Accepted refereed manuscript of: Fukumura K, Pozniak C & Alario F (2022) Avoiding gender ambiguous pronouns in French. *Cognition*, 218, Art. No.: 104909. https://doi.org/10.1016/j.cognition.2021.104909

Avoiding gender ambiguous pronouns in French

Kumiko Fukumura[1]

Céline Pozniak[1&2]

F.-Xavier Alario[3&4]

[1]*Psychology, Faculty of Natural Sciences, University of Stirling, UK*

[2] *Structures Formelles du Langage, Paris 8 University, CNRS, France*

[3]*Aix Marseille University, CNRS, LPC, Marseille, France*

[4]*Department of Neurological Surgery, University of Pittsburgh, USA*

The final version of this article will be available, upon publication, via its DOI: 10

https://doi.org/10.1016/j.cognition.2021.104909

Address correspondence to:

Kumiko Fukumura

Psychology, Faculty of Natural Sciences

University of Stirling

Stirling FK9 4LA

UK

kumiko.fukumura@stir.ac.uk

Abstract

Across many languages, pronouns are the most frequently produced referring expressions. We examined whether and how speakers avoid referential ambiguity that arises when the gender of a pronoun is compatible with more than one entity in the context in French. Experiment 1 showed that speakers use fewer pronouns when human referents have the same gender than when they had different genders, but grammatical gender congruence between inanimate referents did not result in fewer pronouns. Experiment 2 showed that semantic similarity between non-human referents can enhance the likelihood that speakers avoid grammatical-gender ambiguous pronouns. Experiment 3 pitched grammatical gender ambiguity avoidance against the referents' competition in the non-linguistic context, showing that when speakers can base their pronoun choice on non-linguistic competition, they ignore the pronoun's grammatical gender ambiguity even when the referents are semantically related. The results thus indicated that speakers preferentially produce referring expressions based on non-linguistic information; they are more likely to be affected by the referents' non-linguistic similarity than by the linguistic ambiguity of a pronoun.

Keywords: language production; ambiguity avoidance; referential communication; pronoun; semantic similarity; grammatical gender

1. Introduction

Most theories of reference assume that speakers avoid ambiguous referring expressions that could denote more than one referent in the context (e.g., Dale & Reiter, 1995; Frank & Goodman, 2012; Grice, 1975). Yet whether speakers choose linguistic forms to avoid ambiguity has been controversial (e.g., Allbritton et al., 1996; Arnold et al., 2004; Ferreira & Dell, 2000; Ferreira et al., 2005; Fukumura, 2016; Haywood et al., 2005; Kraljic & Brennan, 2005; Snedeker & Trueswell, 2003). For instance, speakers use modifiers when unmodified nouns create referential ambiguity (e.g., *large circle* in the context of more than one circle) (Ferreira et al., 2005; Fukumura, 2018; Horton & Keysar, 1996; Pechmann, 1989), but they also use modifiers even when unmodified nouns are unambiguous (e.g., *red circle* in the context of only one circle) (Belke & Meyer, 2002; Deutsch & Pechmann, 1982; Engelhardt et al., 2006; Fukumura, 2018; Pechmann, 1989). In the domain of syntactic ambiguity, findings suggest that speakers fail to avoid ambiguous sentence structures (e.g., Allbritton et al., 1996; Arnold et al., 2004; Ferreira & Dell, 2000; but see also Haywood et al., 2005; Snedeker & Trueswell, 2003) or that they provide disambiguating cues regardless of whether the context disambiguates (Kraljic & Brennan, 2005).

The current study aimed to contribute to these debates by investigating whether and how speakers avoid ambiguity that arises from the use of semantically impoverished but frequently used referring expressions, such as pronouns. Previous studies have shown that speakers of English use fewer gendered pronouns such as *he* or *she* (relative to nouns or names) when the referential candidates have the same gender than when they have different genders (e.g., Arnold & Griffin, 2007; Fukumura et al., 2010; 2011; 2013); for instance, speakers are less likely to use *he* to refer to a king (relative to more explicit expressions like *the king*) when the king is co-present with a male pilot compared to when the king is co-present with a female flight attendant. Currently, there are two accounts that explain this effect. The first account, which we call **the gender ambiguity avoidance account**, claims that speakers avoid *gender ambiguous* pronouns

for their addressee; in the above example, *he* could refer to the male pilot, as well as to the king, but not to the female flight attendant. Hence, speakers reduce the use of pronouns when the gender on the pronoun does not uniquely identify the referent in the context. Whilst this account assumes that speakers are adept at modelling linguistic ambiguity for their addressee, the second account, which we call the **non-linguistic competition account**, claims otherwise.

Research has shown that the mere presence of a referential alternative in the context, or a *competitor,* results in fewer pronouns (i.e., more nouns and names) (Arnold & Griffin 2007; Clancy, 1980; Fukumura et al., 2010), and this effect is particularly strong when the referential candidates share the same gender compared to when they do not (Fukumura et al., 2010). These findings have led to the proposal that pronoun choice is affected by the degree of competition between the referential candidates in the speaker's message, which is increased by the referents' similarity: The more similar the referents are, the more strongly they compete, which in turn reduces pronoun rates (Arnold & Griffin, 2007, Fukumura et al., 2011; 2013). Under this non-linguistic competition account, speakers use fewer pronouns (relative to full nouns) when the referential candidates have the same gender than otherwise, not because the gender of the pronoun is referentially ambiguous as claimed by the gender ambiguity avoidance account, but rather because the referents of the same gender are *non-linguistically more similar* (e.g., two males) than the referents of different genders (e.g., a male and a female).

Indeed, there is evidence indicating that gender congruence affects competition. Speakers are more likely to make substitution errors (e.g., *the pilot, sorry, the king gets off the horse*) when the potential human referents have the same gender than otherwise (Fukumura et al., 2013; Griffin & Wangerman, 2013), though such an effect may arise at a lexical level, where speakers select specific names or roles, rather than at a non-linguistic level. Stronger support for the non-linguistic competition account comes from findings that speakers tend to use fewer pronouns when human referents have the same gender than otherwise, even when pronouns are not

gender-marked. The Finnish pronoun *hän* can refer to a male or female, so *hän* is ambiguous regardless of whether the referents have the same gender or different genders. Nevertheless, in Fukumura et al. (2013), Finnish speakers produced *hän* less frequently when the referential candidates had the same gender than otherwise. Since the Finnish pronoun does not express the referent's gender, this effect cannot be attributed to gender ambiguity avoidance.

The non-linguistic competition account accords with findings from Ferreira et al. (2005), who argued that whilst speakers are sensitive to the referents' non-linguistic similarity, they often fail to avoid *linguistic ambiguity*: Participants almost always used disambiguating modifiers when two referential candidates had the same identity (*large bat* rather than *bat* in the context of two flying bats), but they frequently failed to do so when the referential candidates had different identities but their referring expressions were homophonic (e.g., *bat* for a flying bat and for a baseball bat)[1]. When the referents share the same identity (a flying bat), the similarity that underlies the ambiguity can be detected non-linguistically without any linguistic encoding. When the referents have different identities, however, speakers do not know whether a referring expression is ambiguous before producing it; they first need to choose and linguistically encode the referring expression before detecting its (phonological) ambiguity. Hence, Ferreira et al. argued that speakers are more likely to avoid ambiguity when the similarity underlying the ambiguity can be determined at a non-linguistic/conceptual level rather than at a linguistic level.

The non-linguistic competition account is supported by other similarity effects. Research has shown that speakers tend to use fewer pronouns when the referential candidates are more similar, even when the use of pronoun is referentially unambiguous. For instance, Fukumura et al. (2011; 2013) reported that participants used fewer pronouns when the referential candidates had similar situational attributes (e.g., both were sitting on a horse) than otherwise (e.g., the

[1] Some may argue that the results can be at least in part explained by the difference in the base rates of *flying bat* and *large bat* (e.g., *large bat* is simpler to say or conceptualize than *flying bat*).

referent was sitting on a horse, whereas the referential competitor was standing on the ground). This occurred in both English and Finnish, and the effect in English was not modulated by the gender ambiguity of the pronouns. Moreover, in a sentence completion task, Fukumura and Van Gompel (2011) reported that participants generally used the pronoun *they* less often when the two referential candidates mentioned in the previous sentence had the same animacy (e.g., boats vs. yachts, swimmers vs. rowers) than when they differed in animacy (e.g., boats vs. swimmers, yachts vs. rowers), and this was the case even when the referential competitor was a singular so the use of *they* was referentially unambiguous. This finding was in accord with the view that referents of the same animacy are more similar and tend to compete more strongly during language production than those that differ in animacy. Gennari et al. (2012) reported that in complex phrase production, competition arising from the referents' animacy congruence results in the inhibition of the agent role, with more agentless passives as in *the man being punched* as opposed to *the man being punched by the woman*.

Fukumura et al. (2011) spelled out the non-linguistic competition account in detail. The authors argued that speakers' pronoun choice is affected by how easily the referent is identified when speakers initiate reference production. The account was inspired by Griffin (2010), who explained naming errors by adapting theories of person recognition (Bruce & Young, 1986; Cohen, 1990): Speakers initiate reference by activating the referent's non-linguistic features, including its gender (e.g., *being a male*). When the referential candidates share a feature, the shared feature sometimes activates the referent, but it activates a competitor on other occasions (cf. Oberauer & Kliegl, 2006). Hence, on average, the referent is less activated and the competitor is more activated, when the referential candidates share a similar feature than otherwise. When the similarity between the referential candidates enhances competition between them, speakers resolve competition by activating more specific features of the referent (e.g., *being a king*) to boost its activation, which inputs to the production of more explicit referring

expressions (e.g., *king* rather than *he*). Although the use of more explicit referring expressions in the presence of similar entities may help listeners' comprehension, assuming that more explicit referring expressions would help listeners discriminate similar referents, it could be at least in part a production-internal effect: The presence of a similar referential competitor results in additional conceptual or message encoding, because speakers themselves require additional information to resolve competition in their message before initiating language production processes, in keeping with the views that more fine-grained information is necessary for speakers to select lexical items in the presence of semantic alternatives (Abel et al., 2009).

The non-linguistic competition account accords with accessibility-based theories of reference (Ariel, 1990; Chafe, 1976; Givón, 1983; Grosz, Joshi, & Weinstein, 1995; Gundel, Hedberg, & Zacharski, 1993), which assumes that pronoun use is affected by the referent's salience or accessibility: Speakers are more likely to use pronouns when the referent is more salient or accessible in the discourse. For instance, speakers are more likely to use pronouns (rather than repeated names or nouns) when the referent has been mentioned as the grammatical subject rather than when it has been mentioned a non-subject role in the preceding sentence (Arnold, 2001; Brennan, 1995; Fletcher, 1984; Fukumura & Van Gompel, 2010, 2011; 2015; Stevenson et al., 1994); it is well established that the grammatical subject is more prominent than other syntactic roles (Brennan et al., 1987; Gordon et al., 1993; Grosz et al., 1995). Furthermore, animate entities are more likely than inanimate entities to be mentioned as the grammatical subject (e.g., Bock et al., 1992; Ferreira, 1994; Gennari et al., 2012; Prat-Sala & Branigan, 2000) and be subsequently pronominalized (Dahl & Fraurud, 1996; Fukumura & Van Gompel, 2011), presumably because animate entities are more salient to language users (Aissen, 2003; Comrie, 1989; Foley & Van Valin, 1985; Silverstein, 1976).

The Current Study

The goal of the current study was to determine whether the referents' gender congruence only affects pronoun use via non-linguistic competition or whether speakers also avoid gender ambiguous pronouns. Previous work in English has focused on the pronouns' gender ambiguity for human referents, which is confounded with the referents' increased non-linguistic similarity. The current study thus focused on *grammatical gender* (Corbett, 1991), investigating whether and how speakers of French avoid gender ambiguous pronouns for non-human referents. French has two grammatical genders, masculine and feminine, and the same pronominal expressions, *il* (he/it) and *elle* (she/it) are used for human and non-human referents. Although grammatical gender can correlate with the word's semantic category (and/or some of its phonological properties), the assignment of grammatical gender is often arbitrary. For instance, different languages assign different grammatical genders to the same meaning (e.g., tomato is feminine in French, whereas it is masculine in Italian). Within the same language, the assignment of grammatical gender is often unpredictable from the noun's meaning; no semantic rule can explain why, in French, an orange and a table are feminine, but a lemon and a desk are masculine. Perhaps most relevant, Vigliocco et al. (2005) found that grammatical gender congruence of inanimate objects does not increase their conceptual similarity (but see Boroditsky et al., 2003; Konishi, 1993). In keeping with this, models of language production assume that grammatical gender is a syntactic property of nouns, represented at the lexical level that is distinct from the conceptual level (e.g., Caramazza, 1997; Levelt et al., 1999; Schriefers, 1993; Vigliocco et al., 1997).

Meyer and Bock (1999) reported that speakers of Dutch tend to make more gender selection errors (e.g., using the neuter pronoun *dat* for a common-gender referent) when the referential candidates have different grammatical genders than when they have the same grammatical gender. This gender mismatch cost is in line with other studies that showed

attraction errors in different languages*:* The gender or number features of irrelevant nouns interfere when speakers implement various agreement features (Bock & Miller, 1991; Bock et al., 1999; 2001; 2004; Barker et al., 2001; Eberhard et al., 2005; Santesteban et al., 2013; Vigliocco et al., 1995; 1996; Vigliocco & Franck, 1999; 2001; Villata & Franck, 2020; inter alia). Although the findings from Meyer and Bock indicate that speakers are sensitive to the grammatical gender congruence (or incongruence) between the referential candidates during pronoun selection, participants in that study were asked to produce pronouns only. Here we asked if speakers *avoid* using pronouns when the referential candidates have the same grammatical gender. The gender ambiguity avoidance account predicts that they should; speakers should take account of the gender ambiguity of the pronoun for their addressee's comprehension, not only when the human referents have the same gender, but also when the non-human referents have the same gender. By contrast, the non-linguistic competition account assumes that speakers use fewer pronouns when human referents share the same gender than otherwise, because the referents' gender congruence increases their non-linguistic similarity, which drives speakers' choice of using a pronoun or a noun. In this view, grammatical gender congruence between non-human referents is unlikely to result in fewer pronouns, because it does not affect the referents' similarity at a non-linguistic level.

1.1.Experiment 1

The first experiment directly contrasted the above-mentioned accounts, investigating whether speakers avoid gender ambiguous pronouns not only for human referents but also for inanimate referents. How the human referents' gender congruence affects competition was discussed in Fukumura et al. (2011; 2013). Suppose speakers refer to a sailor (**Fig.1A**). If they know that the sailor is a male, the referent's gender feature, *being a male,* gets activated as part of the referent's non-linguistic representation. If the context contains another male character, say, a king, then *being a male* also activates the king, who is also a male, resulting in

competition between them. Speakers then resolve this competition by activating a more specific feature of the referent, *being a sailor*, which results in the selection of a more specific concept, MARIN (SAILOR in French). Assuming that speakers select pronouns based on their conceptual gender for human referents (cf., Bock et al., 1999; Meyer & Bock, 1999) [2], this diagram also shows how a pronoun's gender ambiguity could be detected at a non-linguistic level. To refer to the sailor with a pronoun, speakers activate *being a male* and select the concept of the pronoun, IL (HE in English), which is conceptually associated with that gender feature. Because the referential competitor, the king, is also a male, the to-be-selected concept IL can also get associated with the competitor via the shared gender feature, *being a male*.

Unlike for human referents, the grammatical gender ambiguity of the pronoun for inanimate referents can only be determined linguistically. Existing models of pronoun production agree that speakers can only select a specific pronominal expression at the level, where the referent's grammatical information (*lemma*) is processed (Jescheniak et al., 2001; Schmitt et al., 1999). The diagram in **Fig.1B**, adapted from Schmitt et al. (1999) and Jescheniak et al. (2001), illustrates this. To refer to a piece of bread using a pronoun in French, speakers activate the concept, PAIN, which in turn activates the word's lemma, *pain,* along with its associated masculine gender node (M) at the lemma level. This enables the selection of the masculine pronoun, *il*. When the referential context contains a masculine referential competitor, say, a sunflower, *tournesol,* the pronoun, *il,* is gender ambiguous because *il* could be associated with the two lemmas, *pain* and *tournesol* via the shared masculine gender node (M) (cf. Roelofs, 1992). The question is whether speakers can detect this type of ambiguity.

[2] When the noun's grammatical gender diverges from the gender of a person, some languages allow either gender to be marked the pronoun (e.g., in Italian, *vittima,* victim, is grammatically feminine, but it can refer to either a female or a male, Cacciari et al., 1997). In French, the pronoun's gender generally follows the referent's gender, unless the speaker focuses on the referent's role expressed by the noun (Cornish, 1987).

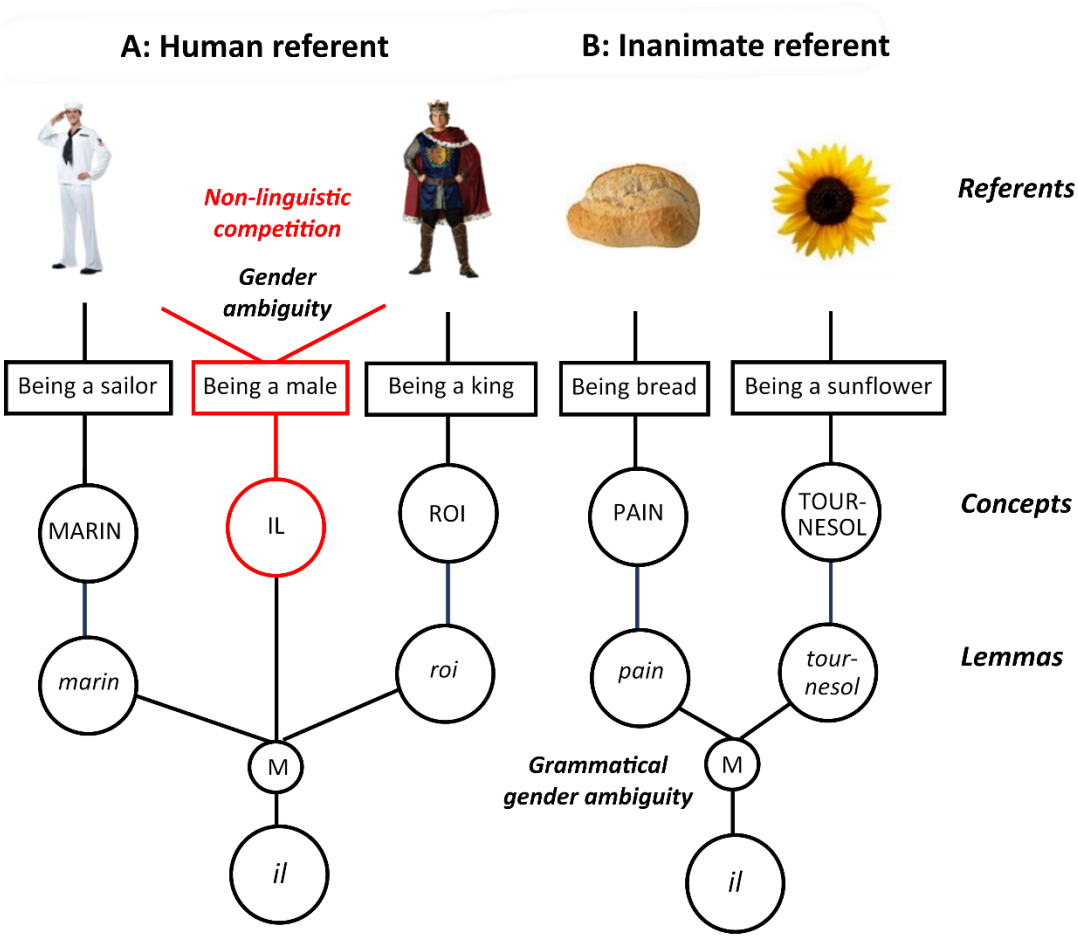

Fig.1. Level of representation at which gender ambiguity can be determined for human and inanimate referents.

As discussed earlier, the gender ambiguity avoidance account predicts that speakers avoid gender ambiguous pronouns for their addressee, regardless of whether the referents' non-linguistic similarity is increased by their gender congruence. This is based on the assumption that speakers monitor their own speech plan so that they can reverse the course of production processes as they deem necessary (Hartsuiker & Kolk 2001; Levelt, 1983; 1989). Whilst questions have been raised about the extent to which speakers engage in such monitoring for ambiguity avoidance (Ferreira et al., 2005; Horton & Keysar, 1996), if speakers could monitor and check for the ambiguity of the to-be-produced pronoun, especially using the comprehension system (Levelt, 1989), they may be able to avoid gender ambiguous pronouns. The rationale for this prediction can be found in evidence in comprehension. Findings from Garnham et al. (1995)

indicate that language users are sensitive to the gender ambiguity of pronouns for both human and non-human referents. In their study, when two inanimate entities mentioned in the preceding clause had the same grammatical gender (*La cape a protégé la veste*, The cape protected the jacket), the subsequent clause that contained a pronoun were read more slowly, compared to when the referents had different grammatical genders (*La cape a protégé le manteau*, The cape protected the coat). The magnitude of this effect did not differ from conditions where the referents were human (see also Cacciari et al., 2011; Carreiras et al., 1993).

By contrast, the non-linguistic competition account assumes that the choice of using a pronoun or noun is made at a non-linguistic level, so pronoun choice is primarily sensitive to the referents' non-linguistic similarity. Whereas gender congruence between the human referents increases the referents' non-linguistic similarity, grammatical gender congruence between the inanimate referents does not. Hence, assuming that speakers rely on production-internal mechanisms when choosing pronouns or nouns, this account predicts that speakers should produce fewer pronouns when the human referents have the same gender than otherwise, but grammatical gender congruence between inanimate entities is unlikely to affect speakers' pronoun choice.

1.1. Method

1.1.1. Participants. Thirty-two participants were recruited from the student community of Aix-Marseille Université in exchange for course credits or cash. They all reported to be French native speakers, aged between 18 and 30, with no visual impairments.

1.1.2. Materials and design. We constructed 40 experimental items for each experiment. Each item comprised a 2 × 3 visual display (see **Fig.2**), containing two referential candidates, the *target* and the *competitor*, and a context sentence (1), which introduced both referents linguistically. The target was always mentioned in the subject position, so the use of pronouns

was pragmatically felicitous (Meyer & Bock, 1999). The competitor was introduced in the prepositional phrases involving *next to*, *above* or *below*, which modified the subject. **Fig.2** shows example displays. In the figure, (A) and (B) represent the *human referent* conditions, in which the target and competitor were both humans, and (C) and (D) represent the *inanimate referent* conditions, in which the referential candidates were both inanimates. In both conditions, the target and competitor either had the same gender or differed in gender. For instance, the sailor (*marin*) in **Fig.2** is paired with a male competitor (*roi*, king) in the same-gender condition (A) and with a female competitor (*reine*, queen) in the different-gender condition (B). The bread (*pain*) in **Fig.2**, which is masculine, is paired with a masculine competitor (*tournesol*, sunflower) in the same-gender condition (C) and with a feminine competitor (*rose*) in the different-gender condition (D).

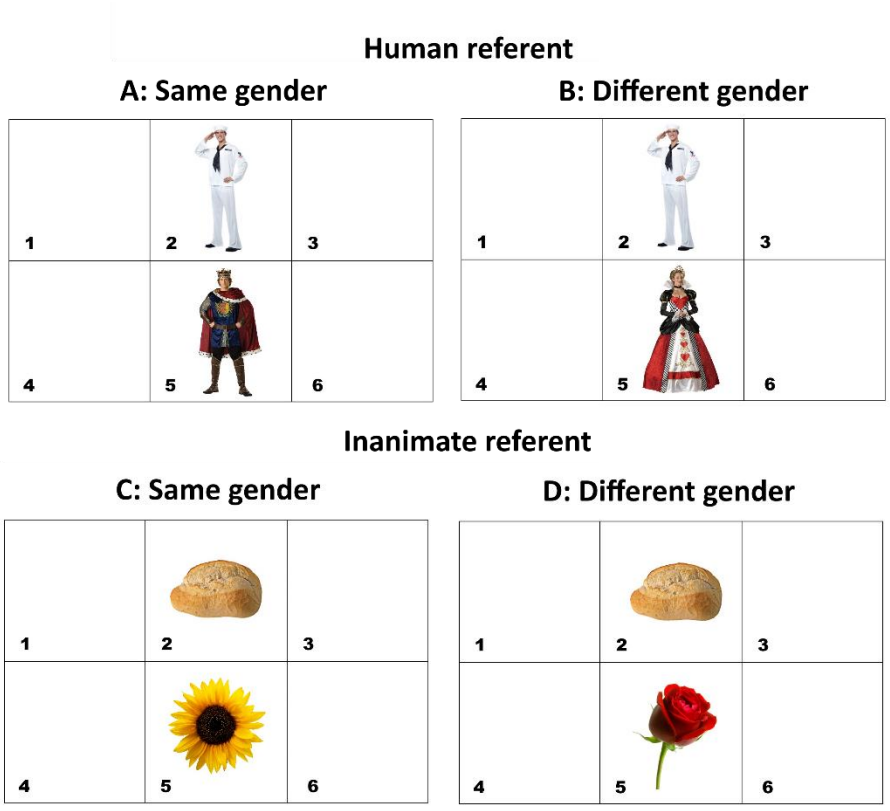

Fig.2. Example displays in Experiment 1.[3]

[3] 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.

(1) Context sentences

a. Le marin au dessus du roi est sur le numéro 2. *The sailor above the king is on No. 2.*

b. Le marin au dessus de la reine est sur le numéro 2. *The sailor above the queen is on No. 2.*

c. Le pain au dessus du tournesol est sur le numéro 2. *The bread above the sunflower is on No. 2.*

d. Le pain au dessus de la rose est sur le numéro 2. *The bread above the rose is on No. 2.*

(2) Target responses

a. Maintenant {le marin / il} est sur le numéro 3. *Now {the sailor / he} is on Number 4.*

b. Maintenant {le pain / il} est sur le numéro 3. *Now {the bread / it} is on Number 4.*

The competitor pairs for the human referent condition had comparable roles or professions (king vs. queen). Likewise, the competitors for the grammatical gender conditions belonged to the same or similar semantic categories (e.g., both were flowers) but they always had a different semantic category from the target. We used 20 male and 20 female human target characters; out of these, 19 had gender-unambiguous character roles (e.g., king, queen) and 21 had gender-typical character roles (e.g., sailor, fairy) (Misersky et al. 2014) (see Appendix A). For the grammatical gender condition, we used 40 target objects; half had a masculine gender and half had a feminine gender. Each character and object appeared as a same-gender competitor and a different-gender competitor across different items, thereby counterbalancing the competitors across conditions. The conditions were distributed such that each participant saw each object/character once as a target and once as a competitor. The positions of the entities were randomly distributed across conditions. Additionally, 80 filler items and 7 practice trials were constructed. These filler items had different sentence structures, and the grammatical role and linear position of the antecedent as well as the number of entities present in the context were also varied; on approx. 50% of the filler trials, the target referent was introduced as the

grammatical subject, whereas on the rest of the filler trials, it was introduced in the prepositional phrase or the subject in expletive constructions (i.e., *There is/are*). On approx. 25% of filler trials, the target referent was mentioned in the second or third position. On some filler items (approx. 25%), there were more than two entities mentioned or present in the display.

1.1.3. Design. We used a 2 (referent: human vs. inanimate) × 2 (gender ambiguity: ambiguous vs. unambiguous) repeated measures design, resulting in the creation of four 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. Each list contained 10 items in each condition, with only one version of each item. Eight participants were randomly assigned to each list.

1.1.4. Procedure. At the beginning of the experiment, the experimenter explained the tasks. Following informed consent, participants took part as the *speaker* and an experimental confederate took part as the *addressee*. The experimenter treated the confederate as a genuine participant throughout, and a post-experimental questionnaire showed no indication that participants realized that the confederate was not genuine. Both the speaker and addressee sat at a table, each facing a computer monitor. An occluder attached to the side of the speaker's screen prevented the addressee from seeing what was shown on the speaker's screen, but the speaker was able to see their addressee's screen. On each trial, both participants were presented with a 2 × 3 display on their monitor (Fig.1). The speaker then pressed a key on the keyboard to proceed, which triggered a presentation of a context sentence (1a-d), appearing below the display. There was no sentence shown on the addressee's screen so the speaker read it aloud, and pressed a key. Then the target changed location, while the competitor remained in the same position. The speaker's task was to describe the change (2), such that the addressee was able to indicate which object moved to which location by pointing. The speaker was instructed to produce a new

sentence, starting with *Maintenant* ("Now"). The experimenter gave example descriptions in the initial practice trial using repeated nouns and pronouns ("for instance, you could say *Now, he is on Number 6* or *Now the man is on Number 6*"), so that speakers did not consider that they would always name each image, and reminded that alternative referring expressions were also possible in the initial practice trials. The speaker was asked to pay attention to the addressee's comprehension, by pressing a "yes" key if they thought their addressee understood their description and a "no" key if they thought that their addressee misunderstood their description.

To control for the feedback from the addressee, prior to the experiment, the addressee (confederate) was instructed to respond as accurately as possible on what they heard, and to identify the entity mentioned in the subject position (i.e., target referents) in cases where participants produced gender ambiguous pronouns. We used E-Prime to present the stimuli and record participants' speech. Each session took approximately 40-45 mins. There was a short break halfway through. Informed consent was obtained from all participants before the experiment.

1.1.5. Scoring. We scored whether participants produced a pronoun (n = 508) or a repeated noun phrase (n = 759) to refer to the target in the first-mentioned subject position in their descriptions. We excluded 13 trials (1.0% of total responses), in which participants inadvertently used the competitor's role name (n = 1) or marked its gender on a pronoun (n = 1); they changed their responses (e.g., *Now he, the boy is on number 3*) (n = 5); or they used other nouns than repeated nouns (n = 3); and there was a technical error (n = 3).

1.2.Results and discussion

Fig.3 reports the percentages of pronoun responses out of pronouns and repeated nouns. As shown in the graph, some participants produced only repeated nouns (i.e., 0% pronouns) or only pronouns (i.e., 100% pronouns), at least in one condition; but most participants produced both types of referring expressions. Throughout the article, the choice of referring expressions

was analyzed using logit mixed effects modelling (e.g., Baayen et al., 2008; Barr et al., 2013). Because some results from the current study were in support of a null hypothesis, Bayesian binomial regression models (Carpenter et al., 2017; Bürkner, 2017; Bürkner & Charpentier, 2020) were also carried out. The results of these analyses corroborated those from the frequentist analyses and the full results are reported in Appendix B.

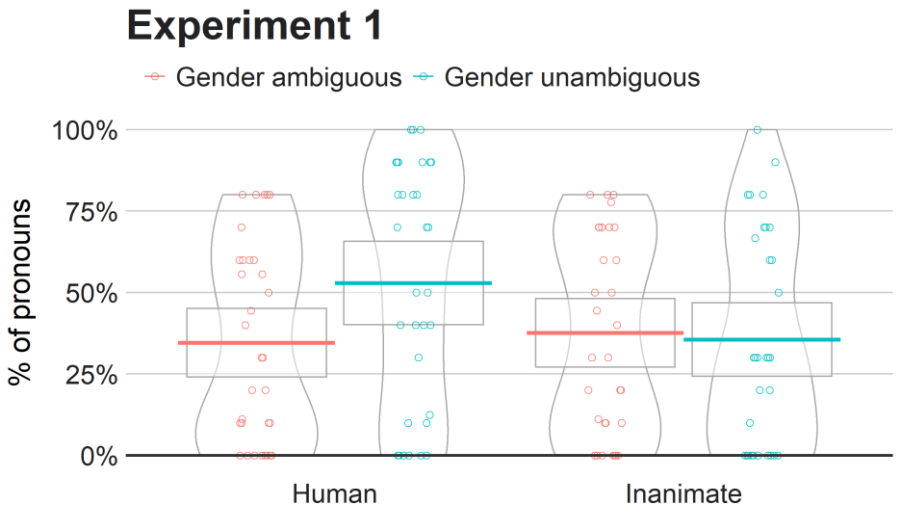

Fig.3. Percentages of pronouns out of pronouns and nouns in Experiment 1. Lines = condition means; boxes = 95% confidence intervals; dots = individual participant means; grey shapes = density estimates.

We used the lme4 package (Bates et al., 2015) in R (version 4.0.3., R Core Team, 2020). All the fixed effects were centered before analysis; like sum coding, centering is a commonly adopted procedure for reducing collinearity between variables, thereby enabling us to interpret coefficients in terms of main effects and interactions (Baayen et al., 2008) as well as easing convergence in R (Gelman & Hill, 2007). As for random effects, we first adopted the maximal random effect structure justified by the design (Barr et al., 2013), by including all the fixed effects (i.e., two main effects and an interaction between them) as well as by-participants and by-items random intercepts and their random slopes for all the relevant fixed factors (Barr et al., 2013), whilst supressing correlations between random effects to avoid overparameterization

(Bates et al., 2015; Bates et al., 2015; Kliegl, 2014; Singmann, & Kellen, 2020). We then removed random effects with (close to) zero variance to avoid potential singularity, though this procedure made no difference to the results, with little or no numerical impact on coefficient estimates or their significance. A pronoun response was coded as 1 and a repeated noun response 0. The analyses included referent (human referent condition coded as 1 and inanimate referent condition coded as 0) and gender (ambiguous condition coded as 1 and unambiguous condition coded as 0) as fixed effects, which were mean-centred and standardized.

Table 1 summarizes the fixed effects. A significant main effect of gender indicated that speakers used fewer pronouns when pronouns were gender ambiguous (36.3%) than when they were gender unambiguous (43.8%). The main effect of referent was also significant; speakers produced more pronouns when referring to humans (44.2%) than when referring to inanimates (36.0%). Crucially, there was a significant interaction between them: Simple effects found that gender ambiguity led to fewer (18.7%) pronouns when the referents were human, whereas it did not affect pronoun use when the referents were inanimate.[4]

Table 1

Analyses of the choice of expressions in Experiment 1

	Estimate	*SE*	*z*	*p*
Main analysis				
Intercept	-0.97	0.41	-2.35	.019
Gender (ambiguous vs. unambiguous)	-0.28	0.07	-3.76	< .001
Referent (human vs. inanimate)	0.26	0.07	3.46	.001
Gender × Referent	-0.36	0.07	-4.79	< .001
Gender effect in the human referent condition				
Intercept	-0.61	0.39	-1.56	.118
Gender	-0.63	0.12	-5.31	< .001
Gender effect in the inanimate referent condition				
Intercept	-1.11	0.39	-2.81	.005
Gender	0.08	0.10	0.77	.443

[4] The gender of the masculine competitor was not explicitly marked on the determiner in the context sentence (e.g., *du roi*, *du tournesol*). An additional analysis confirmed that neither the main effect of gender congruence nor the gender × referent interaction was modulated by referent gender (*Estimate* = -0.01, *SE* = 0.07, *z* = -0.09, *p* = .926; *Estimate* = -0.03, *SE* = 0.07, *z* = -0.42, *p* = .672, respectively).

In sum, the first experiment revealed that speakers of French produce fewer pronouns when human referents have the same gender than when they have different genders, replicating the gender congruence effect reported in the previous studies in English and Finnish (e.g., Arnold & Griffin, 2007; Fukumura et al., 2010; 2011; 2013). We also found that overall, speakers use more pronouns when referring to humans than when referring to inanimates. This replicates the effect of animacy congruence on the use of pronouns in other studies using different methods (Dahl & Fraurud, 1996; Fukumura & Van Gompel, 2012), providing support for the animacy hierarchy theories that claim that animates are more salient than inanimate ones to language users (Aissen, 2003; Comrie, 1989; Foley & Van Valin, 1985; Silverstein, 1976). Critically, we found no evidence that speakers use fewer pronouns when the inanimate referents have the same grammatical gender than when they have different grammatical genders[5]. Hence, the findings are in accord with the non-linguistic competition account, which claims that the referents' gender congruence results in fewer pronouns only when it increases their non-linguistic similarity. By contrast, the findings are incompatible with the gender ambiguity avoidance account, which claims that speakers avoid gender ambiguous pronouns for both human and inanimate referents by checking the pronoun's gender ambiguity for their addressee.

2. Experiment 2

Whilst the discrepancies of the Experiment 1 results with the findings in comprehension (Garnham et al., 1995) might point to the limitations of comprehension-based monitoring for ambiguity avoidance, speakers may not routinely employ comprehension-based monitoring (cf. Ferreira et al., 2005; Horton & Keysar, 1996). Some models of lexical selection assume that

[5] The current findings contrast with an unpublished experiment, presented at the CUNY 2019 32nd Annual CUNY Conference on Human Sentence Processing. These preliminary results indicated that speakers avoid gender ambiguous pronouns for inanimate referents as much as for human referents. This experiment had a gender assignment error to one noun (used across three items) and some minor errors in the filler trials. These shortcomings were rectified in the current experiment, and the current results were replicated, using slightly different stimuli, in an additional experiment, whose findings are reported in Appendix C.

language production involves *conflict monitoring*; speakers minimize lexical selection errors by monitoring interference from alternatives (e.g., Dhooge & Hartsuiker, 2011; 2012; Gauvin & Hartsuiker, 2020; Nozari et al., 2011). For instance, Dhooge and Hartsuiker (2011) show that speakers make fewer naming errors when distractor words were taboo words than otherwise, suggesting that speakers can avoid socially inappropriate errors by monitoring interference. One possibility is therefore that speakers failed to avoid gender ambiguous pronouns for inanimate referents in Experiment 1, because inanimate referential alternatives did not interfere with pronoun selection processes, so speakers did not check for potential ambiguity. Unlike the human referents, the inanimate referents of the same gender were always dissimilar, so the referential alternatives might not have competed for selection. That is, for speakers to detect the grammatical gender ambiguity of a pronoun, they need to engage in monitoring, and they are more likely to do so when there is similarity-based competition during pronoun selection.

The goal of Experiment 2 was thus to test this narrower version of the gender ambiguity avoidance account that speakers are more likely to detect the pronoun's gender ambiguity when the referential alternatives interfere during pronoun selection than otherwise. We did this, by investigating whether speakers avoid grammatical gender ambiguous pronouns when the referential candidates are semantically related. Much evidence indicates that the referents' semantic similarity affects lexical competition; speakers take longer to produce a word when they hear or see a distractor word that is semantically related to the target word than when both words are unrelated (e.g., Damian & Martin, 1999; Schriefers et al., 1990). Such semantic competition is assumed to arise at the lexical level rather than at a non-linguistic level (e.g., Bloem & La Heij, 2003; Caramazza & Hillis, 1990; Damian & Bowers, 2003; Roelofs, 1992; Schriefers et al., 1990): Semantic similarity of a non-target word delays responses only when the task requires naming (Schriefers et al., 1990). Moreover, semantic similarity has been shown to interact with syntactic

choice (Cleland & Pickering; 2003; Gennari et al., 2012) or increase syntactic agreement errors (Barker et al., 2001).

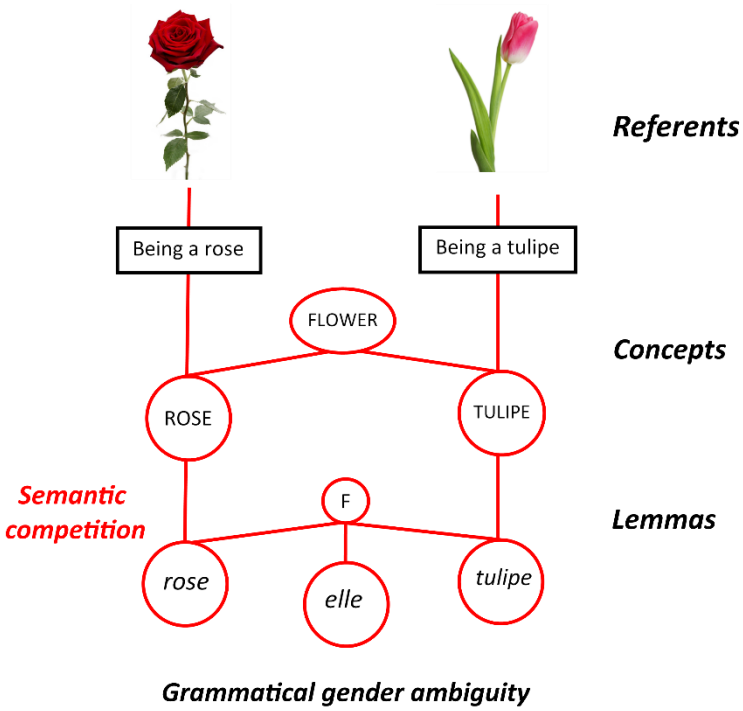

Fig.4. Level of representation at which semantic competition and grammatical gender ambiguity could affect pronoun use.

To consider how the referents' semantic similarity might affect competition, suppose speakers refer to a rose with a feminine pronoun, *elle*, in a context with a semantically related referential competitor, such as *tulipe* (tulip). In **Fig.4**, because a rose and a tulip are both a type of flower, the lexical concepts, ROSE and TULIPE, are linked to the shared concept, FLOWER, at the conceptual level. Moreover, because *rose* and *tulipe* are both feminine, both lemmas are linked to the same feminine gender node. Thus, assuming that speakers activate the lexical concept of the antecedent noun for producing a pronoun (Jescheniak et al., 2001; Schmitt et al., 1999; but see Meyer & Bock, 1999), following the activation of ROSE at the conceptual level, TULIPE receives some activation via FLOWER, and this results in lexical competition between the corresponding lemmas, *rose* and *tulipe*, at the lemma level (Levelt et al., 1999). Hence, when

speakers activate *rose* to select the feminine gender node (F) and the pronoun lemma, *elle*, *tulipe* should interfere, and the co-activation of *rose* and *tulipe* strengthens the links between all the activated lemmas (cf. Cleland & Pickering, 2003; Pickering & Branigan, 1998), enhancing the one-to-many mapping of the pronoun. According to the narrower version of the ambiguity avoidance account, speakers should be able to detect this gender ambiguity, as they carry out comprehension-based analyses in response to the similarity-based competition. Alternatively, speakers may of course focus on avoiding gender selection errors during pronoun selection. If so, they may fail to detect the gender ambiguity regardless of semantic similarity.

Thus, in Experiment 2, the referents were either from the same semantic categories (3a-3b) or from different semantic categories (3c-3d), and they had either the same grammatical gender (3a-3c) or different grammatical genders (3b-3d). If grammatical gender ambiguity avoidance is enhanced by semantic competition, as claimed by the narrower version of the gender ambiguity avoidance account, the referents' grammatical gender congruence is more likely to result in fewer pronouns when the referents are semantically related (3a-3b) than when they are not (3b-3d). Alternatively, speakers may be completely insensitive to the gender ambiguity of pronouns. If so, whilst the referents' semantic similarity might affect pronoun use, their grammatical gender congruence should not lead to fewer pronouns.

(3) Context sentence

a. La rose à côté de la tulipe est sur le numéro 2.	*The rose next to the tulip is on No 2.*
b. La rose à côté du tournesol est sur le numéro 2.	*The rose next to the sunflower is on No. 2.*
c. La rose à côté de la cigogne est sur le numéro 2.	*The rose next to the stork is on No. 2.*
d. La rose à côté du canard est sur le numéro 2.	*The rose next to the duck is on No. 2.*

(4) Target responses

Maintenant, {la rose / elle} est au numéro 6.	*Now {the rose / it} is on No. 6.*

2.1.Method

2.1.1. Participants. Forty participants were recruited from the RISC platform (https://www.risc.cnrs.fr/), social media and the University of Paris-Diderot student community in exchange for cash. They were all reported to be French native speakers, aged between 18 and 30, and to have no visual impairment. Data from one participant who did not follow the instructions were excluded.

2.1.2. Materials. We constructed 44 experimental items (see Appendix A). There were 44 target referents, taken from different categories such as flower, fruits and vegetables (n = 10), furniture (n = 4), music instruments (n = 4), kitchen utensils and tools (n = 6), clothes (n = 4), animals (n = 12), transport (n = 2) and body parts (n = 2). Each target object was combined with four competitor objects as follows. In the *semantically similar* condition (1a-b), the target and the competitor had the same semantic categories (e.g., both were flowers, Léger et al., 2008), whereas in the *semantically dissimilar* condition (1c-d), they had different semantic categories (i.e., the target is a flower, whereas the competitors are birds). Within each similarity condition, the competitor had either the same grammatical gender as the target (*tulipe* in 3a, *cigogne* in 3c) or a different grammatical gender from the target (*tournesol* in 3b, *canard* in 3d). Each competitor object occurred in all four conditions across four different items, to counterbalance the competitors across conditions. The conditions were rotated such that each participant saw each competitor only once. We avoided the target nouns that started with a vowel, so their grammatical gender was marked on the determiner, and the target and competitors had different initial syllables in all conditions.

2.1.3. Design. We used a 2 (semantics similarity: similar vs. dissimilar) × 2 (gender: ambiguous vs. unambiguous) repeated measures design. Together with 83 filler items, 44 experimental items were distributed across four lists, with each list having 11 items in each

condition, with only one version of each item. The conditions within each list were rotated as before. Ten participants were randomly assigned to each list.

2.1.4. Procedure. This was the same as in Experiment 1.

2.1.5. Scoring. We scored whether participants produced a pronoun (n = 545) or a repeated noun (n = 1162) to refer to the target. We excluded cases where participants used the competitor's role name (n = 8) or a pronoun that agreed with the competitor's gender (n = 9) or produced definite articles with wrong gender (n = 2); participants dropped the subject (e.g., …. *now on number 4*) (n = 7); or they changed their responses (e.g. *Now it, the rose is on number 4; Now the …it is on number 4*) (n = 5) or produced a dislocation (*Now it is on number 3 … the strawberry*) (n = 3); they produced non-repeated nouns (n = 3) or demonstrative pronouns (n = 2); they used other constructions (e.g., "There is the *rose* on number 1") (n = 9); or there was a technical error (n = 4). In total, 53 trials (3.0% of total responses) were excluded, and 1707 responses were submitted to the analyses.

2.2. Results and Discussion

Fig.5 reports the percentages of pronouns (out of pronouns and repeated nouns). The analyses were carried out as in Experiment 1; gender (ambiguous condition coded as 1 and unambiguous condition coded as 0) and semantic similarity (similar condition coded as 1 and dissimilar condition coded as 0) and were included as fixed effects, which were mean-centred and standardized. **Table 2** summarizes the results. There was a significant main effect of gender, indicating that speakers used fewer pronouns when the pronouns were gender ambiguous (28.7%) than when the pronouns were gender unambiguous (35.1%). A significant main effect of semantic similarity also indicated that speakers produced fewer pronouns when the referential candidates had the same semantic categories (29.8%) than when they had different semantic categories (34.0%). The gender × semantic similarity interaction was marginally significant.

This reflected the fact that gender ambiguity resulted in fewer (9.5%) pronouns when the referents were semantically related, whereas it did not when the referents were unrelated.[6]

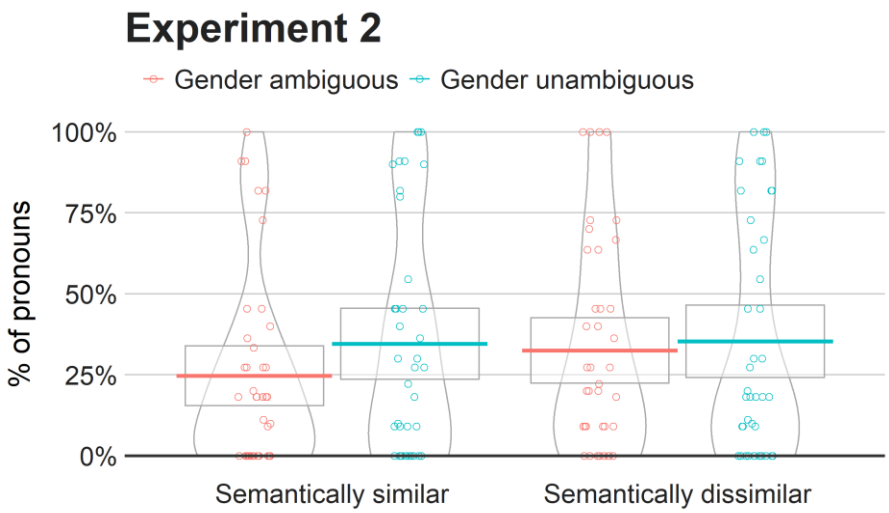

Fig.5. Percentages of pronouns out of pronouns and nouns in Experiment 2. Lines = condition means; boxes = 95% confidence intervals; dots = individual participant means; grey shapes = density estimates.

Table 2

Analyses of the choice of expressions in Experiment 2

	Estimate	*SE*	*z*	*p*
Main analysis				
Intercept	-1.55	0.45	-3.42	<.001
Gender (ambiguous vs. unambiguous)	-0.28	0.10	-2.74	.006
Semantic similarity (similar vs. dissimilar)	-0.19	0.07	-2.65	.008
Gender × Semantic similarity	-0.16	0.08	-1.91	.056
Gender effect in the similar condition				
Intercept	-1.74	0.45	-3.87	<.001
Gender	-0.43	0.13	-3.43	<.001
Gender effect in the dissimilar condition				
Intercept	-1.23	0.43	-2.86	.004
Gender	-0.12	0.12	-0.97	.334

[6] As in Experiment 1, we checked if the effects were modulated by the gender of the referents. Neither the main effect of gender congruence nor the gender × similarity interaction was modulated by target (or competitor) gender (*Estimate* = -0.01, *SE* = 0.08, *z* = -0.14, *p* = .885; *Estimate* = -0.03, *SE* = 0.07, *z* = -0.42, *p* = .677, respectively).

In sum, Experiment 2 showed that speakers avoid grammatical gender ambiguous pronouns for semantically related non-human referents, consistent with the narrower version of the ambiguity avoidance account that semantic competition at the lemma level could facilitate grammatical gender ambiguity avoidance.[7] The grammatical gender × semantic similarity interaction was marginal, however. Unlike gender congruence for human referents, semantic similarity is independent from the referents' grammatical gender congruence; although the referents' category congruence may enhance semantic competition during pronoun selection, it does not determine the gender ambiguity of the pronoun. The third and final experiment thus examined the robustness of this finding in a different set-up.

3. Experiment 3

Although the findings from Experiment 2 are compatible with the narrower version of the ambiguity avoidance account, whether speakers avoid grammatical gender ambiguity critically hinges on whether they check for it at the lexical level. The third and final experiment thus examined the robustness of such processes, by manipulating the referents' competition in the non-linguistic context. One possibility is that speakers always check for the pronoun's gender ambiguity whenever the potential referents are semantically related. This assumes that similarity-based competition during pronoun selection triggers comprehension-based monitoring irrespective of the referential context. An alternative possibility is that whether speakers engage in comprehension-based monitoring is context-dependent. Specifically, the non-linguistic competition account assumes that speakers normally choose whether to use a pronoun or a noun based on the non-linguistic information; if they can base their pronoun choice on the non-

[7] An additional experiment was run to check if the antecedents' phonological similarity might also enable grammatical gender ambiguity avoidance in French. The results showed that neither the referents' phonological similarity nor their grammatical gender congruence affects pronoun use (*p*s > .05) (Fukumura et al., 2020). Hence, it is very unlikely that similarity facilitates grammatical gender ambiguity avoidance at the phonological level.

linguistic context, they will be unlikely to engage in monitoring at the lexical level to backtrack the choice made at a non-linguistic level.

In Experiment 3, we thus manipulated the referential context as follows. In **Fig.6A**, the two objects are both in a red box, signalling to speakers that both could be the referent, whereas in **Fig.6B**, only one object is in a red box, signalling to speakers that only the object in a red box could be the referent. Other data from our work show that speakers produce fewer pronouns when the referential competitor could also be the referent (two box context) than when only the target could be the referent (one box context) (Fukumura et al., in press).

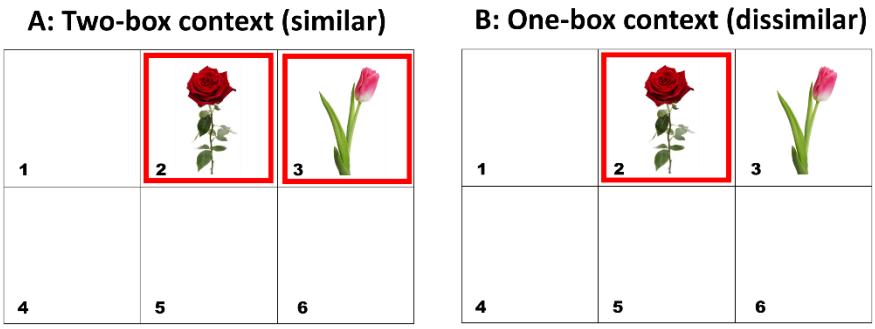

Fig.6. Example displays for the same gender condition in Experiment 3.

Note. The image for tulip was replaced by a sunflower in the different gender condition.

We assume that the box manipulation affects non-linguistic competition rather than linguistic competition, because neither the linguistic context nor the referring expressions themselves indicate whether an entity is in the box or not, and situational attributes such as *being in a box* are not expressed by the antecedent nouns. This is illustrated in **Fig.7**, where *being in a box* is attached to the non-linguistic representations of the two referents, rose and tulip, but not to the nouns' concepts nor their lemmas; *being in a box* is neither a part of the noun's concept nor a syntactic feature of the noun (i.e., a rose is a type of flower and *rose* is a feminine noun, regardless of whether the rose is in a box or not). Consistent with this assumption, Fukumura et al. (in press) found that the referents' semantic similarity delayed the onset latencies of repeated

nouns, but their situational similarity (box manipulation) did not. Likewise, whereas the referents' gender congruence increased the rates of substitution errors (e.g., *the queen, sorry, the witch*), their situational congruence (e.g., whether both entities are sitting on a horse) did not in Fukumura et al. (2013), reflecting the fact that, unlike the referents' gender, their situational attributes are not part of the concepts of the role names (i.e., a queen is a queen, irrespective of whether she is sitting on a horse).

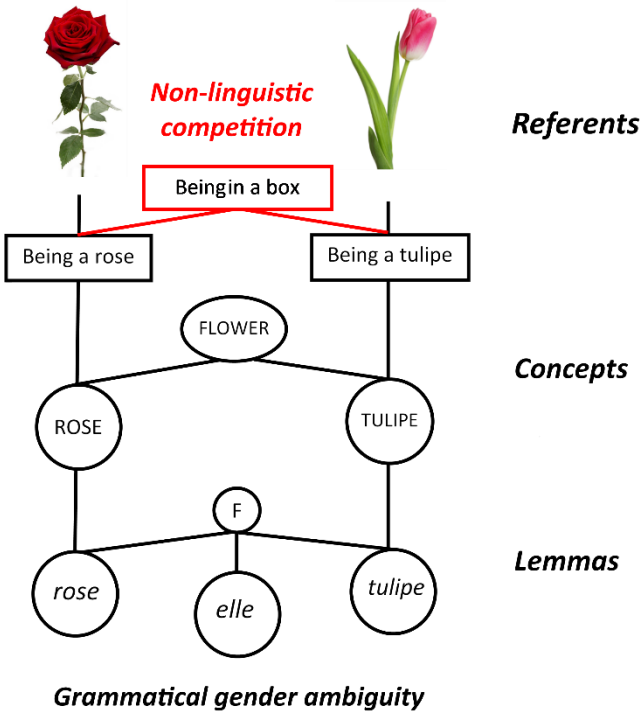

Fig.7. Level of representation at which the box manipulation and grammatical-gender ambiguity might affect pronoun use.

As we have just discussed, the non-linguistic competition account predicts that speakers preferentially resolve competition based on the information that is available prior to linguistic encoding: Speakers are more likely to choose pronouns in the one-box context, where the competitor could not be the referent, than in the two-box context, where the competitor could also be the referent, but once they have made such a choice, they do not monitor and check whether the grammatical gender of the to-be-produced pronoun is compatible with the gender of

the competitor. An alternative possibility, however, is that speakers take account of the pronoun's grammatical gender ambiguity, whenever the referents are semantically. In this view, non-linguistic competition might even facilitate grammatical gender ambiguity avoidance, rather than eliminating it; speakers should avoid grammatical gender ambiguous pronouns more when the referential competitor could also be the referent (two-box context) than otherwise (one-box context).

3.1.Method

3.1.1. Participants. Thirty-two participants were recruited from the same population as in Experiment 1.

3.1.2. Materials and procedure. There were 44 experimental items. These were taken from the semantically similar condition in Experiment 2: The target and competitor had either the same grammatical gender (3a) or different grammatical genders (3b), and they were always semantically related (3a-b). Critically, within each condition, the target and competitor were either both in a red box (*situationally similar* condition, **Fig.6A**), meaning that either can move, or only the target was in a red box (*situationally dissimilar* condition, **Fig.6B**), meaning that only the target could move. The procedure was the same as before, except that participants were informed that the presence of a box signalled which box could possibly move; when more than one entity was in a box, any entity in a red box could move. We used the same filler or practice trials as in Experiment 2, except that the box manipulation was applied; in some filler trials, only the target was in a red box, whereas in other filler trials, non-target objects were also placed in a red box.

3.1.3. Design. We used a 2 (situational similarity: similar vs. dissimilar) × 2 (gender: ambiguous vs. unambiguous) repeated measures design. As before, the experimental items were distributed across four lists, each list having 11 items per condition, and eight participants were randomly assigned to each list.

3.1.4. Scoring. We scored whether participants produced a pronoun (n = 376) or a repeated noun (n = 978) to refer to the target. We excluded cases with substitution errors (n = 3), pronouns with a wrong gender (n = 4); omitted subjects (n = 17); nouns with a demonstrative pronoun (n = 15); a noun conjunct (n = 1); plural nouns (n = 1); changed responses (n = 6); a dislocation (n = 6); technical errors (n = 1). In total, 54 trials (3.8%) were excluded and 1354 responses were analyzed.

3.2.Results and Discussion

Fig.8 reports the means. The analyses were carried out as before, including gender (coded as before) and situational similarity (two-box condition coded as 1 and one-box condition coded as 0). **Table 3** summarizes the results. There was a main effect of situational similarity; participants produced fewer pronouns when the two referential candidates were both in a red box (19.2%) than when only the target was in a red box (36.3%). Yet there was no main effect of grammatical gender ambiguity nor a significant situational similarity × grammatical gender ambiguity interaction. The Bayesian analysis supported these results, with the probability of an effect of situational similarity being smaller than zero was over .999 ($\hat{\beta}$ = -1.76, 95% CrI = [-2.60, -1.00]), but we found no support for gender ambiguity ($\hat{\beta}$ = 0.08, 95% CrI = [-0.39, 0.62], $P(\beta > 0)$ = .622) or the interaction ($\hat{\beta}$ = 0.17, 95% CrI = [-0.91, 1.20], $P(\beta > 0)$ = .632) (see the posterior distributions in **Fig.11** and a full summary of the analysis reported in **Table 5** in Appendix B). Moreover, the Bayes factors, estimated from the Bayesian Information Criterion values (Wagenmakers, 2007), were as low as .027 for gender ambiguity and .028 for the gender ambiguity × situational similarity interaction. These provided very strong support for the null hypotheses.

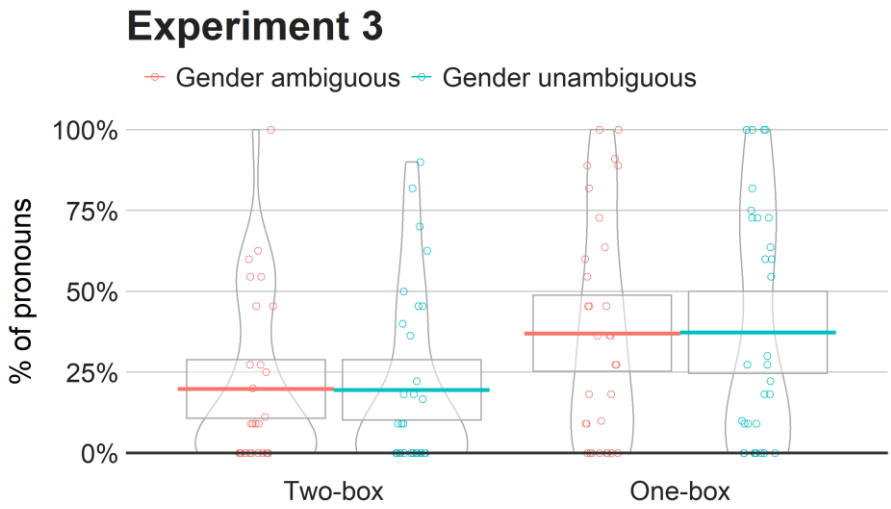

Fig.8. Percentages of pronouns out of pronouns and nouns in Experiment 3. Lines = condition means; boxes = 95% confidence intervals; dots = individual participant means; grey shapes = density estimates.

Table 3
Analyses of the choice of expressions in Experiment 3

	Estimate	*SE*	*z*	*p*
Intercept	-1.92	0.46	-4.13	< .001
Gender (ambiguous vs. unambiguous)	< 0.02	0.09	0.20	.838
Situational similarity (two-box vs. one-box)	-0.82	0.14	-5.86	< .001
Gender × Situational similarity	0.02	0.10	0.23	.815

In summary, speakers used fewer pronouns when both referential candidates could move (two-box condition) compared to when only the target could move (one-box condition), demonstrating that the referents' non-linguistic competition, arising from their situational similarity, results in fewer pronouns. Critically, the effect did not interact with the referents' grammatical gender congruence, and unlike in Experiment 2, speakers did not avoid grammatical gender ambiguous pronouns in any condition. These results provided support for the non-linguistic competition account that speakers preferentially resolve competition based on the information available at the non-linguistic level. Speakers are more likely to choose a pronoun (rather than a noun) when the referential competitor could not be the referent than

otherwise, but once they have chosen to use a pronoun, they do not bother to check the grammatical gender ambiguity of the pronoun for the referential competitor that could not be the referent anyway. Likewise, speakers favor a noun (rather than a pronoun) more when the competitor could also be the referent than otherwise, but they do not check whether the grammatical gender of a pronoun, an alternative referring expression, would cause ambiguity.

We may wonder if speakers exclusively focus on the box manipulation. Other data from our lab, however, show that the box manipulation does not eliminate the gender congruence effect with human referents in English: Speakers use fewer pronouns when the referents have the same gender than otherwise, independent of the box manipulation (Fukumura et al., in press), in keeping with findings that the human referents' situational congruence and gender congruence independently affect pronoun use. That is, *he gets off the horse* is less ambiguous when the referential competitor is standing and unable to get off the horse, than when the competitor is also sitting on the horse and thus able to get off the horse. Nevertheless, the referents' situational congruence does not interact with the pronoun's gender ambiguity; speakers of English use fewer pronouns when the referential alternative has the same gender than otherwise, regardless of the referents' situational congruence (Fukumura et al., 2011; 2013). Hence, grammatical gender ambiguity was ignored in the current experiment, because, unlike gender congruence between human referents, grammatical gender congruence between non-human referents could not be determined at the non-linguistic level when non-linguistic competition strongly biased the speaker's choice of referring expressions.

4. General Discussion

The goal of the current study was to determine whether and how speakers avoid gender ambiguous pronouns in French. Let us first recap our findings. **Table 4** summarizes the key effects. Experiment 1 showed that speakers used fewer pronouns when the human referential

candidates had the same gender than when they had different genders, but grammatical gender congruence between the unrelated inanimate referents did not result in fewer pronouns. These findings disconfirmed the gender ambiguity avoidance account, according to which speakers should avoid gender ambiguous pronouns, regardless of whether the referents' non-linguistic similarity is increased by their gender congruence. By contrast, the results were compatible with the non-linguistic competition account, which predicts that speakers choose referring expressions based on non-linguistic information, such that the referents' gender congruence results in fewer pronouns only when it affects non-linguistic competition.

Table 4
Summary of key findings

Experiment 1
Main effect of gender congruence
Main effect of referent (humanness)
Gender congruence effect for human referents, but no grammatical gender congruence effect for inanimate referents
→ Consistent with the non-linguistic competition account: Pronoun choice is primarily affected by the referents' non-linguistic similarity
Experiment 2
Main effect of grammatical gender congruence
Main effect of semantic similarity
Grammatical gender effect for similar non-human referents but not for dissimilar non-human referents
→ Consistent with a narrower version of the gender ambiguity avoidance account: Speakers avoid grammatical gender ambiguity when the referents are semantically related.
Experiment 3
No main effect of grammatical gender congruence
Main effect of situational similarity
No effect of grammatical gender congruence, irrespective of situational similarity
→ Consistent with the non-linguistic competition account: Competition in the non-linguistic context overrides grammatical gender ambiguity avoidance

Experiments 2 and 3 followed up the Experiment 1 results. Experiment 2 tested a narrower version of the gender ambiguity avoidance account that predicts that speakers are more likely to check for and avoid grammatical gender ambiguity when there is similarity-based

competition during pronoun selection. Consistent with this account, participants used fewer pronouns when the non-human referents had the same grammatical gender than otherwise only when the referents had the same semantic categories. Experiment 3 examined the robustness of the Experiment 2 finding by pitching the referents' competition in the non-linguistic context against their grammatical gender congruence. Speakers used fewer pronouns when the referential competitor could also be the referent than when it could not, but they failed to avoid grammatical gender ambiguous pronouns, even though the referents were semantically related, indicating that grammatical gender ambiguity avoidance was overridden by the resolution of non-linguistic competition. Experiment 3 thus provided support for the non-linguistic competition account: Speakers preferentially choose referring expressions based on non-linguistic information; they do not engage in grammatical gender ambiguity avoidance when they can resolve competition at a non-linguistic level.

Our findings that speakers are more likely to avoid gender ambiguous pronouns for human referents than for non-human referents accord with those of Ferreira et al. (2005); as discussed earlier, they showed that speakers are more likely to avoid ambiguous bare nouns when the similarity that underlies the ambiguity can be represented non-linguistically (e.g., two flying bats for *bat*) than when it can only be determined linguistically (e.g., flying bat and baseball bat for *bat*). Hence, gender ambiguity for human referents may be a type of non-linguistic ambiguity arising from the referents' shared non-linguistic attribute (e.g., *being a male*), which can be conceptually associated with specific pronouns (e.g., *being a male* for the concept of IL in French) (see Fig.1). Consistent with this, Fukumura et al. (2013) showed that when the human referential candidates have the same gender, speakers reduce pronoun use more in English than in Finnish, whilst the referents' situational similarity (e.g., both are sitting on a horse or only one of them is sitting on a horse) reduces the pronoun rates to a similar extent in both languages. Whereas neither English nor Finnish pronouns express the referents' situational

attributes, only English pronouns express the referent's gender; the presence of a same-gender competitor renders the use of pronouns ambiguous in English, whereas Finnish pronouns are ambiguous regardless. Hence, the larger gender effect in English than in Finnish may suggest that gender congruence of the human referents affects ambiguity avoidance as well as competition.

By contrast, grammatical gender ambiguity exemplifies linguistic ambiguity, arising from the syntactic feature shared between the linguistic antecedents: Whereas gender ambiguous pronouns for human referents can be avoided at a non-linguistic/message level, where speakers decide whether to use a pronoun or not, grammatical gender ambiguity for non-human referents cannot be avoided at a non-linguistic level. Importantly, our results suggest that speakers can avoid grammatical gender ambiguous pronouns at least under some circumstances; they use fewer pronouns when the semantically related non-human referents share the same grammatical gender than otherwise (Experiment 2). We have suggested that when the referents are semantically related, the referential competitor interferes more during pronoun selection, and this interference increases the chance of speakers modelling the potential ambiguity of the to-be-produced pronoun. However, the absence of a grammatical gender ambiguity effect in the presence of non-linguistic competition (Experiment 3) underscores a fragility of this effect. Grammatical gender ambiguity avoidance requires an additional check at the lexical level; when speakers have chosen referring expressions to resolve competition at a non-linguistic level, they do not attempt to backtrack the choice made earlier, even when the referents are semantically related and can compete during pronoun selection processes.

Researchers have argued that when speakers fail to avoid linguistic ambiguity, they tend not to be aware of it; if they were, they avoid it (Ferreira et al., 2005; Ferreira & Dell, 2000; Snedeker & Trueswell, 2003). For instance, Ferreira and Dell (2000) found that speakers do not include a sentence complementizer, *that*, to avoid temporal ambiguity as in *The coach knew (that)*

you were…, arguing that speakers are unaware of or unconcerned with the temporal syntactic ambiguity. Snedeker and Trueswell, (2003) showed that speakers use prosody to avoid syntactic ambiguous instructions only when the context does not disambiguate (see also Kraljic & Brennan, 2005), arguing that when the context allows only one interpretation, speakers are unaware that an alternative interpretation is possible. Hence, our participants mostly failed to avoid grammatical gender ambiguous pronouns, possibly because they were unaware of the pronoun's grammatical gender ambiguity. Gender congruence between human referents could result in fewer pronouns *without* the speaker's awareness of the pronoun's ambiguity, however; the similarity between the referents of the same gender increases competition at a non-linguistic level, the resolution of which would favor the use of more explicit expressions (Fukumura et al., 2013), though it may also be that gender ambiguity for human referents can be easily inferred based on the conceptual association between the shared non-linguistic feature and the concept of the pronoun (cf. Ferreira et al., 2005). By contrast, grammatical gender congruence cannot affect pronoun choice via non-linguistic competition, nor can it be detected conceptually. Yet, our participants rarely employed the mechanisms that enabled them to detect the grammatical gender ambiguity of pronouns linguistically (*except* when the referents were semantically related and there was no salient competition in the non-linguistic context). These results therefore indicate that speakers' pronoun choice is largely led by the production-internal constraints that favor the use of the earliest-available information for linguistic choice.

During the reviewing process, it was pointed out that the non-linguistic competition account resonates with *self-organized sentence processing models* proposed by Tabor and colleagues (e.g., Tabor & Hutchins, 2004), which explain how similarity increases competition between relevant syntactic units in the process of structure formation; see Villata et al. (2018) and Villata and Franck (2020), who adopted the models to account for agreement errors. Importantly, these models do not explain how speakers choose between a pronoun and a noun

nor distinguish between non-linguistic, semantic, and grammatical features. Hence, although the models may explain how speakers make agreement errors, they do not provide straightforward explanations as to why gender congruence between the human referents leads to fewer pronouns, whilst grammatical gender congruence between the inanimate referents does not, and why the referents' competition in the non-linguistic context results in fewer pronouns and yet it does not facilitate grammatical gender ambiguity avoidance. By assuming that speakers preferentially choose referring expressions based on the referents' non-linguistic information, the non-linguistic competition account explains why speakers reliably reduce pronoun usage when the human referents share the same gender, but not when non-human referents share the same grammatical gender. The non-linguistic competition account can also explain why the referents' competition in the non-linguistic context results in fewer pronouns whilst it does not facilitate grammatical gender ambiguity avoidance.

In the current study, the antecedent was always the grammatical subject, as in many other pronoun production studies (e.g., Arnold & Griffin, 2007; Fukumura et al., 2011; 2013; Jescheniak et al., 2001; Meyer & Bock, 1999; Schmitt et al., 1999), so participants would sometimes produce pronouns. Some may wonder whether grammatical gender ambiguity did not influence the pronoun use very often, because gender ambiguity could be resolved based on the pronoun's preference for the subject antecedent; when the antecedent is highly prominent, its gender ambiguity with a less prominent non-subject may not be that critical. The rates of pronouns were below ceiling in the current study. The negative intercepts in fact indicated that repeated nouns were generally more favored. Though this may reflect the partcipants' general tendency to try to be precise, if speakers had taken the referent as being highly prominent, they would have used pronouns more frequently. Most critically, if speakers primarily relied upon the referent's position for their pronoun choice, it is not clear why they avoided gender ambiguous pronouns in some conditions but not in other conditions, and why the referents' situational congruence affected the

use of pronouns. We also found no evidence that speakers would avoid gender ambiguous pronouns more when the referential competitor competes more strongly in the referential context: Whilst participants generally reduced the rates of pronouns when the competitor could also be the referent than otherwise, they failed to avoid grammatical-gender ambiguous pronouns regardless (Experiment 3).

An interesting question for future research may concern the role of the addressee. First, in the current study, the addressee (confederate) was asked to identify the referent based on the subject preference when participants produced gender ambiguous pronouns. But what would happen if the addressee misinterpreted them? One possibility is that speakers become aware of the pronoun's gender ambiguity, and they learn to avoid it. However, in Experiment 1, whilst participants were able to avoid gender ambiguous pronouns for human referents, they failed to do so when the pronoun agreed with the grammatical gender of the inanimate referential candidates. Hence, an alternative possibility is that when communication breakdowns, speakers fail to 'work out' the grammatical gender ambiguity of pronouns; instead of avoiding gender ambiguous pronouns, they become generally more explicit by using more nouns. Second, we may ask to which extent the similarity effects were based on the speaker's own knowledge only or on their assumptions of the addressee's knowledge. In Experiment 3, situational similarity was varied for both the speaker and the addressee. Did speakers use fewer pronouns (i.e., more nouns) in the two-box context than in the one-box context because they took into account the addressee's knowledge or primarily because they responded to competition in their own visual display? Fukumura and Van Gompel (2012) showed that speakers' pronoun choice is driven by their own discourse model; participants used fewer pronouns and more nouns when the immediately preceding sentence mentioned a referential competitor than otherwise, regardless of whether the preceding sentence was shared with their addressee, though Fukumura (2015) showed that speakers can adapt pronoun use to the addressee's discourse model when it is

simpler for them to do; speakers use fewer pronouns when the addressee did not hear the sentence that mentioned the referent than when they did. Hence, although it is possible that speakers used their own visual display as a proxy to their addressee's (rather than taking the addressee's visual display into account) in the current study, perspective-taking might be easier with non-linguistic or visual context than with linguistic context. Thus, future research might wish to determine whether the shared status of the box manipulation influences pronoun use.

Finally, although language production models typically assume that grammatical gender is represented at the grammatical level, distinct from the noun's semantic representation, some research suggests that speakers of grammatical gender languages tend to associate grammatical gender with the semantics of conceptual gender (or sex) in various judgements, such as trait attribution (Boroditsky et al., 2003; Konishi, 1993), voice assignment (e.g., Flaherty, 2001; Ramos & Roberson, 2010; Sera et al., 1994; Sera et al., 2002), face categorization (Sato & Athanasopoulos. 2018), and similarity rating (Phillips & Boroditsky, 2003) tasks. For instance, Konishi (1993) reported that both Spanish and German participants rated masculine inanimate nouns as higher in potency than feminine inanimate nouns in their native language. However, questions have been raised concerning whether and the extent to which the attribution of male-like or female-like characteristics to asexual objects in these studies results from the strategic adoption of the participants' grammatical knowledge (e.g., Herold, 1982). Although grammatical gender can influence some cognitive judgements, we cannot be certain if grammatical gender alters the speaker's non-linguistic representations (cf. Boroditsky et al., 2003; Phillips & Boroditsky, 2003). In the current study, the referents' grammatical gender congruence alone did not influence the use of pronouns in French. Hence, even if grammatical gender congruence of non-human entities makes them semantically more similar at some level under some tasks, it does not do so at the level where speakers choose a pronoun or noun.

To conclude, we showed that speakers are more likely to avoid gender ambiguous pronouns when referring to human referents than when referring to non-human referents, providing support for the account that speakers are more successful in avoiding ambiguous referring expressions when the ambiguity can be represented at a non-linguistic level. Although speakers can avoid grammatical gender ambiguity for non-human referents, this depends on whether the referents are semantically related *and* when there is no strong non-linguistic competition that biases speakers' choice of using a pronoun or noun. Overall, pronoun choice is most likely to be affected by non-linguistic information that is available before production processes commence, rather than by the ambiguity of the pronoun that requires speakers to carry out the linguistic analysis of the chosen referring expression.

Acknowledgements

We thank William Woodside for implementing the experiments and Stéphane Dufau for his help with setting up the experiments in Marseille and Solène Gibaud, Estelle Hervé, Suzanne Lesage Broyelle, Quentin Denigot, Célia Richy, Clélie de Witasse, Cécile Reynaud, Elisa Gavard, Ghais Mokhtar-Slimane, Mélody Zira, for their help with data collection and/or transcription. We also thank the Laboratoire de Linguistique Formelle, University of Paris, for kindly letting us running experiments in their lab. This work was supported by the Leverhulme Trust [RPG-2016-253]. F-Xavier Alario was supported by grants ANR-16-CONV-0002 (ILCB) and the Excellence Initiative of Aix-Marseille University (A*MIDEX).

Appendix A: Materials

Experiment 1

For each item, the first line represents the human referent condition and the second line represents the inanimate referent condition, each involving a same-gender competitor before the slash and a different-gender competitor after the slash.

Le prince à côté du maçon/de la caissière est sur le numéro 3.
Le lacet à côté du melon/de la poire est sur le numéro 3.

Le marin au dessus du roi/de la reine est sur le numéro 2.
Le pain au dessus du tournesol/de la rose est sur le numéro 2.

La fille à côté de la grand-mère/du grand-père est sur le numéro 1.
La pomme à côté de la cigarette/du cigare est sur le numéro 1.

Le maçon en dessous du prince/de la princesse est sur le numéro 4.
Le melon en dessous du lacet/de la chaussure est sur le numéro 4.

La secrétaire en dessous de la châtelaine/du chevalier est sur le numéro 6.
La fraise en dessous de la chaise/du lit est sur le numéro 6.

Le soldat au dessus du footballeur/de la majorette est sur le numéro 3.
Le saxophone au dessus du gant/de la ceinture est sur le numéro 3.

Le roi au dessus du marin/de la fée est sur le numéro 1.
Le tournesol au dessus du pain/de la banane est sur le numéro 1.

La coiffeuse à côté de la mariée/du marié est sur le numéro 1.
La chemise à côté de la lampe/du tapis est sur le numéro 1.

Le boxeur au dessus du shérif/de la voyante est sur le numéro 2.
Le seau au dessus du violon/de la guitare est sur le numéro 2.

Le sorcier à côté du pompier/de la fleuriste est sur le numéro 6.

Le verre à côté du coussin/de la couverture est sur le numéro 6.

La grand-mère en dessous de la fille/du garçon est sur le numéro 5.
La cigarette en dessous de la pomme/du radis est sur le numéro 5.

Le pompier à côté du sorcier/de la sorcière est sur le numéro 2.
Le coussin à côté du verre/de la bouteille est sur le numéro 2.

La châtelaine en dessous de la secrétaire/du bagagiste est sur le numéro 5.
La chaise en dessous de la fraise/du poivron est sur le numéro 5.

Le shérif à côté du boxeur/de la danseuse est sur le numéro 6.
Le violon à côté du seau/de la poêle est sur le numéro 6.

Le prêtre à côté du mécanicien/de la maquilleuse est sur le numéro 5.
Le marteau à côté du fromage/de la citrouille est sur le numéro 5.

Le footballeur au dessus du soldat/de la gouvernante est sur le numéro 1.
Le gant au dessus du saxophone/de la trompette est sur le numéro 1.

La mariée à côté de la coiffeuse/du barbier est sur le numéro 4.
La lampe à côté de la chemise/du pull est sur le numéro 4.

La couturière au dessus de la nymphe/du général est sur le numéro 3.
La pêche au dessus de la fourchette/du couteau est sur le numéro 3.

Le mécanicien à côté du prêtre/de la religieuse est sur le numéro 4.
Le fromage à côté du marteau/de la scie est sur le numéro 4.

La nymphe à côté de la couturière/du mineur est sur le numéro 5.
La fourchette à côté de la pêche/du kiwi est sur le numéro 5.

La reine en dessous de la fée/du marin est sur le numéro 6.
La rose en dessous de la banane/du pain est sur le numéro 6.

Le grand-père en dessous du garçon/de la fille est sur le numéro 4.
Le cigare en dessous du radis/de la pomme est sur le numéro 4.

La princesse à côté de la caissière/du maçon est sur le numéro 5.
La chaussure à côté de la poire/du melon est sur le numéro 5.

La caissière à côté de la princesse/du prince est sur le numéro 2.
La poire à côté de la chaussure/du lacet est sur le numéro 2.

La fée en dessous de la reine/du roi est sur le numéro 4.
La banane en dessous de la rose/du tournesol est sur le numéro 4.

Le bagagiste à côté du chevalier/de la châtelaine est sur le numéro 2.
Le poivron à côté du lit/de la chaise est sur le numéro 2.

La sorcière en dessous de la fleuriste/du pompier est sur le numéro 4.
La bouteille en dessous de la couverture/du coussin est sur le numéro 4.

Le général à côté du mineur/de la couturière est sur le numéro 1.
Le couteau à côté du kiwi/de la pêche est sur le numéro 1.

La voyante en dessous de la danseuse/du boxeur est sur le numéro 5.
La guitare en dessous de la poêle/du seau est sur le numéro 5.

La danseuse à côté de la voyante/du shérif est sur le numéro 6.
La poêle à côté de la guitare/du violon est sur le numéro 6.

La gouvernante au dessus de la majorette/du footballeur est sur le numéro 1.
La trompette au dessus de la ceinture/du gant est sur le numéro 1.

Le mineur en dessous du général/de la nymphe est sur le numéro 6.
Le kiwi en dessous du couteau/de la fourchette est sur le numéro 6.

La religieuse à côté de la maquilleuse/du mécanicien est sur le numéro 3.
La scie à côté de la citrouille/du fromage est sur le numéro 3.

La maquilleuse au dessus de la religieuse/du prêtre est sur le numéro 1.
La citrouille au dessus de la scie/du marteau est sur le numéro 1.

La fleuriste en dessous de la sorcière/du sorcier est sur le numéro 4.
La couverture en dessous de la bouteille/du verre est sur le numéro 4.

Le barbier à côté du marié/de la mariée est sur le numéro 3.
Le pull à côté du tapis/de la lampe est sur le numéro 3.

Le chevalier au dessus du bagagiste/de la secrétaire est sur le numéro 2.
Le lit au dessus du poivron/de la fraise est sur le numéro 2.

La majorette en dessous de la gouvernante/du soldat est sur le numéro 4.
La ceinture en dessous de la trompette/du saxophone est sur le numéro 4.

Le marié à côté du barbier/de la coiffeuse est sur le numéro 3.
Le tapis à côté du pull/de la chemise est sur le numéro 3.

Experiment 2

For each item, the first line represents the semantically similar condition and the second line represents the semantically dissimilar condition, each involving a same-gender competitor before the slash and a different-gender competitor after the slash.

La fraise à côté de la pastèque/du melon est sur le numéro 2.
La fraise à côté de la guitare/du violon est sur le numéro 2.

La betterave à côté de la courgette/du navet est sur le numéro 3.
La betterave à côté de la voiture/du camion est sur le numéro 3.

Le pichet à côté du verre/de la tasse est sur le numéro 6.
Le pichet à côté du gorille/de la girafe est sur le numéro 6.

La bouteille au dessus de la tasse/du verre est sur le numéro 2.
La bouteille au dessus de la colombe/du pigeon est sur le numéro 2.

Le brocoli en dessous du navet/de la courgette est sur le numéro 6.
Le brocoli en dessous du gant/de la veste est sur le numéro 6.

Le bus à côté du camion/de la voiture est sur le numéro 2.
Le bus à côté du couteau/de la cuillère est sur le numéro 2.

Le chapeau en dessous du gant/de la veste est sur le numéro 5.
Le chapeau en dessous du verre/de la tasse est sur le numéro 5.

La chaussure à côté de la veste/du gant est sur le numéro 1.
La chaussure à côté de la dent/du nez est sur le numéro 1.

Le citron au dessus du melon/de la pastèque est sur le numéro 3.
Le citron au dessus du short/de la jupe est sur le numéro 3.

La flûte à côté de la guitare/du violon est sur le numéro 5.
La flûte à côté de la pastèque/du melon est sur le numéro 5.

Le hibou à côté du pigeon/de la colombe est sur le numéro 5.
Le hibou à côté du navet/de la courgette est sur le numéro 5.

Le lion au dessus du gorille/de la girafe est sur le numéro 1.
Le lion au dessus du melon/de la pastèque est sur le numéro 1.

La fourchette au dessus de la cuillère/du couteau est sur le numéro 3.
La fourchette au dessus de la veste/du gant est sur le numéro 3.

La main en dessous de la dent/du nez est sur le numéro 4.
La main en dessous de la jupe/du short est sur le numéro 4.

La moto au dessus de la voiture/du camion est sur le numéro 1.
La moto au dessus de la girafe/du gorille est sur le numéro 1.

La mouette à côté de la colombe/du pigeon est sur le numéro 4.

La mouette à côté de la cuillère/du couteau est sur le numéro 4.

Le pantalon en dessous du short/de la jupe est sur le numéro 6.
Le pantalon en dessous du violon/de la guitare est sur le numéro 6.

La panthère en dessous de la girafe/du gorille est sur le numéro 5.
La panthère en dessous de la courgette/du navet est sur le numéro 5.

Le piano à côté du violon/de la guitare est sur le numéro 1.
Le piano à côté du nez/de la dent est sur le numéro 1.

Le pied au dessus du nez/de la dent est sur le numéro 1.
Le pied au dessus du camion/de la voiture est sur le numéro 1.

La robe au dessus de la jupe/du short est sur le numéro 1.
La robe au dessus de la tasse/du verre est sur le numéro 1.

Le décapsuleur en dessous du couteau/de la cuillère est sur le numéro 4.
Le décapsuleur en dessous du pigeon/de la colombe est sur le numéro 4.

Le tournevis en dessous du clou/de la vis est sur le numéro 6.
Le tournevis en dessous du cochon/de la chèvre est sur le numéro 6.

La perceuse à côté de la vis/du clou est sur le numéro 5.
La perceuse à côté de la tulipe/du tournesol est sur le numéro 5.

La pomme à côté de la banane/du kiwi est sur le numéro 5.
La pomme à côté de la chèvre/du cochon est sur le numéro 5.

La vache à côté de la chèvre/du cochon est sur le numéro 3.
La vache à côté de la fenêtre/du rideau est sur le numéro 3.

La porte à côté de la fenêtre/du rideau est sur le numéro 3.
La porte à côté de la trompette/du saxophone est sur le numéro 3.

Le mouton à côté du cochon/de la chèvre est sur le numéro 3.
Le mouton à côté du saxophone/de la trompette est sur le numéro 3.

Le volet au dessus du rideau/de la fenêtre est sur le numéro 3.
Le volet au dessus du céleri/de la carotte est sur le numéro 3.

La libellule au dessus de la coccinelle/du papillon est sur le numéro 2.
La libellule au dessus de la carotte/du céleri est sur le numéro 2.

Le scarabée en dessous du papillon/de la coccinelle est sur le numéro 4.
Le scarabée en dessous du clou/de la vis est sur le numéro 4.

La tortue en dessous de la baleine/du dauphin est sur le numéro 4.
La tortue en dessous de la vis/du clou est sur le numéro 4.

La harpe à côté de la trompette/du saxophone est sur le numéro 5.
La harpe à côté de la banane/du kiwi est sur le numéro 5.

Le banjo à côté du saxophone/de la trompette est sur le numéro 6.
Le banjo à côté du tournesol/de la tulipe est sur le numéro 6.

Le requin à côté du dauphin/de la baleine est sur le numéro 5.
Le requin à côté du fauteuil/de la commode est sur le numéro 5.

Le bureau en dessous du fauteuil/de la commode est sur le numéro 5.
Le bureau en dessous du papillon/de la coccinelle est sur le numéro 5.

La table au dessus de la commode/du fauteuil est sur le numéro 3.
La table au dessus de la baleine/du dauphin est sur le numéro 3.

Le poireau au dessus du céleri/de la carotte est sur le numéro 2.
Le poireau au dessus du canard/de la cigogne est sur le numéro 2.

La tomate en dessous de la carotte/du céleri est sur le numéro 5.
La tomate en dessous de la coccinelle/du papillon est sur le numéro 5.

Le pamplemousse à côté du kiwi/de la banane est sur le numéro 4.
Le pamplemousse à côté du rideau/de la fenêtre est sur le numéro 4.

La poule au dessus de la cigogne/du canard est sur le numéro 2.

La poule au dessus de la commode/du fauteuil est sur le numéro 2.

Le pingouin à côté du canard/de la cigogne est sur le numéro 2.
Le pingouin à côté du kiwi/de la banane est sur le numéro 2.

La rose à côté de la tulipe/du tournesol est sur le numéro 2.
La rose à côté de la cigogne/du canard est sur le numéro 2.

Le coquelicot en dessous du tournesol/de la tulipe est sur le numéro 6.
Le coquelicot en dessous du dauphin/de la baleine est sur le numéro 6.

Appendix B: Bayesian analyses

Each model was run with 4 chains with 6000 iterations by chain. Because there were no previous studies on grammatical gender ambiguity avoidance, we had no information about the distributions of the parameters. We thus used only weakly informative priors of *Normal* ($\mu = 0$, $\sigma = 10$), assuming a normal distribution with mean zero and a large standard deviation on logit scale. Random effects included by-participants and by-items random intercepts and slopes for all predictors (Barr et al., 2013). We report the estimated mean ($\hat{\beta}$), the range (95% credible), and the probability of the parameter effect ($P(\beta)$) being smaller than (for negative estimates) or greater than (for positive estimates) zero. Following Engelmann et al. (2019), we will interpret $P(\beta) \geq .95$ as strong evidence for the effect and $P(\beta) \geq .85$ as weak evidence.

Experiment 1

Fig.9 shows the posterior distributions of the parameters. The output from the main model revealed that the probabilities of the effect of gender ($\hat{\beta} = -0.52$, 95% CrI = [-0.91, -0.07]) and the interaction ($\hat{\beta} = -1.54$, 95% CrI = [-2.37, -0.75]) being smaller than zero were .986 and over .999, respectively. In addition, the probability of the effect of referent (animacy) being larger than zero was .998. The interaction reflected the fact that, whereas the probability of the gender effect for the human referents being smaller than zero was over .999 ($\hat{\beta} = -1.25$, 95% CrI = [-1.85, -0.63]), we observed little evidence for the effect of gender ambiguity for the inanimate referents ($\hat{\beta} = 0.27$, 95% CrI = [-0.25, 0.92]), $P(\beta > 0) = .831$), where, if anything, the effect went into the opposite direction (see **Table** 5 for a full summary).

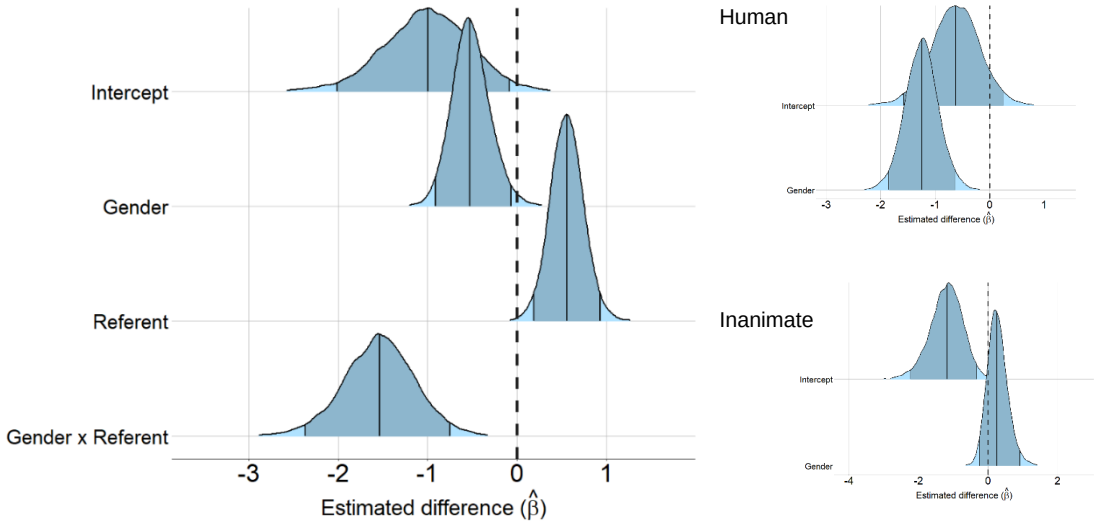

Fig.9. Posterior distributions of parameters for pronoun choice in Experiment 1.

Table 5
Summary of the Bayesian analyses

Experiment 1	β	***P*(*β*)>**	***P*(*β*)<**	**Min**	**Max**	**L95**	**H95**
Intercept	-1.01	.017	.983	-3.27	0.94	-2.01	-0.09
Gender	-0.52	.014	.986	-1.25	0.50	-0.91	-0.07
Referent	0.56	.998	.002	-0.21	1.60	0.19	0.93
Gender × Referent	-1.54	< .001	> .999	-3.46	0.14	-2.37	-0.75
Gender ambiguity in the human referent condition							
Intercept	-0.63	.077	.923	-2.81	1.22	-1.58	0.25
Gender	-1.25	< .001	> .999	-2.59	0.08	-1.85	-0.63
Gender ambiguity in the inanimate referent condition							
Intercept	-1.20	.005	.995	-3.93	0.99	-2.23	-0.32
Gender	0.27	.831	.169	-0.69	2.56	-0.25	0.92
Experiment 2	β	***P*(*β*)>**	***P*(*β*)<**	**Min**	**Max**	**L95**	**H95**
Intercept	-1.61	.001	.999	-3.92	0.30	-2.70	-0.60
Gender	-0.49	.039	.961	-1.53	0.55	-1.00	0.06
Semantic similarity	-0.40	.018	.982	-1.15	0.51	-0.77	-0.03
Gender × similarity	-0.63	.056	.944	-2.78	1.42	-1.43	0.16
Gender ambiguity in the semantically similar condition							
Intercept	-1.86	< .001	> .999	-4.44	0.23	-3.01	-0.84
Gender	-0.78	.022	.978	-2.72	1.03	-1.52	-0.03
Gender ambiguity in the semantically dissimilar condition							
Intercept	-1.30	.007	.993	-3.69	0.48	-2.35	-0.29
Gender	-0.14	.316	.684	-1.72	1.43	-0.75	0.54
Experiment 3	β	***P*(*β*)>**	***P*(*β*)<**	**Min**	**Max**	**L95**	**H95**
Intercept	-2.07	< .001	> .999	-4.86	0.16	-3.24	-0.98
Gender	0.08	.622	.378	-1.16	1.24	-0.39	0.62
Box	-1.76	< .001	> .999	-3.74	-0.11	-2.60	-1.00
Gender × Box	0.17	.632	.368	-2.35	2.51	-0.91	1.20

Experiment 2

Fig. 10 reports the posterior distributions. The analyses revealed strong support for gender ($\hat{\beta}$ = -0.49, 95% CrI = [-1.00, 0.06]) and semantic similarity ($\hat{\beta}$ = -0.40, 95% CrI = [-0.77, -0.03]), as the probabilities of these main effects being smaller than zero were .961 and .982, respectively. The probability of the interaction ($\hat{\beta}$ = -0.63, 95% CrI = [-1.43, 0.16]) being smaller than zero was .944. Follow-up analyses revealed strong evidence for gender in the semantically similar condition ($\hat{\beta}$ = -0.78, 95% CrI = [-1.52, -0.03], $P(\beta < 0)$ = .978) and very weak evidence for gender in the semantically dissimilar condition ($\hat{\beta}$ = -0.14, 95% CrI = [-0.75, 0.54], $P(\beta < 0)$ = .684) (see **Table 5** for a full summary).

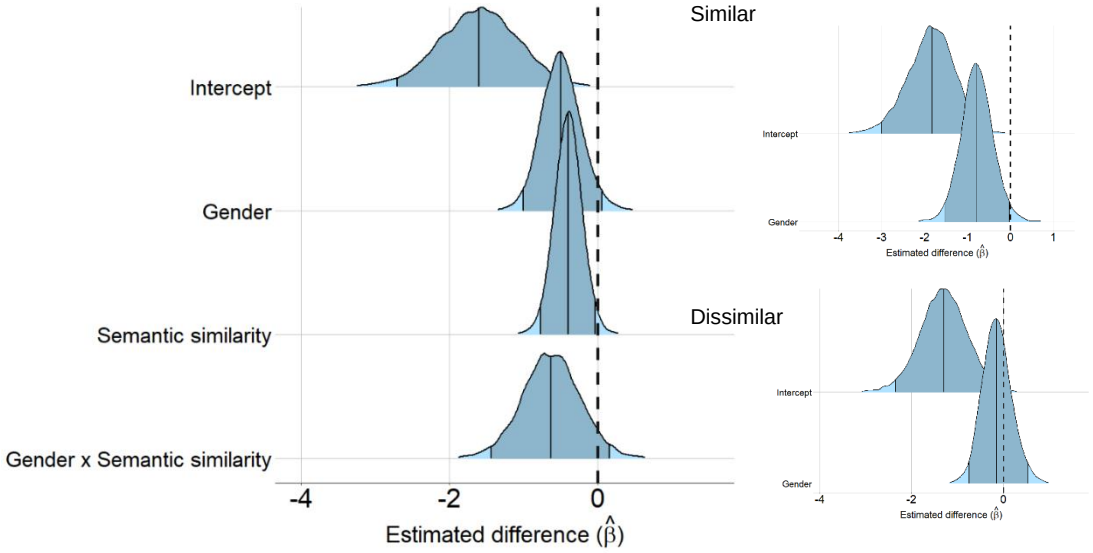

Fig. 10. Posterior distributions of parameters for pronoun choice in Experiment 2.

Experiment 3

Fig. 11 reports the posterior distributions, and see the main manuscript for the analyses.

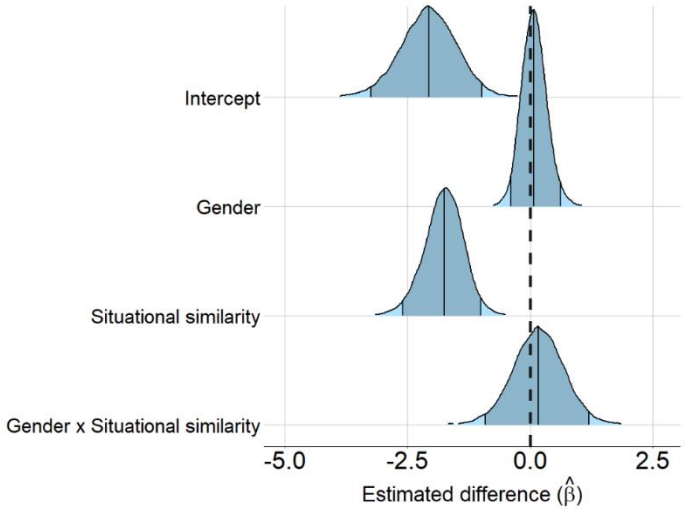

Fig. 11. Posterior distributions of parameters for pronoun choice in Experiment 3.

Appendix C: Experiment 4 (Replication of Experiment 1)

As footnoted in Experiment 1, unpublished findings from our lab suggested that gender congruence between unrelated inanimate referents led to fewer pronouns as much as gender congruence between human referents did. We therefore ran an additional experiment to assess the generalizability of Experiment 1 findings. We used the same human referents as in Experiment 1. The non-human referents were adopted from another experiment, where we found neither the referents' phonological similarity nor grammatical gender congruence affected pronoun use.

Method

Participants. Thirty-two native speakers of French were recruited from the same population as before, in exchange for course credits or cash.

Materials and procedure. There were 40 experimental items. In the *human referent* condition (7a-7b), the target and competitor were both humans, and in the *non-human referent* condition (7c-7d), they were non-human with grammatical gender. In both conditions, the target and competitor either had the same gender or differed in gender, and we examined participants' choice of referring expressions when they produced the target descriptions in (8). The materials in the human referent condition were the same as in Experiment 1. For the non-human referent condition, we used a different set of target and competitor pairs. Four items had animate targets (dove, pigeon, pig, and cockroach) and 36 items had inanimate targets. The target and competitors were always unrelated both semantically and phonologically (as was the case in the dissimilar condition in Experiment 3). As before, each competitor occurred in the both same gender and different gender conditions. The procedure was the same as before.

(7) Context sentences

a. Le marin au dessus du roi est sur le numéro 2. *The sailor above the king is on No. 2.*

b. Le marin au dessus de la reine est sur le numéro 2. *The sailor above the queen is on No. 2.*

c. Le baril au dessus du collier est sur le numéro 2. *The barrel above the necklace is on No. 2.*

d. Le baril au dessus de la comète est sur le numéro 2. *The barrel above the comet is on No. 2.*

(8) Target responses

a. Maintenant {le marin / il} est sur le numéro 3. *Now {the sailor / he} is on Number 3.*

b. Maintenant {le baril / il} est sur le numéro 3. *Now {the barrel / it} is on Number 3.*

Design. We used a 2 (referent: human vs. non-human) × 2 (gender: ambiguous vs. unambiguous) repeated measures design, resulting in the creation of four lists. Together with 80 filler items, 40 experimental items were distributed in a fixed quasi-random order as in Experiment 1. Eight participants were randomly assigned to each list.

Scoring. We scored whether participants produced a pronoun (n = 410) or a repeated noun (n = 834) phrase to refer to the target as before. We excluded cases with other nouns than repeated nouns (n = 9); a pronoun with a wrong gender (n = 1) or a noun with a wrong gender determiner (n = 6); substitution errors (n = 1); omitted subjects (n = 3), changed responses (n = 6); dislocations (n = 8); and technical errors (n = 2). In total, 36 trials (2.8%) were excluded and 1244 responses were analyzed.

Results

Fig.12 presents the means and **Table 6** reports the results. The analyses included gender (coded as before) and referent (human referent coded as 1 and non-human referent coded as 0). The main effect of gender ambiguity found fewer pronouns in the same gender condition (28.0%) than in the different gender condition (37.9%). There was no significant main effect of referent or a significant gender × referent interaction. The referents' gender ambiguity led to fewer (16.8%) pronouns for human referents, but not for non-human referents.

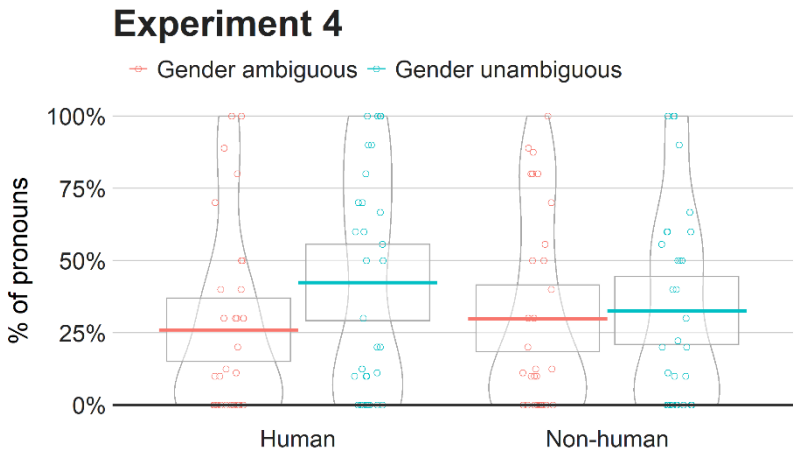

Fig.12. Percentages of pronouns out of pronouns and nouns in Experiment 4. Lines = condition means; boxes = 95% confidence intervals; dots = individual participant means; grey shapes = density estimates

Table 6

Analyses of the choice of expressions in Experiment 4

	Estimate	*SE*	*z*	*p*
Main analysis				
Intercept	-1.64	0.55	-2.95	.003
Gender (ambiguous vs. unambiguous)	-0.45	0.09	-4.71	< .001
Referent (human vs. non-human)	0.12	0.08	1.36	.175
Gender × Referent	-0.30	0.09	-3.49	< .001
Gender effect in the human referent condition				
Intercept	-1.50	0.59	-2.53	.011
Gender	-0.75	0.15	-4.88	< .001
Gender effect in the non-human referent condition				
Intercept	-1.65	0.51	-3.24	.001
Gender	-0.13	0.12	-1.10	.273

Fig. 13 reports the posterior distributions of the parameters. The results from the Bayesian analysis revealed that the probabilities of the effect of gender ambiguity ($\hat{\beta}$ = -0.93, 95% CrI = [-1.45, -0.44]) and the gender ambiguity x referent interaction ($\hat{\beta}$ = -1.26, 95% CrI = [-2.18, -0.37]) being smaller than zero were as large as .999 and .995, respectively. We found strong evidence for the effect of gender ambiguity for human referents; the probability of the gender effect being smaller than zero was over .999 ($\hat{\beta}$ = -1.56, 95% CrI = [-2.44, 0.72]). By contrast, the effect of

gender ambiguity for non-human referents was very weak ($\hat{\beta}$ = -0.26, 95% CrI = [-0.87, 0.37], $P(\beta > 0)$ = .804). The analysis also showed some weak evidence for the effect of referent ($\hat{\beta}$ = 0.23, 95% CrI = [-0.22, 0.68], $P(\beta > 0)$ = .856), reflecting slightly more pronouns when the referents were human (34.4%) than when they were non-human (31.5%).

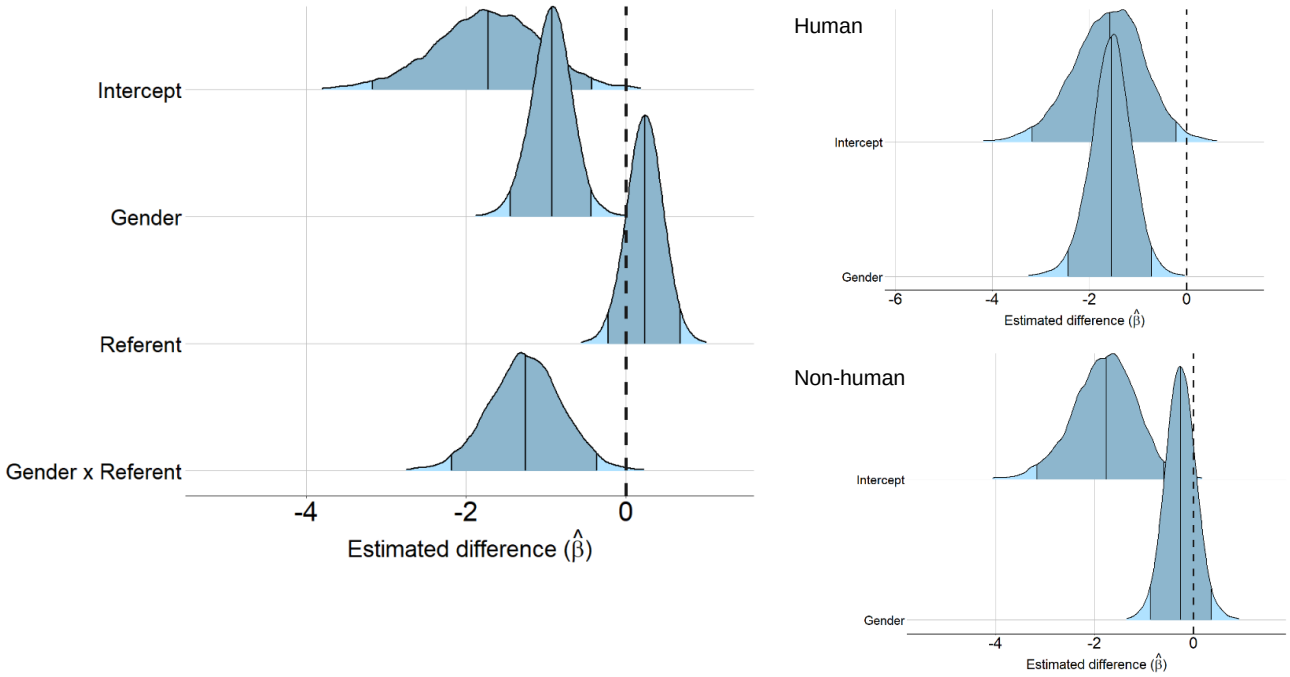

Fig.13. Posterior distributions of parameters for pronoun choice in Experiment 4**.**

Hence, whilst speakers reliably avoided gender ambiguous pronouns for human referents, this did not avoid gender ambiguous pronouns for unrelated non-human referents. The absence of a grammatical gender ambiguity effect for semantically unrelated entities was in accord with the findings reported in this manuscript, and the results confirmed that the unpublished findings mentioned earlier (i.e., the absence of gender × referent interaction) do not generalize. Most importantly, the results provide further support for the non-linguistic competition account: Speakers are more likely to avoid gender ambiguous pronouns for human referents whose gender can be determined non-linguistically, rather than for non-human referents whose gender can only be determined linguistically.

References

Abel, S., Dressel, K., Bitzer, R., Kummerer, D., Mader, I., Weiller, C., & Huber, W. (2009). The separation of processing stages in a lexical interference fMRI-paradigm. *NeuroImage, 44*, 1113–1124.

Aissen, J. (2003). Differential object marking: Iconicity vs. Economy. *Natural Language & Linguistic Theory, 21*, 435-483.

Allbritton, D. W., McKoon, G., & Ratcliff, R. (1996). Reliability of prosodic cues for resolving syntactic ambiguity. *Journal of experimental psychology. Learning, memory, and cognition, 22*, 714–735.

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

Arnold, J. E. (2001). The effect of thematic roles on pronoun use and frequency of reference continuation. *Discourse Processes, 31*, 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*, 521-536.

Arnold, J. E., Wasow, T., Asudeh, A., & Alrenga, P. (2004). Avoiding attachment ambiguities: The role of constituent ordering. *Journal of Memory and Language, 51*, 55-70.

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

Barker, J., Nicol, J., & Garrett, M. (2001). Semantic factors in the production of number agreement. *Journal of Psycholinguistic Research, 30*, 91–114.

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*, 255-278.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. arXiv preprint arXiv:1506.04967.

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

Belke, E., & Meyer, A. S. (2002). Tracking the time course of multidimensional stimulus discrimination: Analyses of viewing patterns and processing times during "same"-"different" decisions. *European Journal of Cognitive Psychology*, *14*, 237-266.

Bloem, I., & La Heij. (2003). Semantic facilitation and semantic interference in word translation: Implications for models of lexical access in language production, *Journal of Memory and Language*, *48*, 468-488.

Bock, J. K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review, 99*, 150–171.

Bock, J. K., Nicol, J., & Cutting, J. C. (1999). The ties that bind: Creating number agreement in speech. *Journal of Memory and Language, 40***,** 330–346.

Bock, K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology, 23*, 45–93.

Bock, K., Eberhard, K. M., & Cutting, J. C. (2004). Producing number agreement: How pronouns equal verbs. *Journal of Memory and Language, 51*, 251–278.

Bock, K., Eberhard, K. M., Cutting, J. C., Meyer, A. S., & Schriefers, H. (2001). Some attractions of verb agreement. *Cognitive Psychology, 43*, 83–128.

Boroditsky, L., Schmidt, L., & Phillips, W. (2003). Sex, syntax and semantics. In D. Gentner and S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp.61-79), Cambridge, MA: MIT Press.

Boroditsky, L., Schmidt, L., & Phillips, W. (2003). Sex, syntax and semantics. In D. Gentner and S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp.61-79), Cambridge, MA: MIT Press.

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

Brennan, S. E., Friedman, M. W., & Pollard, C. J. (1987). A centering approach to pronouns. In *Proceedings of the 25th annual meeting of the association for computational linguistics* (pp. 155-162). Cambridge, MA: Association for Computational Linguistics.

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology, 77*, 305–327.

Bürkner, P.-C. & Charpentier, E. (2020). Modelling monotonic effects of ordinal predictors in Bayesian regression models. *British Journal of Mathematical and Statistical Psychology*, 1-32.

Bürkner, P.-C. (2017). Brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software, 8*, 1–28.

Cacciari, C., Carreiras, M., & Barbolini Cionini, C. (1997). When words have two genders: Anaphor resolution for Italian functionally ambiguous words. *Journal of Memory and Language, 37*, 517532.

Cacciari, C., Corradini, P., Padovani, R., & Carreiras, M. (2011). Pronoun resolution in Italian: The role of grammatical gender and context. *Journal of Cognitive Psychology, 23*, 416-434.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology, 14*, 177-208.

Caramazza, A., Hillis, A.E. (1990). Where do semantic errors come from? *Cortex, 26*, 95–122.

Carreiras, M., Garnham, A., & Oakhill, J. (1993). The use of superficial and meaning-based representations in interpreting pronouns: Evidence from Spanish. *European Journal of Cognitive Psychology, 5*, 93-116.

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software, 76*, 1–32.

Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. N. Li (Ed.), *Subject and topic* (pp. 25–56). New York: Academic Press.

Clancy, P. M. (1980). Referential choice in English and Japanese narrative discourse. In W. Chafe (Ed.), *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production* (Vol. 3, pp. 127–201). Norwood, NJ: Ablex.

Cleland, A. A., & Pickering, M. J. (2003). The use of lexical and syntactic information in language production: Evidence from the priming of noun-phrase structure. *Journal of Memory and Language, 49*, 214–230.

Cohen, G. (1990). Why is it difficult to put names to faces? *British Journal of Psychology, 81*, 287–297.

Comrie, B., 1989. *Language Universals and Linguistic Typology*. Chicago: University of Chicago Press.

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

Cornish, F. (1987). Anaphoric pronouns: under linguistic control, or signalling particular discourse representations? *Journal of Semantics, 5*, 233-260.

Dahl, Ö., & Fraurud, K. (1996). Animacy in grammar and discourse. In T. Fretheim & J.K. Gundel (Eds.), *Reference and referent accessibility* (pp. 47-64). Amsterdam: John Benjamins.

Dale, R., & Reiter, E. (1995). Computational interpretation of the Gricean maxims in the generation of referring expressions. *Cognitive Science, 18,* 233-263.

Damian, M. F., & Bowers, J. S. (2003). Locus of semantic interference in picture–word interference tasks. *Psychonomic Bulletin & Review, 10,* 111-117.

Damian, M. F., & Martin, R. C. (1999). Semantic and phonological codes interact in single word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 345–361.

Deutsch, W., & Pechmann, T. (1982). Social interaction and the development of definite descriptions. *Cognition, 11*, 159-184.

Dhooge, E., & Hartsuiker, R. J. (2011). How do speakers resist distraction?: Evidence from a taboo picture-word interference task. *Psychological Science, 22*, 855–859.

Dhooge, E., & Hartsuiker, R. J. (2012). Lexical selection and verbal self-monitoring: Effects of lexicality, context, and time pressure in in picture-word interference. *Journal of Memory and Language, 66*, 163–176.

Eberhard, K. M., Cutting, J. C., & Bock, K. (2005). Making syntax of sense: Number agreement in sentence production. *Psychological Review, 112*, 531–559.

Engelhardt, P.E., Bailey, K., & F.Ferreira, F. (2006). Do speakers and listeners observe the Gricean Maxim of Quantity? *Journal of Memory and Language, 54*, 554-573.

Engelmann, F., Granlund, S., Kolak, J., Szreder, M., Ambridge, B., Pine, J., ... & Lieven, E. (2019). How the input shapes the acquisition of verb morphology: Elicited production and computational modelling in two highly inflected languages. *Cognitive psychology*, *110*, 30-69.

Ferreira, F. (1994). Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language, 33*, 715-736.

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*, 263-284.

Flaherty, M. (2001). How a language gender system creeps into perception. *Journal of Cross-Cultural Psychology, 32*, 18–31.

Fletcher, C. R. (1984). Markedness and topic continuity in discourse processing. *Journal of Verbal Learning and Verbal Behavior, 23*, 487-493.

Foley, W. A., & Van Valin, R.D. (1985). Information packaging in the clause. In T. Shopen (Ed.), *Language typology and syntactic description* (Vol. 1, pp. 282-364). Cambridge: Cambridge University Press.

Frank, M., & Goodman, N. (2012). Predicting pragmatic reasoning in language games. *Science, 336*, 998.

Fukumura, K. (2015). Interface between visual and linguistic information during audience design. *Cognitive Science, 39*, 1419–1433.

Fukumura, K. (2016). Development of audience design in children with and without ASD. *Developmental Psychology, 52*, 71-87.

Fukumura, K. (2018). Ordering adjectives in referential communication. *Journal of Memory and Language, 10*, 37-50.

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*, 52–66.

Fukumura, K., & Van Gompel, R.P.G. (2011). The effects of animacy in the choice of referring expressions. *Language and Cognitive Processes, 26*, 1472-1504.

Fukumura, K., & Van Gompel, R.P.G. (2012). Producing pronouns and definite noun phrases: Do speakers use the addressee's discourse model? *Cognitive Science, 36*, 1289-1311.

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*, 501-525

Fukumura, K., Hervé, C., Villata, S., Zhang, S., & Foppolo, F. (in press). Representations underling pronoun choice in Italian and English. *The Quarterly Journal of Experimental Psychology*. https://doi.org/10.1177/17470218211051989

Fukumura, K., Pozniak, C., & Alario, F.X. (2020). Representations underlining grammatical gender ambiguity avoidance. *Proceeding of the 26th AMLaP Architectures and Mechanisms of Language Processing*, Potsdam, Germany.

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*, 1012-1021.

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*, 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*, 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*, 41-62.

Gauvin, H. S., & Hartsuiker, R. J. (2020). Towards A New Model of Verbal Monitoring. *Journal of Cognition, 3*, 1–37.

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

Gennari, S. P., Mirkovic, J., & MacDonald, M. C. (2012). Animacy and competition in relative clause production: a cross-linguistic investigation. *Cognitive Psychology, 65*, 141–176.

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). Amsterdam: John Benjamins.

Gordon, P.C., Grosz, B.J. & Gilliom, L.A. (1993). Pronouns, names, and centering of attention is discourse. *Cognitive Science, 17,* 311-347.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics Vol. III: Speech acts* (pp. 41-58). New York: Academic Press.

Griffin, Z. M. (2010). Retrieving personal names, referring expressions, and terms of address. *Psychology of Learning and Motivation: Advances in Research and Theory*, *53*, 345–387.

Griffin, Z.M., & Wangerman, T. (2013). Parents Accidentally Substitute Similar Sounding Sibling Names More Often than Dissimilar Names. *PLoS ONE, 8,* 1-6.

Grosz, B. J., Joshi, A. K., & Weinstein, S. (1995). Centering: A framework for modelling the local coherence of discourse. *Computational Linguistics, 21*, 203-225.

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

Hartsuiker, R. J. & Kolk, H. H. J. (2001). Error monitoring in speech production: A computational test of the perceptual loop theory. *Cognitive Psychology, 42,* 113–57.

Haywood, S. L., Pickering, M. J., & Branigan, H. P. (2005). Do speakers avoid ambiguities during dialogue? *Psychological Science, 16,* 362–366.

Herold, A. L. (1982). Linguistic relativity: Transforming the relative into the insensible: A reply to Clarke et al. *Language Learning, 32,* 201–207.

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition, 59,* 91–117.

Jescheniak, J.D., Schriefers, H.J., & Hantsch, A. (2001). Semantic and phonological activation in noun and pronoun production. *Journal of Experimental Psychology: Learning, Memory and Cognition, 27,* 1058-1078.

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

Konishi, T (1993). The semantics of grammatical gender: A cross-cultural study. *Journal of Psycholinguistic Research, 22*, 519–534.

Kraljic, T., & Brennan, S. E. (2005). Prosodic disambiguation of syntactic structure: For the speaker or for the addressee? *Cognitive Psychology, 50*, 194–231.

Léger, L., Boumlak, H., & Tijus, C. (2008). BASETY: Extension et typicalité des exemplaires pour 21 catégories d'objets [BASETY: Extension and typicality of the specimens for 21 categories of objects]. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 62*, 223–232.

Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition, 14*, 41–104.

Levelt, W.J.M. (1989). *Speaking: From intention to articulation.* Cambridge, MA: MIT Press.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1-38.

Meyer, A. S., & Bock, K. (1999). Representations and processes in the production of pronouns: Some perspectives from Dutch. *Journal of Memory and Language, 41*, 281-301.

Misersky, J., Gygax, P. M., Canal, P., Gabriel, U., Garnham, A., Braun, F., et al. (2014). Norms on the gender perception of role nouns in Czech, English, French, German, Italian, Norwegian, and Slovak. *Behav. Res. Methods, 46*, 841–871.

Nozari, N., Dell, G. S., & Schwartz, M. F. (2011). Is comprehension the basis for error detection? A conflict-based theory of error detection in speech production. *Cognitive Psychology, 63*, 1–33.

Oberauer, K., & Kliegl, R. (2006). A formal model of capacity limits in working memory. *Journal of Memory and Language, 55*, 601–626.

Pechmann, T. (1989). Incremental speech production and referential overspecification. *Linguistics, 27,* 89-110.

Phillips, W., & Boroditsky, L. (2003). Can quirks of grammar affect the way you think? Grammatical gender and object concepts. In R. Alterman, D. Kirsh (Eds.), *Proceedings of the 25th Annual Meeting of the Cognitive Science Society* (pp. 928–933). Boston, MA: Cognitive Science Society.

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language, 39*, 633–651.

Prat-Sala, M., & Branigan, H. P. (2000). Discourse constraints on syntactic processing in language production: A cross-linguistic study in English and Spanish. *Journal of Memory and Language, 42*, 168-182.

R Core Team (2019). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/.

Ramos, S., & Roberson, D. (2010). What constrains grammatical gender effects on semantic judgements? Evidence from Portuguese. *Journal of Cognitive Psychology, 23*, 102–117.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition, 42,* 107–42.

Santesteban, M., Pickering, M. J., & Branigan, H. P. (2013). The effects of word order on subject–verb and object–verb agreement: Evidence from Basque. *Journal of Memory and Language, 68,* 160 –179.

Sato, S., & Athanasopoulos, P. (2018). Grammatical gender affects gender perception: Evidence for the structural-feedback hypothesis. *Cognition, 176,* 220-231.

Schmitt, B. M., Meyer, A. S., & Levelt, W. J. M. (1999). Lexical access in the production of pronouns. *Cognition, 69,* 313–335.

Schriefers, H. (1993). Syntactic processes in the production of noun phrases. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 841-850.

Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in production: Picture-word interference studies. *Journal of Memory & Language, 29*, 86-102.

Sera, M. D., Berge, C., & del Castillo Pintado, J. (1994). Grammatical and conceptual forces in the attribution of gender by English and Spanish speakers. *Cognitive Development, 9*, 261–292.

Sera, M., Elieff, C., Forbes, J., Burch, M. C., Rodríguez, W., & Dubois, D. P. (2002). When language affects cognition and when it does not: An analysis of grammatical gender and classification. *Journal of Experimental Psychology: General, 131*, 377–397.

Silverstein, M. (1976). Hierarchy of features an ergativity. In R. M. W. Dixon (Ed.), *Grammatical categories in Australian languages* (pp. 112-171). New Jersey: Humanities Press.

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). New York: Routledge.

Snedeker, J., & Trueswell, J. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language, 48,* 103–130.

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

Tabor, W., & Hutchins, S. (2004). Evidence for Self-Organized Sentence Processing: Digging-In Effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 431–450.

Vigliocco, G., & Franck, J. (1999). When sex and syntax go hand in hand: Gender agreement in language production. *Journal of Memory and Language, 40*, 455–478.

Vigliocco, G., & Franck, J. (2001). When sex affects syntax: Contextual influences in sentence production. *Journal of Memory and Language, 45*, 368 –390.

Vigliocco, G., Antonini, T., & Garrett, M. F. (1997). Grammatical gender is on the tip of Italian tongues. *Psychological Science, 8*, 314–317.

Vigliocco, G., Butterworth, B., & Garrett, M. F. (1996). Subject–verb agreement in Spanish and English: Differences in the role of conceptual constraints. *Cognition, 61*, 261–298.

Vigliocco, G., Butterworth, B., & Semenza, C. (1995). Constructing subject–verb agreement in speech: The role of semantic and morphological factors. J*ournal of Memory and Language, 34*, 186–215.

Vigliocco, G., Vinson, D. P., Paganelli, F., & Dworzynski, K. (2005). Grammatical gender effects on cognition: Implications for language learning and language use. *Journal of Experimental Psychology: General, 134*, 501–520.

Villata, S., & Franck, J. (2020). Similarity-based interference in agreement comprehension and production: Evidence from object agreement. *Journal of Experimental Psychology: Learning Memory and Cognition, 46*, 170-188.

Villata, S., Tabor, W., & Franck, J. (2018). Encoding and retrieval interference in sentence comprehension: Evidence from agreement. *Frontiers in Psychology, 9*, 2.

Wagenmakers, E. J. A practical solution to the pervasive problems of *p* values. *Psychonomic Bulletin & Review, 14*, 779–804.

Zhang, S., & Fukumura (2020). Similarity-based interference and the choice of referring expressions in Chinese and English. *The 26th Architectures and Mechanisms for Language Processing Conference.*

Open Practices Statement

The data and R scripts are available at https://osf.io/jwrt5/?view_only=eb6705fbf162486b86d4ee03bb11dd01