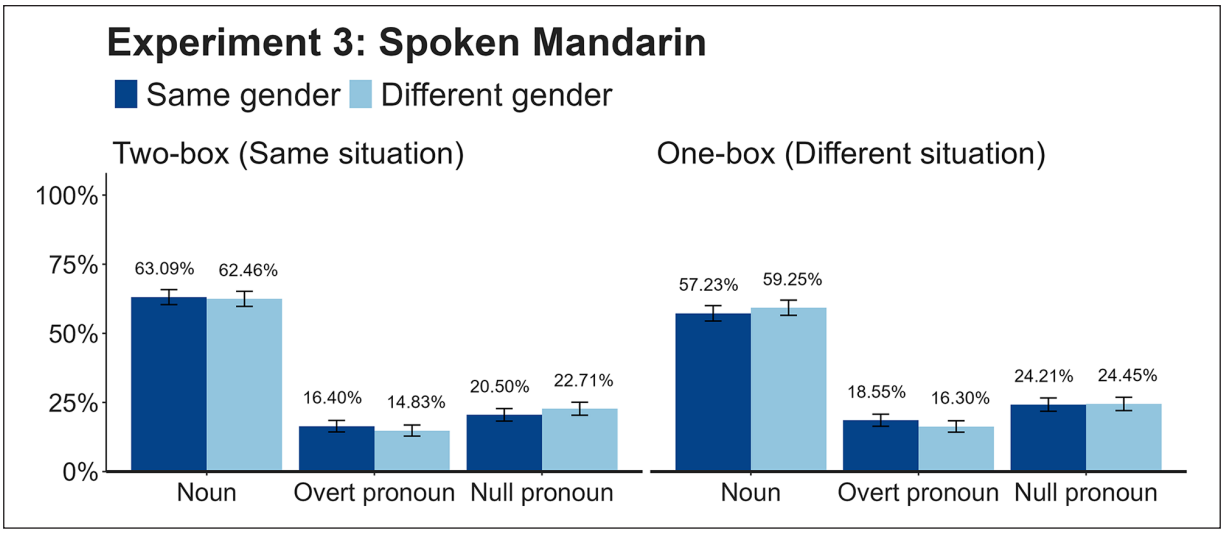

Figure 6. Percentages of referring expressions in Experiment 3.

Materials and procedure. The procedure was the same as in Experiments 1 and 2. Forty experimental items were adapted from Experiment 4 in Fukumura et al. (2021), with the context sentences translated into Mandarin Chinese. Each experimental item comprised 2 × 3 visual displays, which included photographs of two human characters (see Figure 5) and context sentences (3) that linguistically introduced the characters. The target referent (sailor) was introduced as the subject in the context sentence, while its competitor (king or queen) was mentioned in a prenominal phrase modifying the target. The spatial relations between the two entities were described using one of three phrases: *xiafang* ("below"), *shangfang* ("above"), or *pangbian* ("next to"). In each display, the target referent and its competitor were either of the same gender (e.g., *shuishou* "sailor" vs. *guowang* "king") (Figure 5A) or of different genders (e.g., *shuishou* "sailor" vs. *wanghou* "queen") (Figure 5B). In addition, the entities were either in the *two-box context* (i.e., *situationally congruent* as both were in the red box) (Figure 5A) or in the *one-box context* (i.e., *situationally incongruent* as only the target referent was in the red box) (Figure 5B). In 80 filler items, the syntactic structures of the context sentences, the syntactic positions of the target referents, and the number of displayed entities were varied.

Design. We used a 2 (*Situational Congruence*: Two-Box vs. One-Box) × 2 (*Gender Congruence*: Same vs. Different) repeated measures design. Along with the 80 fillers, the 40 experimental items were distributed across four lists, following quasi-randomized orders, with at least one filler item occurring between each experimental item. Each list contained 10 experimental items of each condition, and only one version of each item was presented in a given list.

Scoring. Responses (n = 1,280 in total) were scored based on whether the participants produced a repeated noun (n = 769), an overt pronoun (n = 210), or a null pronoun (n = 292) to refer to the target referent. Responses were excluded if they contained changes in referring expressions (n = 3, e.g., *Xianzai Ø zai . . . nvjueshi zai liuhao wei* "Now Ø is on . . . the countess is on Number 6"); participants used other nouns than those used in the context sentences (n = 2) or non-target descriptions, that is, non-canonical clauses, for example, *xianzai zai sihao wei de shi balei wuzhe* "Now on number 4 there is a ballerina." (n = 4). In total, 9 trials (0.78% of total responses) were excluded.

Results and discussion

Figure 6 reports the percentages of repeated nouns, overt pronouns, and null pronouns for spoken production. The analyses were carried out as before. Table 3 provides a summary of the results. The analysis of repeated nouns indicated more repeated nouns (relative to all pronouns) in the two-box condition (62.8%, SE = 1.9%) than in the one-box condition (58.2%, SE = 2.0%), with no effect of gender congruence. The analysis of overt pronouns showed no effect of situational congruence or gender congruence, although both factors slightly decreased the rates of overt pronouns numerically. In contrast, the analysis of null pronouns revealed a main effect of situational congruence, with fewer null pronouns (relative to others) in the two-box context (21.6%, SE = 1.6%) than in the one-box context (24.3%, SE = 1.7%), with no other reliable effects.

The absence of a significant effect of gender congruence on the use of overt pronouns in Mandarin aligns with Hwang (2021). The current study involved a visual context, so the absence of a gender effect in Hwang's study cannot be explained by the absence of a visual context in their paradigm. The gender congruence effect reported in Zhou (2021) may therefore be due to the lack of within-participant variation, as the finding was based on a

Table 3. Summary of analysis of each response type in Experiment 3.

Fixed effect	Estimate	*SE*	*z*	*p*
Repeated noun				
(Intercept)	1.03	0.51	2.04	.042
Situational congruence	0.38	0.16	2.38	.017
Gender congruence	−0.08	0.16	−0.53	.596
Situation × Gender Congruence	0.22	0.32	0.71	.480
Overt pronoun				
(Intercept)	−3.68	0.75	−4.89	<.001
Situational congruence	−0.21	0.21	−0.98	.325
Gender congruence	0.27	0.20	1.33	.185
Situation × Gender Congruence	−0.06	0.40	−0.16	.872
Null pronoun				
(Intercept)	−4.56	1.24	−3.68	<.001
Situational congruence	−0.48	0.23	−2.06	.040
Gender congruence	−0.11	0.24	−0.48	.633
Situation × Gender Congruence	−0.39	0.48	−0.82	.414

between-experiment comparison. The question then is why no gender-based competition effect was observed. Critically, we also found no effect of situation-based competition on overt pronouns, as their usage did not reliably decrease when both referential candidates were enclosed in a red box. In contrast, situational congruence significantly reduced the use of null pronouns and increased the use of repeated nouns. These findings are consistent with Zhou et al. (2023), who reported no reliable effect of semantic similarity on overt pronouns under comparable conditions. Taken together, the results from Experiment 3 suggest that overt pronouns are generally insensitive to similarity-based competition in spoken Mandarin.

Experiment 4

A fourth and final experiment was conducted to investigate whether overt pronouns show sensitivity to similarity-based competition in the written modality, specifically whether they are influenced by both gender- and situation-based competition. This experiment was identical to Experiment 3, except that it was conducted in the written modality.

Method

Participants. A further 32 native speakers of Mandarin Chinese were recruited from Prolific (www.prolific.com) and from various UK university student communities.

Materials, procedure, and design. These were the same as in Experiment 3, except that the experiment was conducted in written modality. The experiment was conducted online using a web-based platform, Penn Controller for Internet Based Experiments (Zehr & Schwarz, 2018). Participants were asked to read aloud the context sentence first. Once they finished reading it, they pressed a key, which then presented a second display, where the target referent changed location. Participants described the change in writing in a box appearing below the display, starting with *Xianzai* "now."

Scoring. This was done in the same way as before. We scored whether the participants produced a repeated noun (n=1,085), an overt pronoun (n=127), or a null pronoun (n=49) when referring to the target referent. We excluded cases where participants used labels other than those mentioned in the context sentences (n=10), the inanimate third-person pronoun 它 "it" (n=1), or noun phrase coordinates (n=3); there was a recording failure (n=1); participants used expletive or other non-target constructions (n=4). In total, 19 trials (1.49% of total responses) were excluded.

Results and discussion

Figure 7 shows the rates of repeated nouns, overt pronouns, and null pronouns as a function of gender congruence and situational congruence. The analyses were conducted as before. Table 4 summarizes the results.

First, the analyses of repeated nouns showed a main effect of gender congruence, indicating more repeated nouns and fewer pronouns (including both null and overt pronouns) in the same-gender condition (87.7%, SE=1.3%) than in the different-gender condition (84.4%, SE=1.4%). In addition, a main effect of situational congruence indicated more repeated nouns (i.e., fewer pronouns, including both overt and null pronouns) in the two-box context (88.5%, SE=1.3%) than in the one-box context (83.6%, SE=1.5%). Second, the analyses of overt pronouns showed a marginal effect of gender congruence, with fewer overt pronouns in the same-gender condition (8.6%, SE=1.1%) than in the different-gender condition (11.5%, SE=1.3%), and a significant main effect of situational congruence, with fewer overt pronouns in the two-box context (8.0%, SE=1.1%) than in the one-box context (12.1%, SE=1.3%). Finally, the analyses of null pronouns revealed a marginal effect of situational congruence, with slightly fewer null pronouns in the two-box context (3.5%, SE=0.7%) than in the one-box context (4.3%, SE=0.8%).

The results of Experiment 4 contrast with those of Experiment 3 in two important ways. First, gender congruence between referents marginally reduced overt pronouns and significantly increased repeated nouns relative to null pronouns and overt pronouns, unlike in Experiment 3. This finding aligns with Hwang (2021), who also observed a gender effect in written Mandarin. Second, the referents' situational congruence significantly reduced the use of overt pronouns and only marginally reduced the use of null pronouns in Experiment 4. In contrast, in Experiment 3,

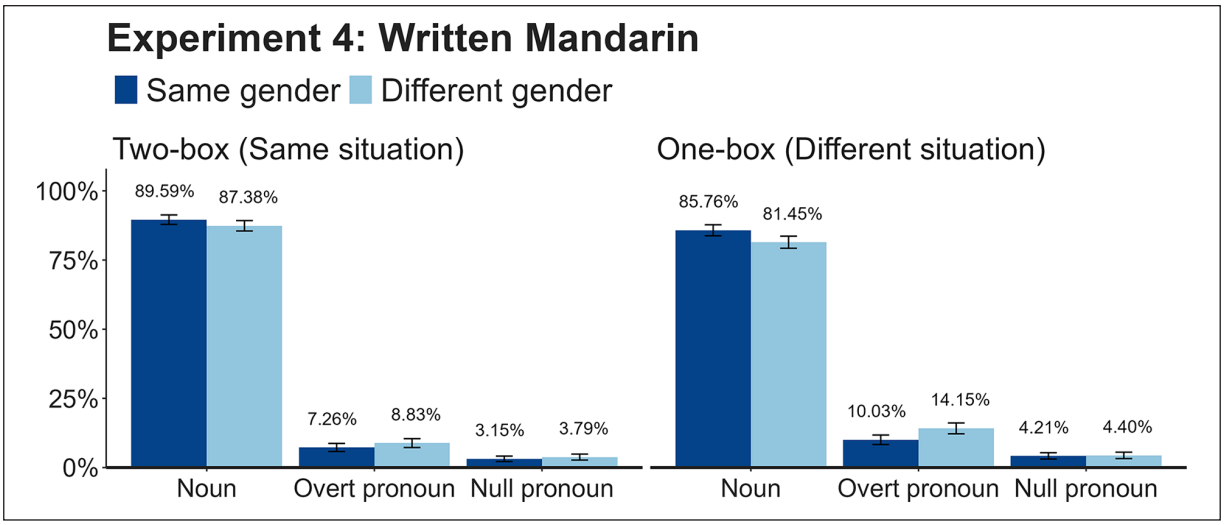

Figure 7. Percentages of referring expressions in Experiment 4.

Table 4. Summary of analyses of each response type in Experiment 4.

Fixed effect	Estimate	*SE*	*z*	*p*
Repeated noun				
(Intercept)	4.85	1.06	4.59	<.001
Situational congruence	0.83	0.24	3.43	.001
Gender congruence	0.53	0.24	2.22	.026
Situation × Gender Congruence	−0.26	0.48	−0.55	.582
Overt pronoun				
(Intercept)	−5.40	1.12	−4.80	<.001
Situational congruence	−0.85	0.41	−2.05	.040
Gender congruence	−0.53	0.31	−1.70	.089
Situation × Gender Congruence	0.44	0.53	0.83	.408
Null pronoun				
(Intercept)	−12.21	2.33	−5.24	<.001
Situational congruence	−1.37	0.79	−1.73	.084
Gender congruence	−0.81	0.79	−1.04	.301
Situation × Gender Congruence	−1.04	1.50	−0.69	.489

situational congruence primarily affected null pronouns and had no significant effect on overt pronouns. These findings suggest that modality may influence which forms are more susceptible to competition. We discuss this further in the next section.

General discussion

The current study examined ambiguity avoidance strategies in pronoun use, specifically testing the non-linguistic competition account. This account posits that speakers avoid pronouns when referential candidates are highly similar. Increased similarity enhances competition between referents, and to resolve this competition, speakers activate more specific information about their intended referent, resulting in the use of more explicit referring expressions. For example, in English, speakers are predicted to use fewer third-person pronouns like *he* and *she* when human referential candidates share the same gender, as the increased similarity from gender congruence amplifies competition at a non-linguistic level.

In contrast, the *ambiguity avoidance account* posits that speakers avoid pronouns to prevent linguistic ambiguity. Specifically, gender congruence reduces pronoun use—specifically in languages where pronouns encode gender—because it makes the gender information in the pronoun referentially ambiguous. According to this view, pronoun use is affected by gender-based ambiguity rather than referential similarity. To distinguish these accounts, we investigated whether speakers of Hindi and Mandarin Chinese—languages with non-gendered pronouns—reduce pronoun use as referential similarity increases with gender congruence between human referents.

We found support for the non-linguistic competition account in our Hindi experiments. In Experiment 1, human referents' gender congruence increased the use of repeated nouns relative to pronoun-based alternatives for subject antecedents, but not for postpositional object antecedents. Focusing on subject antecedents only, Experiment 2 further showed that shared gender led to a reduction in the use of the non-gendered pronoun *vah* relative to alternative forms. This gender effect on Hindi non-gendered pronouns was independent of verb gender agreement, suggesting that the reduction in pronoun usage was not caused by gender ambiguity in the verb. The gender effect in Hindi aligns with the findings from Finnish, where gender-based competition led to a reduction in the use of the non-gendered pronoun *hän* (Fukumura et al., 2013).

The absence of a gender-by-verb agreement interaction in Hindi mirrors the absence of a gender effect on Italian

null pronouns regardless of gender agreement on the verb (Fukumura et al., 2021). However, this does not mean that verb-related information can never influence pronoun usage. Research has shown that pronoun usage decreases to a greater extent when a competitor's situational similarity (e.g., also sitting on a horse) makes it compatible with the action attributed to the intended referent (e.g., getting off the horse) than when it does not (e.g., taking off a hat) (Fukumura et al., 2011). This effect may occur because, unlike verb agreement features, action-related information is activated during the speaker's conceptualization of the event—before lexical selection. In contrast, gender agreement features on the verb are processed as part of the grammatical encoding stage, after speakers have decided whether to use a pronoun.

In contrast to Hindi, we found no significant effect of gender-based competition on the non-gendered pronoun *ta* in spoken Mandarin (Experiment 3). A marginal effect was observed, however, in written Mandarin, where *ta* is gender-marked (Experiment 4). While the absence of a gender effect in spoken Mandarin may support the ambiguity avoidance account (Hwang, 2021), overt pronoun usage was also unaffected by situational similarity. In contrast, in written Mandarin, situational similarity decreased the use of overt pronouns while only marginally reducing null pronouns. These findings suggest that overt pronouns in Chinese are more strongly influenced by similarity-based competition in the written modality than in the spoken modality, whereas the reverse pattern holds for null pronouns. Thus, the lack of a reliable effect of gender congruence on overt pronoun use in Experiment 3 is likely because overt pronouns are generally less susceptible to similarity-based competition in spoken Mandarin.

According to the non-linguistic competition account, similarity-based competition occurs when speakers activate non-linguistic features of referents during the early stages of reference production. One possibility is that similarity-based competition is more likely to affect the usage of preferred pronouns in a given modality, as these pronouns are more likely to be planned early—at a stage when the referents' non-linguistic features influence competition. In spoken Mandarin, null pronouns were preferred over overt pronouns as substitutes for repeated nouns (58% vs. 42%), whereas in written Mandarin, overt pronouns were strongly favored (72% vs. 28%). Thus, in spoken Mandarin, where null pronouns are highly prevalent, overt pronouns are less likely to be planned early than in written Mandarin, while null pronouns may be planned earlier in speech than in writing. Consequently, situation-based competition tended to affect null pronouns in the spoken modality and overt pronouns in the written modality.

Similar to Italian null pronouns (Fukumura et al., 2021), gender similarity between referents did not reliably decrease the rate of Mandarin Chinese null pronouns, although null pronouns were influenced by situation-based competition. Fukumura et al. (2021) suggested that a referent's gender is not strongly activated when speakers plan to produce null pronouns, since null pronouns do not encode conceptual features like gender. This idea is supported by findings in second-language processing, which suggest that second-language speakers are more likely to make gender selection errors (e.g., *he* vs. *she* in English) if their first language permits null pronouns (Antón-Méndez, 2010; Tsoukala et al., 2017). Additionally, eye-tracking research has shown that speakers fixate on the referent less when using a pronoun than when using a repeated noun (Van der Meulen et al., 2001), raising the possibility that they attend to the referent even less when using a null pronoun than an overt pronoun. Situational similarity—such as both characters being enclosed in red boxes—tends to influence similarity-based competition regardless of pronoun type, as it involves a highly salient contextual feature that is accessible independently of how referents are conceptually encoded.

In Hindi, null pronouns were used relatively infrequently, and, unlike overt pronouns and demonstrative-noun responses, their usage was unaffected by the grammatical role of the antecedent. This contrasts with findings from previous studies in languages such as Italian and Spanish (e.g., Contemori & Di Domenico, 2021, 2023) and Mandarin Chinese (e.g., Lam & Hwang, 2022), which typically show a decline in null pronoun usage with non-subject antecedents. The low preference for null pronouns in Hindi, in contrast to spoken Mandarin, might explain the observed gender effect on Hindi overt pronouns. For subject antecedents, Hindi speakers often planned overt pronouns early, allowing gender similarity to influence their use through similarity-based competition. The current study did not explore the impact of situational similarity in Hindi. This was because Hindi overt pronouns were already shown to be affected by gender-based competition. To date, all overt pronouns that show gender-based competition have also exhibited effects of similarity-based competition (Fukumura et al., 2011; 2013; 2021; 2022). Future research could nevertheless examine how Hindi pronouns and demonstrative-noun expressions are affected by situation-based competition. Since gender congruence led to reduced overt pronoun use and increased use of more explicit forms, we expect situation-based competition to produce a similar effect on referential choice.

As discussed earlier, theories of reference typically assume that different pronominal expressions signal varying levels of accessibility (Ariel, 1990) or types of cognitive status (Gundel et al., 1993). While these theories are compatible with the effects of antecedent subjecthood on the use of pronouns, demonstrative-nouns, and repeated nouns (i.e., more pronoun and demonstrative-noun responses, and fewer repeated noun responses) observed in Experiment 1, they do not account for several key findings. Specifically, they do not explain why gender-based competition reduces the use of pronoun responses but not demonstrative-noun responses in Hindi, despite both being

similarly influenced by the subjecthood of their antecedents. Moreover, if pronoun usage was determined solely by referent accessibility, it is unclear why null pronouns are influenced by situation-based but not gender-based competition, why Mandarin overt pronouns are affected by situation-based competition in writing but not in speech, and why such modality-specific patterns emerge at all.

Thus, we propose updating the non-linguistic competition account as follows. According to this account, the use of semantically impoverished pronominal forms—such as null and overt pronouns—is cross-linguistically influenced by similarity-based competition. When a referent is similar to alternative entities, speakers tend to activate more specific semantic information, regardless of whether the similarity affects the linguistic ambiguity of the pronouns. This leads to the increased usage of more explicit expressions (e.g., repeated nouns or demonstrative-noun constructions in Hindi), which can help prevent potential ambiguity. Importantly, though, when a language allows multiple pronominal forms, similarity-based competition is more likely to influence the form preferred in that language or modality, as it is typically planned earlier in reference production at a stage when competition reduces the referent's accessibility. In some languages or modalities, null pronouns are more frequent than overt pronouns, or vice versa, making them more susceptible to similarity-based competition. Furthermore, different pronominal forms may be sensitive to different types of competition. Specifically, null pronouns are cross-linguistically less sensitive to gender-based competition than overt pronouns, as speakers tend not to activate detailed referential features—such as gender—when planning null pronouns.

Acknowledgments

We thank Niralee Gupta, Meghna Bose, Prachi Khandekar, Satyam Kumar, and Sana Hariyani for their help with data collection and coding.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by the Leverhulme Trust under Grant [RPG-2016-253].

Ethical considerations

This study was approved by the University of Stirling General University Ethics Panel (approval number GUEP 330) on 17 January 2018. All in-person participants (Experiments 1–3) provided written informed consent prior to participation, while online participants (Experiment 4) indicated consent by ticking a consent box.

Data accessibility statement

The data and materials from the present experiment are publicly available at the Open Science Framework website: https://osf.io/rsu7b/.

ORCID iDs

Kumiko Fukumura https://orcid.org/0000-0002-1475-9304
Shi Zhang https://orcid.org/0000-0002-1459-5215
Sakshi Bhatia https://orcid.org/0000-0002-4867-0076
Samar Husain https://orcid.org/0000-0002-4336-4897

Notes

1. वो (vo) is commonly used in colloquial speech to mean "that", instead of the more formal वह (*vah*).
2. This was based on the requirements that (a) both referential candidates must be mentioned; (b) the sentence introduces the referent's current position in the display; (c) the sentence can elicit pronoun responses frequently by making the target referent more prominent than the referential alternative. Although the use of the postpositional object phrase may sound over-specified without another category exemplar (e.g., no other policewoman in Figure 1), any potential awkwardness should be consistent across all conditions.
3. *Zai* serves the function of a copular verb (Chao, 1968).

References

Alonso-Ovalle, L., Fernández-Solera, S., Frazier, L., & Clifton, C. (2002). Null vs. overt pronouns and the topic-focus articulation in Spanish. *Italian Journal of Linguistics*, *14*, 151–170.

Antón-Méndez, I. (2010). Gender bender: Gender errors in L2 pronoun production. *Journal of Psycholinguistic Research*, *39*(2), 119–139.

Ariel, M. (1990). *Accessing noun-phrase antecedents*. Routledge.

Arnold, J. E. (2001). The effect of thematic roles on pronoun use and frequency of reference continuation. *Discourse Processes*, *31*(2), 137–162.

Arnold, J. E., & Griffin, Z. M. (2007). The effect of additional characters on choice of referring expression: Everyone counts. *Journal of Memory and Language*, *56*(4), 521–536.

Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S., & Trueswell, J. C. (2000). The rapid use of gender information: Evidence of the time course of pronoun resolution from eye tracking. *Cognition*, *76*(1), B13–B26.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2018). *Parsimonious mixed models* (arXiv:1506.04967). arXiv. http://arxiv.org/abs/1506.04967

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Bernolet, S., Hartsuiker, R. J., & Pickering, M. J. (2009). Persistence of emphasis in language production: A cross-linguistic approach. *Cognition*, *112*(2), 300–317.

Brennan, S. E. (1995). Centering attention in discourse. *Language and Cognitive Processes*, *10*(2), 137–167.

Caramazza, A., Grober, E., Garvey, C., & Yates, J. (1977). Comprehension of anaphoric pronouns. *Journal of Verbal Learning and Verbal Behavior*, *16*(5), 109–601.

Carminati, M. N. (2002). *The processing of Italian subject pronouns* [Unpublished doctoral thesis]. University of Massachusetts, Amherst.

Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics and point of view. In C. N. Li (Ed.), *Subject and topic* (pp. 26–55). Academic Press.

Chao, Y. R. (1968). *A grammar of modern spoken Chinese*. University of California Press.

Clancy, P. M. (1980). Referential choice in English and Japanese narrative discourse. In W. L. Chafe (Ed.), *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production* (pp. 127–201). Ablex.

Contemori, C., & Di Domenico, E. (2021). Microvariation in the division of labor between null- and overt-subject pronouns: the case of Italian and Spanish. *Applied Psycholinguistics*, *42*(4), 997–1028.

Contemori, C., & Di Domenico, E. (2023). The production of subject anaphoric expressions in Italian and Mexican Spanish: A forced-choice experimental study. *Journal of Psycholinguistic Research*, *52*(6), 2257–2285.

Corbett, G. (1991). *Gender*. Cambridge University Press.

Crawley, R. A., Stevenson, R. J., & Kleinman, D. (1990). The use of heuristic strategies in the interpretation of pronouns. *Journal of Psycholinguistic Research*, *19*, 245–264.

Di Eugenio, B. (1998). Centering in Italian. In M. Walker, A. Joshi, & E. Prince (Eds.), *Centering theory in discourse* (pp. 115–138). Oxford University Press.

Ferreira, V. S., & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology*, *40*, 296–340.

Ferreira, V. S., Slevc, L. R., & Rogers, E. S. (2005). How do speakers avoid ambiguous linguistic expressions? *Cognition*, *96*(3), 263–284.

Filiaci, F., Sorace, A., & Carreiras, M. (2014). Anaphoric biases of null and overt subjects in Italian and Spanish: a cross-linguistic comparison. *Language and Cognitive Processes*, *29*(7), 825–843.

Frederiksen, J. (1981). Understanding anaphora: Rules used by readers in assigning pronominal referents. *Discourse Processes*, *4*(4), 323–347.

Fukumura, K., & van Gompel, R. P. G. (2010). Choosing anaphoric expressions: Do people take into account likelihood of reference? *Journal of Memory and Language*, *62*(1), 52–66.

Fukumura, K., & van Gompel, R. P. G. (2011). The effect of animacy on the choice of referring expression. *Language and Cognitive Processes*, *26*(10), 1472–1504.

Fukumura, K., & van Gompel, R. P. G. (2015). Effects of order of mention and grammatical role on anaphor resolution. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(2), 501–525.

Fukumura, K., Hervé, C., Villata, S., Zhang, S., & Foppolo, F. (2021). Representations underlying pronoun choice in Italian and English. *Quarterly Journal of Experimental Psychology*, *75*(8), 1428–1447.

Fukumura, K., Hyönä, J., & Scholfield, M. (2013). Gender affects semantic competition: The effect of gender in a non-gender-marking language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(4), 1012–1021.

Fukumura, K., Pozniak, C., & Alario, F.-X. (2022). Avoiding gender ambiguous pronouns in French. *Cognition*, *218*, 104909.

Fukumura, K., van Gompel, R. P. G., & Pickering, M. J. (2010). The use of visual context during the production of referring expressions. *Quarterly Journal of Experimental Psychology*, *63*(9), 1700–1715.

Fukumura, K., van Gompel, R. P. G., Harley, T., & Pickering, M. J. (2011). How does similarity-based interference affect the choice of referring expression? *Journal of Memory and Language*, *65*(3), 331–344.

Garnham, A., Oakhill, J., Erlich, M. F., & Carreiras, M. (1995). Representations and processes in the interpretation of pronouns: New evidence from Spanish and French. *Journal of Memory and Language*, *34*(1), 41–62.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Givón, T. (1983). Topic continuity in discourse: An introduction. In T. Givón (Ed.), *Topic continuity in discourse: A quantitative cross-language study* (pp. 1–41). John Benjamins.

Gordon, P. C., & Hendrick, R. (1998). The representation and processing of coreference in discourse. *Cognitive Science*, *22*(4), 389–424.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of anaphoric expressions in discourse. *Language*, *69*(2), 274–307.

Huang, J. C. T. (1984). On the Distribution and Reference of Empty Pronouns. *Linguistics Inquiry*, *15*(4), 531–574.

Hwang, H. (2018). Semantic properties of pronouns modulate pronoun use: Evidence from Cantonese. *Discourse Processes*, *55*(1), 92–102.

Hwang, H. (2021). Avoidance of gender-ambiguous pronouns as a consequence of ambiguity-avoidance strategy. *Discourse Processes*, *58*(3), 251–259.

Hwang, H., Lam, S. Y., Ni, W., & Ren, H. (2022). The role of grammatical role and thematic role predictability in reference form production in Mandarin Chinese. *Frontiers in Psychology*, *13*, 930572.

Kachru, Y. (2006). *Hindi*. John Benjamins Publishing Company.

Kaiser, E., & Trueswell, J.C. (2008). Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. *Language and Cognitive Processes*, *23*(5), 709–748.

Karmiloff-Smith, A. (1985). Language and cognitive processes from a developmental perspective. *Language and Cognitive Processes*, *1*(1), 61–85.

Kliegl, R. (2014). *Reduction of complexity of linear mixed models with double-bar syntax*. https://rpubs.com/Reinhold/22193

Lam, S.-Y., & Hwang, H. (2022). How does topicality affect the choice of referential form? Evidence from mandarin. *Cognitive Science*, *46*(10), e13190.

Li, C. N., & Thompson, S. A. (1979). Third-person pronouns and zero-anaphora in Chinese discourse. In T. Givón (Ed.), *Discourse and syntax* (pp. 311–335). Brill.

Li, R. (2016). Hanyu disan rencheng huizhiyu xingshi yandong yuedushiyan [The eye-tracking experiment of the third

person anaphor in Mandarin texts]. *Zhongguo Yuwen*, *1*(1), 83–92.

Li, R. (2020). *Hanyu yupian zhutixing disan rencheng daici huizhi fenxi* [An analysis of anaphoric subject pronouns in Chinese discourse]. Zhongxi Shuju.

Lozano, C. (2016). Pragmatic principles in anaphora resolution at the syntax-discourse interface: advanced English learners of Spanish in the CEDEL2 corpus. In M. Alonso-Ramos (Ed.), *Spanish learner corpus research: Current trends and future perspectives* (pp. 235–265). John Benjamins Publishing Company.

Misersky, J., Gygax, P. M., Canal, P., Gabriel, U., Garnham, A., Braun, F., Chiarini, T., Englund, K., Hanulikova, A., Ottl, A., Valdrova, J., Von Stockhausen, L., & Sczesny, S. (2014). Norms on the gender perception of role nouns in Czech, English, French, German, Italian, Norwegian, and Slovak. *Behavior Research Methods*, *46*, 841–871.

Neeleman, A., & Szendrői, K. (2007). Radical pro drop and the morphology of pronouns. *Linguistic Inquiry*, *38*(4), 671–714.

Prasad, R., & Strube, M. (2000). Discourse salience and pronoun resolution in Hindi. *Penn Working Papers in Linguistics*, *6*, 189–208.

Prasad, R. (2000). A corpus study of zero pronouns in Hindi: An account based on centering transition preferences. In *Proceedings of DAARC 2000* (pp. 66–71). Lancaster University.

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Rizzi, L. (1986). Null objects in Italian and the theory of pro. *Linguistic Inquiry*, *17*(3), 501–557.

Rohde, H., & Kehler, A. (2014). Grammatical and information-structural influences on pronoun production. *Language, Cognition and Neuroscience*, *29*(8), 912–927.

Sanford, A. J., & Garrod, S. (1981). *Understanding written language: Explorations in comprehension beyond the sentence*. Wiley.

Singmann, H., & Kellen, D. (2020). An introduction to mixed models for experimental psychology. In D. H. Spieler & E. Schumacher (Eds.), *New methods in cognitive psychology* (pp. 4–31). Routledge.

Stevenson, R. J., Crawley, R. A., & Kleinman, D. (1994). Thematic roles, focus and the representation of events. *Language and Cognitive Processes*, *9*(4), 519–548.

Tomioka, S. (2003). The semantics of Japanese null pronouns and its cross-linguistic implications. In K. Schwabe & S. Winkler (Eds.), *The interfaces: Deriving and interpreting omitted structures* (pp. 321–339). John Benjamins.

Tsoukala, C., Frank, S. L., & Broersma, M. (2017). "He's pregnant": Simulating the confusing case of gender pronoun errors in L2 English. In *Proceedings of the 39th Annual Meeting of the Cognitive Science Society* (pp. 3392–3397). Cognitive Science Society.

Van der Meulen, F. F., Meyer, A. S., & Levelt, W. J. M. (2001). Eye movements during the production of nouns and pronouns. *Memory & Cognition*, *29*(3), 512–521.

Zehr, J., & Schwarz, F. (2018). *PennController for Internet Based Experiments (IBEX)*. https://doi.org/10.17605/OSF.IO/MD832

Zhang, A, & Kwon, N. (2022). The interpretational preferences of null and overt pronouns in Chinese. *Journal of Linguistics*, *58*(3), 649–676.

Zhou, Y. (2021). *Effects of similarity on speakers' production of referring expressions*. The University of Edinburgh.

Zhou, Y., Branigan, H. P., Yu, Y., & Pickering, M. J. (2023). The effects of semantic similarity on Mandarin speakers' referential expressions. *Quarterly Journal of Experimental Psychology*, *76*(11), 2579–2595.