PSICOLÓGICA

What a transparent Romance language with a Germanic gender-determiner mapping tells us about gender retrieval: Insights from European Portuguese

Ana Rita Sá-Leite[1], Ângela Tomaz[2], Juan Andrés Hernández-Cabrera[3], Isabel Fraga[1], Carlos Acuña-Fariña[4], and Montserrat Comesaña[5, 6]

[1] Cognitive Processes & Behaviour Research Group, Department of Social Psychology, Basic Psychology, and Methodology, University of Santiago de Compostela, Spain
[2] Sciences Cognitives et Sciences Affectives (SCALab), University of Lille, France
[3] Department of Psychobiology and Methodology of Behavioral Sciences, University of La Laguna, Spain
[4] Cognitive Processes & Behaviour Research Group, Department of English and German Philology, University of Santiago de Compostela, Spain
[5] Psycholinguistics Research Line, CIPsi, School of Psychology, University of Minho, Portugal
[6] Centro de Investigación Nebrija en Cognición (CINC), Universidad Nebrija, Spain

Abstract

The study of the representation and processing of grammatical gender during language production has encountered mixed results regarding which conditions must be met to observe gender effects and whether these reflect the selection of gender values or competition between elements of agreement. The answer seems to depend on the number of determiners associated with each gender and on the language being explored. The present study aims to assess this issue through three picture-word interference tasks in European Portuguese. This is a transparent Romance language featuring a one-to-one gender-determiner mapping system similar to opaque Germanic languages. Conditions of gender in/congruency between targets and distractors were considered, along with gender transparency and agreement. We observed a gender congruency effect restricted to noun phrases. Importantly, the effect was modulated by transparency, which seems relevant regardless of agreement. To explain the results, we adapted the Dual-Route Model of language comprehension to production.

Keywords: grammatical gender, gender congruency effect, gender transparency, European Portuguese, dual-route model, late selection hypothesis, picture-word interference paradigm

Psicológica (2022)

DOI
10.20350/digitalCSIC/14777

Corresponding author
Ana Rita Sá-Leite
Campus Vida
Calle Xosé María Suárez Núñez
s/n, 15782
Santiago de Compostela
A Coruña, Spain
anarita.saleite.dias@usc.es

Edited by
Carlos Romero Rivas

Reviewed by
Alba Casado, Sendy Caffarra

Introduction

Languages with gender systems classify nouns according to a certain number of values. European Portuguese has a bipartite gender system in which nouns such as "*castelo*" (castle) belong to the masculine gender, and nouns such as "*mesa*" (table) belong to the feminine. Other languages have their own gender systems; thus, in German, "*Schloss*" (castle) is neuter, whereas "*Tisch*" (table) is masculine. Gendered languages thus arbitrarily assign nouns to a certain gender value. These gender values inherently classify nouns and determine the form of many other words contained in a sentence. The process by which other words change their form depending on the gender value of the main noun is called agreement. For instance, to comply with the rules of the Portuguese grammar, a native speaker would say "*O castelo amarelo*" (the yellow castle) but "*A mesa amarela*" (the yellow table). We can therefore define grammatical gender as an abstract lexico-syntactic feature of nouns whose value has no semantic consequences, and which has form-type repercussions only for the other words contained in an utterance (but for evidence of certain repercussions of abstract gender values at the level of conceptual encoding, see Casado et al., 2018, 2021).[1] Note that the noun form is not impacted by grammatical gender, but can partially correlate with it (Bates et al., 1995). This is known as ortho-phonological gender transparency, and languages show different degrees of gender transparency depending on the number of nouns showing this correlation. For instance, in European Portuguese, following calculations conducted using the P-PAL extensive word database (Soares et al., 2018), 61% of nouns are transparent, as most masculine nouns end in "-o" ("*castelo*") and most feminine nouns end in "-a" ("*casa*"). Approximately 22% of the nouns are *opaque* since they contain not-so-frequent endings that do not clearly correlate with any gender value (e.g., ending "-e" of "*amizade*" [friendship] which is feminine, or "*monte*" [hill] which is masculine). Finally, a few nouns (11%) end in the typical cues for gender "-o" and "-a" but are of the opposite gender; hence, they are called *irregular* nouns (feminine noun "*tribo*" [tribe]). The remaining (6%) have double gender.

Grammatical gender is thus a complex feature with important lexico-syntactic and ortho-phonological implications for language. Indeed, when exploring the cognitive representation and retrieval of gender values during language production researchers have faced interesting challenges, such as understanding the extent to which an agreement context is mandatory for the retrieval of a noun's gender value to occur. Besides, it is still unclear how exactly gender is selected. Some argue that it is a process based in facilitative priming or competitive mechanisms. In contrast, others are skeptical about the possibility that gender effects are obtained experimentally and usually defend that it is automatically retrieved as a direct consequence of lexical access. Likewise, the role of gender transparency is unknown; hence, the question remains whether gender cues have an active role during language production. Current research seems to suggest that there is no unique answer to these questions, as the nature of grammatical gender representation and processing tends to be driven by language-specific idiosyncrasies. In this study, we conducted three different experiments in European Portuguese to address some of the most controversial results in the field by exploring the role of the agreement context in gender retrieval and testing the replicability of certain effects.

Modeling the representation and processing of grammatical gender

Two influential models of language production have included gender in their tenets, namely WEAVER++ (Levelt et al., 1999) and the Independent Network model (IN; Caramazza, 1997; Caramazza & Miozzo,

[1] We should not mistake grammatical gender with natural gender, which is not the focus of this study as it has a semantic basis grounded on (usually) a sex distinction. Natural gender has an optional gender value depending on the biological or perceived sex of the referent, and has other types of repercussions for the morphology of nouns and lexical access itself.

1997). Both share a common ground in which they postulate the existence of three independent levels of language representation defined by the kind of information they encode: semantic, syntactic, and morpho-phonological. Grammatical gender is located at the syntactic level, along with characteristics such as number and word class (see Figure 1). Yet, there are certain relevant differences between both models as regards gender. WEAVER++ considers that syntactic information is processed at the "lemma" level of representation, and once the lemma is processed, the morpho-phonological information is encoded at the lexeme level. Thus, in the production of "*mala*" (feminine, handbag), meaning-related information is first processed at the conceptual level, at which point the concept node for "*MALA*" is selected. Following this, activation spreads to the so-called abstract lemma level, where one lemma node exists for every concept node, but there are also nodes for each syntactic property (e.g., "masculine" and "feminine") connected via labeled links (e.g., "gender") to each lemma. Thus, the lemma node "*mala*" will send activation to its grammatical gender node (feminine, to which all feminine nouns are linked). For gender retrieval to occur, the degree of activation of this node must exceed an absolute threshold of selection. However, Levelt and colleagues argue that the activation of syntactic features is not the same as their selection: selection only occurs when and where necessary (Roelofs et al., 1998). Hence, in the case of grammatical gender, if agreement does not have to be fulfilled to determine the form of other words contained in speech, none of the gender nodes will reach the threshold for selection because encoding of gender is not necessary. In this way, activation will spread to the morpho-phonological encoding level, and word form will be retrieved (/*mala*/) without gender being selected.

The IN model makes similar claims. It also assumes the existence of three levels of language processing, but the spread of activation is not discrete and unidirectional between the syntactic and the morpho-phonological strata[2]. Rather, the authors propose that the conceptual level has direct connections with the morpho-phonological level, and a noun can be produced without the syntactic level intervening. In fact, only after the morpho-phonological information has been selected is activation sent to the syntactic stratum for the full selection of the grammatical features to occur. Regarding gender, this model considers that abstract grammatical features that are inherent to nouns are selected as a consequence of lexical access itself. This selection involves competition between gender nodes, so that the retrieval of one value induces the inhibition of the other.

[2] Note that the most recent formulations of the WEAVER++ model (Roelofs, 2004, 2008; Roelofs & Piai, 2015) allow for a cascaded flow of activation to occur from the lemma to the lexeme and assume back spreading of activation from sounds to lexical items. However, these mechanisms were conceived for self-monitoring and executive control purposes to explain mainly task-based effects.

In sum, WEAVER++ states that gender retrieval occurs when gender nodes reach an absolute activation threshold of selection but argues that it happens previously to morpho-phonological encoding and only when agreement has to be fulfilled. On the contrary, the IN model argues that gender selection occurs through a competitive process that takes places after morpho-phonological encoding, and that any syntactic process can be skipped if it is not necessary but makes no further comments on whether this necessity is based on agreement.

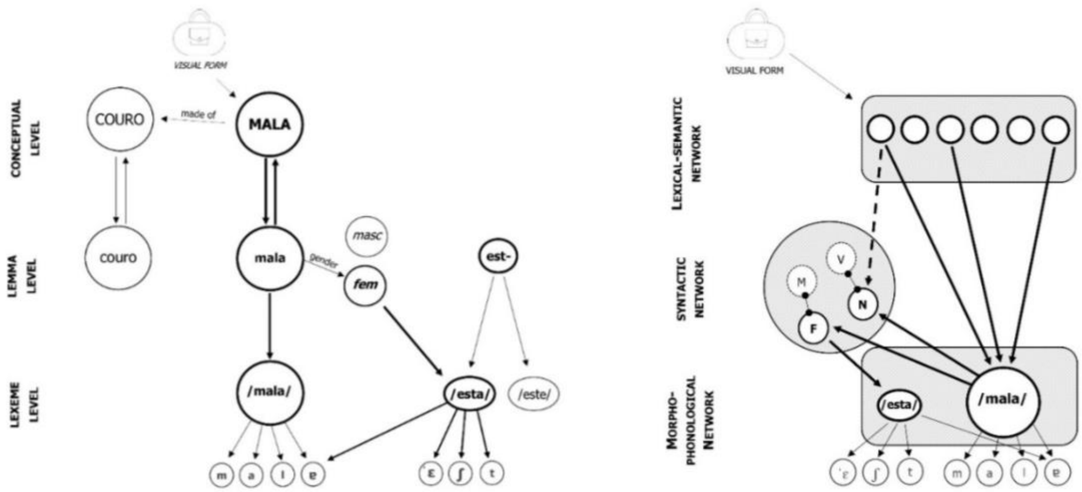

Figure 1. Simplified structure of lexical access according to WEAVER++ (left) and the IN model (right).

Note. WEAVER++ on the left: Processing of the Portuguese noun phrase "*esta mala*" (this purse). Bold nouns and thicker lines/circles represent selected nodes; non-bold and finer lines represent activation. Adapted from Levelt et al. (1999). IN model on the right: Only the Phonological-lexeme is shown, to simplify the figure. Dotted circles indicate inhibition. F = feminine; M = masculine; N = Noun; V = Verb. Adapted from Caramazza (1997).

Importantly, none of the two models refer to ortho-phonological gender transparency and to the impact of this variable on gender retrieval. Yet, there is a vast amount of evidence supporting an active role of transparency in gender acquisition and retrieval during language comprehension. Infants acquiring highly transparent languages rely mostly on the form-based regularities present in nouns (i.e., "-o" for the masculine and "-a" for the feminine in Spanish, Italian, and European Portuguese) along with the elements of agreement, such as articles, to acquire gender (Arias-Trejo & Alva, 2013; Pérez-Pereira, 1991). Importantly, the existence of transparent cues has been systematically shown to facilitate and increase the accuracy in the use of the elements of agreement, with transparent nouns showing a clear advantage over opaque nouns (Arias-Trejo et al., 2013; Caselli et al., 1993; Pérez-Pereira, 1991; Perona-Jara, 2015). Critically, differences in lexical processing as a function of transparency persist in life. In languages like Spanish and Italian, transparent nouns and pseudowords that show the transparent cues for gender benefit in terms of accuracy in lexical decision tasks as well as in tasks in which participants have to classify them according to gender or to choose between different elements of agreement. Similarly, empirical event-related potential (ERP) evidence consistently shows that gender-to-ending correspondences are detected at an early stage of processing and that the processing of transparent nouns requires more cognitive resources due to a larger amount of gender information being available (Caffarra & Barber, 2015; Caffarra et al., 2014, 2015). However, there is a lack of studies on the role of form regularities in gender retrieval during language production, and little has been theoretically advanced about what should be expected on this matter. Indeed, classical models of language production were created mainly from languages with low levels of transparency (such as Dutch and German, Velnić, 2020). This has probably contributed to their disregard of transparency when addressing gender retrieval. To our knowledge, the only proposal that acknowledges gender transparency in language production is the Double-Selection Model (DSM) by Cubelli et al. (2005). Yet, this is a highly controversial model that is centered exclusively on gender and is not based on other types of evidence as the WEAVER++ and the IN models are; hence, its experimental scope is limited. The DSM states that gender regularities cognitively depend on gender itself. More specifically, it states that in Italian and Spanish, in nouns such as "*casa*" (house, in Spanish), the typical cue for gender "-a" is retrieved once the feminine gender value is processed (i.e., Italian and Spanish speakers automatically process gender in order to process the form of a noun, the same would happen with the masculine and the "-o" ending). Since Germanic languages are opaque, gender is not retrieved because it is not needed to encode the form of the nouns. In these cases, it is thus only retrieved when agreement is necessary. Nonetheless, this model diverges from other classic views of grammar and does not locate gender at a separate syntactic level nor represents it as nodes. Conversely, the authors consider grammatical gender a lexical-semantic property of nouns associated with each noun as a feature. Competition for selection would take place between those lemmas that are related or similar in terms of their associated lexico-semantic features. Importantly, they do not propose any type of activation-based mechanism underlying the selection of gender. Although an interesting proposal, the DSM has been criticized for its conception of morpho-phonological decomposition and the organization that it proposes for grammatical gender (Finocchiaro et al., 2011; Sá-Leite et al., 2021).

Besides the DSM, no other model of language production has assessed so far how gender would be retrieved in interaction with noun form. In fact, the proposal that has accumulated the greatest amount of evidence comes from language comprehension, the Gollan and Frost's (2001) Dual-Route Model for gender retrieval. In a nutshell, according to the authors, there are two available sources for the selection of gender values when reading or hearing a noun: (1) a lexical source based on abstract information stored within the noun, and (2) a non-lexical, form-based source that relies on morphological and phonological information. During the processing of transparent nouns, our cognitive system can rapidly detect the phonological cues for gender via the form-based route. Then, checking processes take place to confirm that the selection of the gender value made through the form route matches the information obtained afterwards through the lexical route, as confirmed by Caffarra and colleagues (e.g., Caffarra et al., 2015; Caffarra & Barber, 2015). On the other hand, the selection of the gender value of opaque nouns has been shown to rely only on the lexical route, as the nominal endings are not identified as reliable cues for gender. Yet, it is unclear whether and how these routes would operate in language production.

Experimental assessment of grammatical gender in language production

The study of the retrieval of grammatical gender during the production of nouns has relied mainly on the picture-word interference (PWI) paradigm. In this task, participants are asked to name a picture using either a bare noun or a noun within a phrase (the target; e.g., in Dutch, "*huis*" [house] or "*het huis*" [the house]) while ignoring a superimposed noun (the distractor). Target and distractor nouns may have the same or a different gender value, creating conditions of gender (in)congruency that are expected to cause variations in the response times of the participants. More specifically, researchers predict a gender congruency effect (GCE), which consists of slower naming latencies when target and distractor have different gender (target "*het huis*" [neuter] with distractor "*boom*" [common, "tree"]) compared to when they coincide (same target with distractor "*oog*" [neuter], "eye"). Yet, results vary widely depending on the language family as well as on the type of phrase used to name the picture.

In Germanic and Slavic languages, the GCE has been obtained only when participants are asked to name the pictures using a noun phrase composed by an element of agreement that varies across gender values (e.g., "*het huis*" vs. "*de boom*", the house [neuter] vs. the tree [common]; Bordag & Pechmann, 2008; La Heij et al., 1998; Schiller & Caramazza, 2003; Schriefers, 1993; Schriefers & Teruel, 2000; van Berkum, 1997). The GCE was absent with (1) noun phrases in which the elements of agreement do not vary across gender values (e.g., the plural in German and Dutch, for which the same definite article is used regardless of the gender value of the main noun; for instance, in Dutch, "*de huizen*" vs. "*de bomen*" [the houses vs. the trees]; Schiller & Caramazza, 2003, 2006) or (2) bare nouns (Finocchiaro et al., 2011; La Heij et al., 1998; Starreveld & La Heij, 2004). Following the WEAVER++ model, it could be hypothesized that because agreement does not have to be fulfilled in the case of the plural (e.g., "*de*" is used for all gender values), then gender is not being selected, and thus, the GCE is absent. However, a series of consistent results with the singular-plural paradigm (a simple picture-naming task) showed that even during the production of plural noun phrases, the singular is active, and competition occurs between the singular determiners (e.g., "*de*" [common] and "*het*" [neuter]) and the plural determiner (e.g., "*de*") depending on the gender of the noun (facilitation for the common "*de*"; for a review, see Jescheniak et al., 2014). In view of these results, it was concluded that (1) the GCE is actually a determiner congruency effect occurring at the level of morpho-phonological encoding of determiners, and (2) genuine gender effects cannot be observed; hence gender is a feature whose retrieval is automatic and does not depend on contexts of agreement (Caramazza et al., 2001).

The picture is slightly more complicated in Romance languages. The GCE has been observed with bare nouns in European Portuguese (Sá-Leite et al., 2021, Experiment 1). In this study, the authors manipulated the animacy of the targets, on the assumption that animate nouns may capture the attention of the participants to a greater degree than inanimate nouns, as well as prioritize semantic processing (Branigan et al., 2008; New et al., 2007). This would decrease the capability of distractor nouns to interfere, as well as affect syntactic processing, perhaps inducing a tendency to skip it if not necessary (in line with the tenets of the IN model). Note that except for a couple of cases, all nouns included in Sá-Leite et al. (2021) were animate nouns with grammatical gender, not natural gender, as their gender value does not refer to their sex ("*morcego*" [masculine, bat], "*avestruz*" [feminine, ostrich]). Hence, there is not any kind of overlapping between semantic and gender information there. Indeed, the GCE was found only for inanimate targets. Besides informing on the role of animacy on the PWI paradigm and gender processing, these results support the idea that the selection of gender actually entails effects at the level of the lemma in the absence of agreement. More specifically, the authors explain their results following the structure proposed by WEAVER++ by which gender nodes accumulate activation and need to reach a threshold for selection to occur. Yet, they propose that the activation of more than one gender node will cause competition for selection and that agreement is not compulsory for gender selection to occur. This way, target and distractor lemmas of the same gender would see their activation converging on the same gender node, speeding and facilitating gender selection, but those of the opposite gender would create a situation of interference for selection. However, in other Romance languages like Spanish and Italian, the results have also taken the form of a gender incongruency effect, i.e., lower reaction times for target and distractor pairs of different (incongruent) gender. The DSM of Cubelli and colleagues (2005) explains this effect of incongruency as a lemma competition between nouns of the same gender, which occurs because Italian and Spanish are transparent languages; hence, gender has to be retrieved for the form of nouns to be encoded ("-*o*" and "-*a*"). Still, (a) it leaves unexplained why the effect is of congruency in European Portuguese, b) why the effect is also present for opaque nouns, whose form does not depend on gender, and (c) it is at odds with a recent meta-analysis which offers little support for this effect (Sá-Leite et al., in press). Indeed, null effects that contradict both the GCE and the incongruency effect have been found in Italian and Spanish (Finocchiaro et al., 2011; Sá-Leite et al., in press).

A completely different discussion emerges when speakers of Romance languages are asked to use noun phrases to name the pictures. In this case, effects are systematically null and hence do not support any type of competition occurring between determiners (Catalan: Costa et al., 1999; French: Alario & Caramazza, 2002; Italian: Cubelli et al., 2005; Miozzo & Caramazza, 1999; Miozzo et al., 2002; Spanish: Costa et al., 1999; Miozzo et al., 2002; O'Rourke, 2007). To explain these contrasting findings between Germanic/Slavic languages (in which a GCE that reflects competition between determiners is consistently obtained) and Romance languages, Miozzo and Caramazza (1999) proposed the Late Selection Hypothesis (LSH). Take, for example, the Italian definite articles. Unlike Dutch, there are various definite articles per gender value (for Dutch, "*de*" [common] and "*het*" [neuter]; for Italian "*lo*", "*il*", and "*l'*" [masculine], and "*la*" and "*l'* " [feminine]). Article selection depends on the phonological form of the following word. Thus, for instance, whereas "*lo*" is used when the next word starts with a consonant cluster of the form "*s* + consonant" or "*gn*-", or an affricate, "*il*" is used in all other cases. In Spanish, the feminine value shows similar variability: when the next word starts with a stressed "*a*-" ("*agua*" [feminine, water]), the definite article "*el*" instead of the typical "*la*" is used. Besides, in these languages, adjectives can be used prenominally, and hence the first phonological segment of the noun may not be the only candidate for determining the form of the definite article (e.g., "*lo zaino*" [the backpack] but "*il grande zaino*" [the big backpack]). Consequently, a mapping consisting of one-to-multiple gender-determiner forms in languages like Spanish or Italian demands a series of processes to occur before the correct determiner within a specific gender value is ready for selection. The processes masking the GCE are presumably the computation of the stressing patterns in Spanish and word identification and organization in both Spanish and Italian. It is unclear, however, if besides these processes of stressing and word identification/organization, the phonological encoding of the next word following the determiner (i.e., a noun or an adjective) is another additional process to take place before determiner selection. This is because it is unknown when it occurs, as it depends on the view we adopt: the WEAVER++ model proposes that gender is encoded before morpho-phonology, but the IN model supports the opposite idea (for more information, see Bürki et al., 2016). Thus, the absence of the GCE for these Romance languages has been explained as a matter of the late selection of the determiner form. More precisely, any competitive effect between the determiner forms of opposite genders is rendered invisible due to the delay that occurs from the moment the gender of target and distractor is processed to the moment the right article is selected. At that late stage, the only competition that hypothetically could be observed would be between determiner forms of the same gender (e.g., "*lo*" vs "*li*").

European Portuguese is an interesting language when putting the LSH to the test. It is a Romance language like Italian and Spanish but has a one-to-one gender-determiner form mapping for definite articles, as the tested Germanic and Slavic languages do ("*o*" for the masculine; "*a*" for the feminine). Following the LSH, if the null results with noun phrases are a matter of mapping and not of any other differences between both groups of languages, a GCE should be obtained when testing native speakers of European Portuguese producing noun phrases (definite article + noun). Yet, note that even if we obtain a GCE with noun phrases in European Portuguese, the current state of results does not allow us to affirm that this is uniquely a reflection of competition between determiners. This is because Sá-Leite et al. (2021) obtained a GCE with bare nouns, as was mentioned above.

The aim of the present study is to put the LSH to the test as well as to check the replicability of the GCE observed by Sá-Leite et al. (2021). We intend to do so while considering the ortho-phonological transparency category of both targets and distractors as a way of thoroughly exploring the processing of transparency in language production. Two PWI paradigm studies were carried out in which gender and nominal endings were manipulated for the first time in an orthogonal way using noun phrases (Experiment 1) and bare nouns (Experiments 2 and 3). Our hypotheses were the following:

(1) If the LSH is right, a classic GCE will emerge in Experiment 1 with noun phrases since European Portuguese, like German and Dutch, has a one-to-one gender-determiner form mapping.
(2) A GCE will also emerge when bare nouns are used, in line with the mechanisms of gender selection proposed by WEAVER++ and prior results with native speakers of European Portuguese (Sá-Leite et al., 2021).

The present study

Experiment 1

Method

Participants. Seventy-six native speakers of European Portuguese studying at the University of Minho (69 female; M*age* = 21.53 years, SD = 4.45) participated in this experiment. All were required to sign an informed consent for experimentation with human subjects approved by the ethical committee of the University of Minho (SECSH 017/2015). Extra course credits were given to participants.

Materials. Sixty-four pictures were selected as targets from the International Picture Naming Project (IPNP) database (Szekely et al., 2004). Targets were distributed evenly in four categories according to their grammatical gender value and their transparency: (a) 16 feminine and transparent words (e.g., "*palmeira*" [palmtree]), (b) 16 masculine and transparent words, (e.g., "*barco*" [boat]), (c) 16 feminine and opaque words (e.g., "*ponte*" [bridge]), and (d) 16 masculine and opaque words (e.g., "*tapete*" [rug]).

Pictures from the four categories were matched through a one-way ANOVA with four conditions (masculine-transparent, feminine-transparent, masculine-opaque, feminine-opaque) for visual complexity (objectively defined by the digital size of the drawing), $Z = 1.25$, $p = .299$, and goodness-of-depiction (i.e., how well each picture illustrated the target nouns), $Z = .29$, $p = .830$, as obtained from the dataset in Szekely et al. (2005). Target words were matched through the same one-way ANOVA across categories on per-million frequency (Log10), obtained from the SubtLex-PT database (Soares et al., 2015). They were also matched on the number of phonological and orthographic neighbors (N), word length (in letters), and mean logarithmic bigram frequency, obtained from the P-PAL database (Soares et al., 2018). Finally, they were also controlled for imageability, concreteness, and subjective frequency obtained from the Minho Word Pool database (MWP; Soares et al., 2017). See the mean values of every controlled variable per target type in Table 1 (all *ps* > .110).

Table 2. Means and standard errors (SE) as well as Z and p values of the controlled variables per Target type for the 64 experimental pictures in Experiments 1 and 2.

	Feminine Transparent	**Masculine Transparent**	**Feminine Opaque**	**Masculine Opaque**	***Z* (*p*)**
FpM	3.10 (.15)	3.40 (.12)	3.08 (.16)	2.89 (.13)	2.10 (.11)
PN	2.50 (.47)	5.19 (1.72)	3.19 (1.03)	2.31 (.95)	1.34 (.27)
ON	2.75 (.59)	5.63 (1.49)	3.25 (1.00)	2.44 (1.08)	7.83 (.51)
L	6.69 (.31)	6.31 (.53)	5.75 (.40)	6.06 (.38)	.92 (.44)
MLBF	2.55 (.11)	2.81 (.29)	2.45 (.14)	2.42 (.14)	.95 (.42)
VC	17265.13 (1891.93)	14295.12 (1407.58)	18067.06 (2590.53)	13230.19 (2210.94)	1.25 (.30)
GoD	5.88 (.11)	5.96 (.09)	5.81 (.23)	5.76 (.68)	.29 (.83)
I	5.91 (.08)	6.15 (.07)	5.97 (.08)	5.96 (.09)	1.78 (.16)
C	6.52 (.06)	6.55 (.05)	6.27 (.15)	6.44 (.08)	1.88 (.144)
SF	5.10 (.31)	5.13 (.32)	4.63 (.34)	4.39 (.32)	1.27 (.29)

Note. FpM = Frequency per million (Log_{10}). Obtained from the SubtLex-PT database (Soares et al., 2015). PN = Phonological number of neighbors (N); ON = Orthographic number of neighbors (N); L = Length (in letters); MLBF = Mean log bigram frequency. All obtained from the P-PAL database (Soares et al., 2018). VC = Visual complexity; GoD = Goodness-of-depiction. All obtained from the IPNP database (Szekely et al., 2004). I = Imageability; C = Concreteness; SF = Subjective Frequency (obtained from the Minho Word Pool database, Soares et al., 2017).

For each picture, four European Portuguese distractor nouns were selected (256 in total, 64 per condition); two of these had the same grammatical gender, and the other two had different grammatical gender. All target and distractor nouns were inanimate and thus had grammatical gender. Note that none of these distractors showed any semantic relationship with the targets with which they were paired, as they were from different semantic categories, nor was there any significant orthographic overlap between them, according to the NIM database (Guasch et al., 2013). In addition, all targets and their respective distractors differed in their initial phoneme to avoid any kind of facilitation during the naming task (Bi et al., 2009).

Targets were assigned to four experimental conditions as a function of their relation to Gender Congruency (GC) and Transparency Congruency (TC) with the distractors. For instance, with the feminine transparent target "*palmeira*" (palmtree), four distractors were selected: a) congruent in gender and in transparency (GCTC), such as "*carroça*" (carriage); b) incongruent in gender but congruent in transparency (GITC), such as "*ginásio*" (gymnasium); c) congruent in gender but incongruent in transparency (GCTI), such as "*grafite*" (graphite); and d) incongruent in both gender and transparency (GITI), such as "*foguete*" (rocket). The same conditions were created for feminine opaque targets as for transparent and opaque masculine ones.

Distractors were matched through a one-way ANOVA across 16 conditions per lexical frequency (log10), which was obtained from the SubtLex-PT database (Soares et al., 2015), on the number of phonological and orthographic neighbors (N), word length (in letters), and mean logarithmic bigram frequency, obtained from the P-PAL database (Soares et al., 2018; see mean values per target type in Table 2 and Table 3, all *ps* > .062 but see *Z* and *p* values in Table 4).

Table 2. *Means and standard errors (SE) of the controlled variables for distractors per experimental condition of feminine targets in Experiments 1 and 2.*

	FT GCTC	**FT GCTI**	**FT GITC**	**FT GITI**	**FO GCTC**	**FO GCTI**	**FO GITC**	**FO GITI**
FpM	2.63 (.19)	2.33 (.25)	2.56 (.17)	2.55 (.19)	2.68 (.30)	2.98 (.25)	2.75 (.21)	2.56 (.20)
PN	2.5 (.54)	2.56 (.57)	2.19 (.63)	2.50 (.58)	2.13 (.86)	3.44 (.93)	2.38 (.89)	2.56 (.87)
ON	2.94 (.57)	2.25 (.57)	2.44 (.71)	2.19 (.47)	2.13 (.68)	3.06 (.84)	2.63 (.92)	2.25 (.61)
L	6.38 (.30)	6.44 (.35)	6.50 (.27)	6.31 (.31)	6.13 (.43)	6.00 (.29)	6.06 (.44)	6.38 (.30)
MLBF	2.54 (.07)	2.55 (.09)	2.58 (.10)	2.54 (.10)	2.52 (.13)	2.60 (.09)	2.47 (.08)	2.49 (.12)
OVTD	.16 (.01)	.14 (.03)	.10 (.02)	.14 (.03)	.14 (.03)	.12 (.02)	.16 (.03)	.12 (.03)

Note. FT and FO stand for Feminine Transparent and Feminine Opaque. GCTC, GCTI, GITC and GITI stand for Gender Congruent and Transparency Congruent, Gender Congruent and Transparency Incongruent, Gender Incongruent and Transparency Congruent, and Gender Incongruent and Transparency Incongruent, respectively. FpM = Frequency per million (Log_{10}). Obtained from the SubtLex-PT database (Soares et al., 2015). PN = Phonological number of neighbors (N); ON = Orthographic number of neighbors (N); L = Length (in letters); MLBF = Mean log bigram frequency. All obtained from P-Pal database. OVPN = Orthographic overlap between targets and distractors was obtained from the NIM database (Guasch et al., 2013).

Table 3. Means and standard errors (SE) of the controlled variables for distractors per experimental condition of masculine targets in Experiments 1 and 2.

	MT GCTC	MT GCTI	MT GITC	MT GITI	MO GCTC	MO GCTI	MO GITC	MO GITI
FpM	2.79 (.21)	2.87 (.19)	2.80 (.14)	2.66 (.13)	2.30 (.20)	2.62 (.19)	2.21 (.17)	2.60 (.15)
PN	3.75 (.99)	3.63 (1.08)	4.50 (1.14)	3.31 (1.16)	1.69 (.59)	1.81 (.44)	2.13 (.73)	2.13 (.60)
ON	4.38 (1.29)	3.00 (1.02)	4.31 (1.05)	2.81 (.98)	1.38 (.55)	1.50 (.67)	1.63 (.49)	2.56 (.86)
L	6.13 (.36)	6.00 (.35)	5.94 (.39)	6.63 (.44)	6.25 (.36)	6.44 (.24)	6.81 (.34)	6.13 (.30)
MLBF	2.57 (.09)	2.55 (.09)	2.60 (.12)	2.56 (.10)	2.13 (.06)	2.28 (.12)	2.52 (.12)	2.35 (.09)
OVTD	.16 (.01)	.08 (.02)	.06 (.03)	.10 (0.2)	.15 (.03)	.16 (.03)	.19 (.03)	.14 (.04)

Note. MT and MO stand for Masculine Transparent and Masculine Opaque. GCTC, GCTI, GITC and GITI stand for Gender Congruent and Transparency Congruent, Gender Congruent and Transparency Incongruent, Gender Incongruent and Transparency Congruent, and Gender Incongruent and Transparency Incongruent, respectively. FpM = Frequency per million (Log_{10}). Obtained from the SubtLex-PT database (Soares et al., 2015). PN = Phonological number of neighbors (N); ON = Orthographic number of neighbors (N); L = Length (in letters); MLBF = Mean log bigram frequency. All obtained from P-Pal database. OVPN = Orthographic overlap between targets and distractors was obtained from the NIM database (Guasch et al., 2013).

Table 4. Z and p values of the controlled variables per Target type for distractors per experimental condition in Experiments 1 and 2

	Z (*p*)
FpM	1.05 (.40)
PN	.93 (.53)
ON	1.14 (.32)
L	.50 (.94)
MLBF	1.65 (.06)
OVTD	1.43 (.14)

Note. FpM = Frequency per million (Log_{10}). Obtained from the SubtLex-PT database (Soares et al., 2015). PN = Phonological number of neighbors (N); ON = Orthographic number of neighbors (N); L = Length (in letters); MLBF = Mean log bigram frequency. All obtained from P-Pal database. OVPN = Orthographic overlap between targets and distractors was obtained from the NIM database (Guasch et al., 2013).

Finally, eight different pictures were selected as training items from the IPNP database, along with eight new distractors. Four different lists were created for counterbalancing purposes. Each of these featured the same four conditions based on gender and transparency congruency and included the 64 target pictures (16 per condition). Half of the experimental items were masculine (32) and half were feminine (32). Likewise, half of the masculine and feminine pictures had a transparent ending (16), and half had an opaque ending (16). The same structure on gender value and transparency category was also followed for the distractors. Thus, each target picture was presented four times, one time per list and each time associated with a different distractor. Participants were randomly assigned to one of the four lists, assuring the same number of participants per list.

Procedure. Participants were tested individually in a soundproof booth. As in previous studies (e.g., Cubelli et al., 2005), each participant was first familiarized with the set of pictures to guarantee high naming agreement scores in the experimental task. To this end, all pictures were presented in single slides, along with the nouns that participants would have to use to name them in the task. Each picture and the corresponding noun were presented for 2 s and no overt response was required. The pictures were presented using Microsoft PowerPoint 2010. Words were shown in lowercase letters (Agency FB 36) below the respective picture. Stimuli presentation was randomized across participants.

In the PWI paradigm, participants were instructed to name each picture using a noun phrase (an article followed by a noun; e.g., "*a palmeira*" [the palmtree]) as quickly and accurately as possible, while ignoring the distractor noun superimposed onto the picture. All the pictures were presented at the center of the computer screen. The distractors, presented at the center of the pictures with a 0 ms stimulus onset asynchrony (SOA), were shown in lowercase letters in Courier New font, 14 points (see Cubelli et al., 2005, for a similar procedure). Each experimental trial had the following structure: a fixation point (+) at the center of the computer screen, for 500 ms; the target picture with the superimposed distractor for 2,000 ms or until response; a blank space for 500 ms as an inter-trial interval. Trials were presented randomly per participant with a pause after the first 32 trials. Response latencies were measured from the onset of the stimulus to the beginning of the naming response. Stimuli presentation was done using the DMDX software (Forster & Forster, 2003). Naming reaction times (RTs) were recorded by the voicekey from the presentation of the target to the onset of the naming response and then checked offline using the CheckVocal software (Protopapas, 2007). The session lasted approximately 20 min.

Results and interim discussion

Data corresponding to incorrect responses (naming errors [3.31%] and non-responses [2.12%]) were not included in the analysis. The pictures "*câmara*" (camera) and "*abóbora*" (pumpkin) were excluded for being incorrectly recognized by at least one-third of the participants (33.33%). Importantly, the masculine distractor "*vinho*" (wine), belonging to the second list and paired with feminine target "*ponte*" (bridge), was incorrectly classified as gender congruent, and thus, in the analysis, the condition of gender congruency lost one item per participant and the condition of gender incongruency gained one on the second list.

In addition, RTs that were more than 2.5 *SD*s above or below the mean for each participant in each condition were replaced by the mean response latency for the respective participant (3.03% in total). Mean response latencies (in milliseconds) per condition are presented in Table 5. As in previous studies

(Bordag & Pechmann, 2008; Costa et al., 1999; Cubelli et al., 2005; Paolieri et al., 2010, 2011), accuracy was not analyzed due to the absence of gender agreement errors (none of the participants used a masculine article for a feminine target or vice versa). See Table 6 for the model outcome, including estimates, *t*, *F* and *p* values.

Table 5. Raw and estimated mean response latencies (in milliseconds, standard deviations in parentheses) for Experiment 1 (noun phrases).

		Target type							
		Feminine transparent		**Masculine transparent**		**Feminine opaque**		**Masculine opaque**	
		Gender congruent	**Gender incongruent**	**Gender congruent**	**Gender incongruent**	**Gender congruent**	**Gender incongruent**	**Gender congruent**	**Gender incongruent**
Transparency congruency	Raw	867 (189.55)	861 (188.04)	840 (187.42)	836 (179.97)	843 (203.45)	847 (199.25)	848 (194.89)	838 (173.36)
	Est.	869 (22.2)	862 (20.8)	841 (21.6)	836 (20.2)	843 (21.6)	852 (20.2)	854 (21.4)	848 (20.0)
difference (estimated)		-7		-5		9		-6	
Transparency incongruency	Raw	860 (175.42)	874 (188.14)	829 (165.07)	818 (167.66)	822 (170.44)	873 (201.28)	835 (189.19)	857 (200.12)
	Est.	861 (22.1)	873 (20.6)	832 (21.4)	821 (20.0)	826 (21.6)	876 (19.9)	850 (21.3)	867 (19.7)
difference (estimated)		12		-10		50		17	

Note. Response latencies and standard deviations for each target type considering its congruency or incongruency with the distractor both in terms of gender and transparency congruency.

We analyzed RTs for target naming with linear mixed-effect (LME) models. The LMEs on RTs were conducted with participants and items as crossed random factors. The models were fit using the lme4 R package (Bates et al., 2015) as well as the LMerTest (Kuznetsova et al., 2017) and the lsmeans package (Lenth, 2016). This allowed us to contrast simple effects with differences in predicted least squares means from the mixed model estimated. There was no averaging of the data prior to the analyses. We investigated a 2x2x2x2 interaction between the design factors Target gender (feminine, masculine)[3], Target transparency (transparent, opaque), Gender congruency (GC: congruent [GC], incongruent [GI]) and Transparency congruency (TC: congruent [TC], incongruent [TI]). Effect coding (1 -1) was used for the four research factors. Because of the strong non-normality of residuals, a box-cox transformation of

[3] The gender of the target was included in the analysis models as differential effects depending on the gender value is a possibility we should not dismiss and was only considered once in the literature (see Paolieri et al., 2010). Yet, the differences in the acquisition of the masculine and feminine values (the feminine is usually acquired later, including in Portuguese, e.g., Barreña, 1997; Côrrea et al., 2011; Kuchenbrandt, 2008) as well the default vs. marked gender theories (Adamson, 2019; Corbett, 2000) provide enough reasons for research to consider differences on the retrieval of each gender value that may impact any related gender effect.

RT was performed with a power transformation of -0.22 (Box & Cox, 1964). Following Barr and colleagues (2013), a maximum random-effects model was carried out with random slopes for all factors by participant, plus Gender congruency and Transparency congruency (factors within by items) by item. This model shows perfect collinearity (correlation of 1 and -1 between random intercepts and slope for both random parts of the model), resulting in serious convergence problems and negative eigenvalues. The final estimated model that reached convergence had random intercepts for participants and items and random slopes only for Transparency Congruency by participants and Gender Congruency by items[4]:
model.p='RT.power ~ Target_Gender*GC*TC*Transparency + (TC|Participant) + (GC|Target)'

The coefficient of determination was $R^2 = 0.49$. Importantly, in order to assess the predictive power of item properties on the RTs, a stepwise linear regression was carried out with nine items' covariate variables as predictors (the number of phonological and orthographical neighbors, length [number of letters], and the mean log bigram frequency of the distractors, as obtained through the P-PAL database; the orthographic overlap between targets and distractors as obtained through the NIM database; the degree of imageability, concreteness, and subjective frequency of the distractors as obtained from the Minho Word Pool database, and the logarithmic frequency per million as obtained by the SubtLex-PT database). None of the nine candidate variables was significant.

The results on the RTs revealed a significant two-fold interaction between gender congruency and target transparency, $F(1, 54.9) = 5.51$; $p = .02$. Post-hoc comparisons revealed an effect of GC (GC < GI) restricted to opaque targets (GC opaque targets = 837 ms; GI opaque targets = 854 ms, $p = .002$). A visual inspection of the raw values of RT seemed to indicate that the greatest disparity between gender congruency conditions is found in opaque targets when paired with transparent distractors (51ms for feminine opaque targets and 22ms for masculine opaque targets). Although Gender Congruency x Transparency Congruency did not reach significance, it was marginally significant $F(1, 4241.5) = 3.32$; $p = .068$. Planned comparisons revealed that the GCE was due mainly to transparency-incongruent distractors (when the target and distractor were from different categories of transparency, GC = 836 ms, GI = 855 ms; for instance, an opaque target such as "*tapete*" [rug] paired with a transparent distractor such as "*joelho*" [knee], $p = .0095$; conversely, when target and distractor had the same transparency category, the effect was not significant, for instance, "tapete" and "xadrez" [chess] $p = .90$).

[4] To avoid unnecessary saturation of the model to be estimated, the factor "List" was finally not included since it was not significant ($p > 0.40$), and the previous results were not modified. Model for Experiment 1 including List: model.p='RT.power ~ as,factor(List) + Target_Gender*GC*TC*Transparency + (TC|Participant) + (GC|Target)

Table 6. Model outcome including estimates, t values, degrees of freedom (DenDF) for F, F values, and p values for each factor and interaction for Experiment 1.

	Estimate	*t* value	DenDF *F*	*F* value	*p*
Target Gender	-0.0018775220	-0.943	59.9	0.8891	0.34952
Gender Congruency	-0.0010733803	-1.855	54.6	34.405	0.06902.
Transparency Congruency	-0.0001602275	-0.297	72.8	0.0883	0.76723
Target Transparency	-0.0000001272	0.000	59.9	0.0000	0.99995
Target Gender: Gender Congruency	0.0010191611	1.761	54.6	31.007	0.08386.
Target Gender: Transparency Congruency	0.0004076214	0.762	4249.0	0.5803	0.44625
Gender Congruency: Transparency Congruency	0.0009788972	1.823	4241.2	33.242	0.06834.
Target Gender: Target Transparency	-0.0026522143	-1.332	59.9	17.740	0.18793
GenderCongruency: Target Transparency	0.0013583403	2.346	54.7	55.059	0.02260*
Transparency Congruency: Target Transparency	0.0006765105	1.265	4244.3	15.991	0.20610
Target Gender:Gender Congruency: Transparency Congruency	-0.0007705304	-1.436	4240.5	20.631	0.15097
Target Gender:Gender Congruency: Target Transparency	-0.0003296552	-0.570	54.6	0.3245	0.57127
Target Gender:Transparency Congruency: Target Transparency	0.0003681366	0.688	4246.8	0.4736	0.49138
Gender Congruency: Transparency Congruency:Target Transparency	-0.0008008069	-1.494	4237.9	22.320	0.13526
Target Gender: Gender Congruency:Transparency Congruency: Target Transparency	-0.0000915673	-0.171	4236.2	0.0292	0.86435

In short, as predicted by the LSH (Miozzo & Caramazza, 1999), a GCE was obtained for the first time in a PWI paradigm study with speakers of a Romance language producing noun phrases (determiner + noun), hence confirming that European Portuguese indeed behaves as an early-selection language and that is hence a matter of gender-form mapping. However, the GCE in European Portuguese is not as simple as the one reported for Germanic and Slavic languages. Our orthogonal design allowed us to obtain evidence supporting the notion that transparency matters in language production: the GCE interacted with the transparency category of the targets. More specifically, the GCE was due mainly to opaque targets. Furthermore, planned comparisons suggest that the effect tended to be more robust for targets paired with distractors of another transparency category (i.e., for opaque targets, the effect was greater when they were paired with transparent distractors). As displayed in the mean RTs of Table 5, the disparity on the RTs across gender congruency conditions for the transparency incongruent conditions, especially for opaque targets, is indeed the greatest. We will explain these results in more detail in the General Discussion. Importantly, taken solely, the results of this experiment are insufficient to decide if this GCE is only a reflection of determiner competition or also of gender competition. The next two experiments will address this issue by asking participants to use bare nouns to name the pictures.

Experiment 2

Method

Participants. Eighty native speakers of European Portuguese studying at the University of Minho (72 female; M*age* = 21.53 years, SD = 4.68) from the same population as those in Experiment 1 participated in this experiment. Participants were required to sign an informed consent for experimentation with human subjects approved by the ethical committee of the University of Minho (SECSH 017/2015). Extra course credits were given to participants.

Materials. We used the same materials as in Experiment 1.

Procedure. We followed the same procedure used in Experiment 1, except that participants were required to produce bare nouns (without articles).

Results and interim discussion

Data corresponding to incorrect responses (naming errors [2.50%] and non-responses [3.43%]) were not included in the analysis. The pictures "*câmara*" (camera) and "*abóbora*" (pumpkin) were again excluded for being incorrectly recognized by at least one-third of the participants (33.33%). Given that the same materials as those from Experiment 1 were used, the condition of GC had one less item per participant and the condition of GI had one extra item on the second list due to an incorrect categorization (distractor "*vinho*" with target "*ponte*"). Again, RTs exceeding 2.5 *SD*s above or below the mean for each participant in each condition were replaced by the mean response latency for the respective participant (2.88% in total). Mean response latencies (in milliseconds) per condition are presented in Table 7. As in Experiment 1, accuracy was not analyzed due to the absence of gender-agreement errors. See Table 8 for the model outcome, including estimates, t, F, and p values.

Table 7. Raw and estimated mean response latencies (in milliseconds, standard deviations in parentheses) for Experiment 2 (bare nouns).

		Target type							
		Feminine transparent		**Masculine transparent**		**Feminine opaque**		**Masculine opaque**	
		Gender congruent	**Gender incongruent**	**Gender congruent**	**Gender incongruent**	**Gender congruent**	**Gender incongruent**	**Gender congruent**	**Gender incongruent**
Transparency congruency	Raw	894 (187.36)	898 (189.59)	878 (181.06)	859 (171.52)	852 (183.89)	882 (211.22)	874 (196.91)	876 (207.77)
	Est.	900 (23.7)	893 (22.4)	882 (22.5)	861 (21.1)	856 (23.0)	883 (21.7)	883 (22.2)	889 (20.9)
difference (estimated)		7		-21		27		6	
Transparency incongruency	Raw	884 (188.01)	901 (194.78)	873 (182.98)	851 (169.84)	865 (199.01)	878 (195.88)	887 (192.31)	895 (195.64)
	Est.	877 (23.5)	896 (22.1)	877 (22.1)	853 (20.8)	871 (22.8)	881 (21.3)	897 (21.9)	902 (20.5)
difference (estimated)		19		-24		10		5	

Note. Response latencies and standard deviations for each target type considering its congruency or incongruency with the distractor both in terms of gender and transparency congruency.

The same analysis with the same factors included in Experiment 1 was replicated here. Effect coding (1 -1) was used for the four research factors. Because of the strong non-normality of residuals, a box-cox transformation of RT was performed with a power transformation of -1 (Box & Cox, 1964). Following Barr and colleagues (2013), a maximum random-effects model was carried out with random slopes for all factors by participants and Gender congruency and Transparency congruency (factors within by items) by items. Because of non-convergence problems and negative eigenvalues, the confidence intervals of random effects were checked graphically (Pinheiro & Bates, 2000). The final estimated model had random intercepts for participants and items, and random slopes for Target Gender and Transparency Congruency by participants and Gender Congruency by items:

model.p='RT.p ~ Target_Gender*GC*TC*Transparency + (Target_Gender + TC|Participant) + (GC|Target)'[5]

The coefficient of determination was $R^2 = 0.49$. As in Experiment 1, a stepwise regression with RTs as outcome and the same nine items' covariate variables as predictors was carried out. None of the candidate covariate variables was significant.

[5] To avoid unnecessary saturation of the model to be estimated, the factor "List" was finally not included since it was not significant ($p > 0.40$) and the previous results were not modified. Model for Experiment 2 including List: model.p='RT.p ~ as.factor(List) + Target_Gender*GC*TC*Transparency + (Target_Gender + TC|Participant) + (GC|Target)'

Table 8. Model outcome including estimates, t values, degrees of freedom (DenDF) for F, F values, and p values for each factor and interaction for Experiment 2.

	Estimate	***t* value**	**DenDF *F***	***F* value**	***p***
Target Gender	0.0000005455	0.046	60.7	0.0021	0.963519
Gender Congruency	-0.0000003989	-0.126	57.4	0.0159	0.899943
Transparency Congruency	-0.0000012645	-0.451	223.7	0.2036	0.652259
Target Transparency	0.0000022145	0.187	59.8	0.0350	0.852132
Target Gender: Gender Congruency	0.0000060010	1.900	57.3	36.101	0.062463.
Target Gender: Transparency Congruency	-0.0000020924	-0.773	4406.0	0.5979	0.439436
Gender Congruency: Transparency Congruency	0.0000009547	0.352	4417.4	0.1237	0.725071
Target Gender: Target Transparency	-0.0000141137	-1.193	59.8	14.235	0.237548
GenderCongruency: Target Transparency	0.0000055133	1.745	57.4	30.462	0.086278.
Transparency Congruency: Target Transparency	0.0000071368	2.637	4404.0	69.562	0.008382**
Target Gender:Gender Congruency: Transparency Congruency	-0.0000012470	-0.460	4417.7	0.2113	0.645768
Target Gender: Gender Congruency: Target Transparency	0.0000021299	0.674	57.3	0.4548	0.502783
Target Gender: Transparency Congruency: Target Transparency	0.0000010971	0.406	4404.4	0.1644	0.685118
Gender Congruency: Transparency Congruency: Target Transparency	0.0000019375	0.715	4396.4	0.5115	0.474544
Target Gender: Gender Congruency: Transparency Congruency: Target Transparency	-0.0000034387	-1.269	4399.8	16.104	0.204497

The results for the RTs revealed non-significant effects involving the factor gender congruency. A significant two-fold interaction between transparency congruency and target transparency was observed, $F(1, 4404) = 6.96$; $p = .008$. Planned comparisons revealed that when targets were opaque, they were named faster when paired with distractors of the same transparency category (also opaque) than when paired with transparent distractors (for opaque targets, TC < TI, 871 ms and 881 ms, respectively, $p = .03$).

The GCE obtained by Sá-Leite et al. (2021) with bare nouns was not replicated. This absence of gender effects with bare nouns fits the null results obtained by Finocchiaro et al. (2011) for Spanish and Italian. Importantly, they support the interpretation of the GCE of our first experiment as a reflection of competition between determiners. However, although there were no significant effects of GC, there was an effect of transparency that might suggest that although pure gender effects did not emerge in the RTs, gender nodes are somehow activated during bare noun production, creating different degrees of activation in different gender nodes and inducing disruption on lexical access. More specifically, opaque nouns here benefited from being paired with distractors of the same transparency category (also opaque). This leads us to believe that Experiment 2 reflects a similar pattern of results as those reported in Experiment 1 in terms of transparency. Therefore, and because opaque targets seemed to somehow have a special role in both experiments, we decided to conduct an additional experiment with bare nouns (Experiment 3) using a simpler design in order to increase statistical power. In this experiment, we selected only opaque targets (i.e., the factor "Target transparency" was dropped) with a partially new set of materials to keep experimental conditions matched on variables like frequency per million. This time, we had four lists, but each participant completed every list in four different sessions, which allowed us to obtain many more entries per stimuli. Importantly, to increase the similarity of our experiments to those that obtained a gender incongruency effect with Romance languages (Cubelli et al., 2005; Paolieri et al., 2010, 2011), we included a similar percentage of animate stimuli (28% of target nouns; 9% of distractor nouns) within the list, as this was the main differential factor in the control of the stimuli between our experiments and theirs.

Experiment 3

Method

Participants. Forty-eight native speakers of European Portuguese studying at the University of Minho (45 female; M_{age} = 20.48 years, *SD* = 2.70) from the same population as previous experiments participated in this experiment. Participants were required to sign an informed consent for experimentation with human subjects approved by the ethical committee of the University of Minho (SECSH 017/2015). Extra course credits were given to participants.

Materials. We selected 32 opaque pictures as targets, including 22 of those used in Experiments 1 and 2. The procedure was the same as that in the previous experiments. However, we also selected 32 items as transparent fillers. Every picture was presented four times, each time in a different list with a different distractor (see Tables 5 and 6 for the control on targets and distractors). With the aim of increasing the stimuli pool, this time 13% of pictures and distractors were animate. Furthermore, in this way, we were able to emulate previous studies in which approximately 8 to 33% of animate stimuli were included, and effects were systematically obtained with bare nouns (e.g., Cubelli et al., 2005; Paolieri et al., 2010, 2011). However, instead of having a different set of participants per list, each participant saw all the lists, as was done in previous studies (Cubelli et al., 2005; Finocchiaro et al., 2011; La Heij et al., 1998; O'Rourke, 2007; Paolieri et al., 2010; Paolieri et al., 2011). This allowed us to increase the statistical power. The order of presentation was counterbalanced across participants.

Masculine and feminine target pictures and nouns were matched through a one-way ANOVA for visual complexity and goodness-of-depiction as obtained from the dataset in Szekely et al. (2005), per-million frequency (Log10), number of phonological and orthographic neighbors (N), word length (in letters), and mean logarithmic bigram frequency, this obtained from the P-PAL database (Soares et al., 2018). All *ps* > .121 (see mean values in Table 9). Finally, imageability, concreteness, and subjective frequency for only the 81% of available target nouns in the Minho Word Pool database (MWP; Soares et al., 2017). All *p*s > .05 (see Table 9), yet, the absence of values for six targets across both groups (Feminine and Masculine targets) makes these data not reliable enough. Distractor nouns were matched for the same variables plus orthographic overlap (NIM; Guasch et al., 2013) except the picture-related ones (visual complexity and goodness-of-depiction). All *ps* > .50 (see mean values in Table 10). Analyses for imageability, concreteness, and subjective frequency could be only carried out for 58% of the distractor nouns, those available in the Minho Word Pool database (*p* = .063) and thus are not informative enough.

Procedure. We followed the procedure used in Experiment 2; hence, participants were required to produce bare nouns (without the use of articles). In addition, each list was presented on a different day (with a time interval ranging between two to three days) to avoid any biases related to residual memory effects; the presentation of the lists was counterbalanced.

Results and interim discussion

Data corresponding to incorrect responses (naming errors and non-responses) were not included in the analysis (3.34%). Also, RTs that were more than 2.5 *SDs* above or below the mean for each participant in each condition were removed (2.76% of the total). Mean response latencies are presented in Table 11. The model outcome is presented in Table 12.

We analyzed RTs (power transformed with a lambda of -0.5, Box & Cox, 1964) for target naming with linear mixed effect (LME) models. As in the two previous experiments, a maximum random-effects model was carried out with random slopes for all factors by participants and Gender congruency (factor within by items) by items. Effect coding (1 -1) was used for the three research factors. Because of collinearity and convergence problems, the final model had a random slope only for the GC factor by item[6]:

[6] To avoid unnecessary saturation of the model to be estimated, the factor "List" was finally not included since it was not significant (*p* > 0.40), and the previous results were not modified. Model for Experiment 3: model.p='RT.p ~ as.factor(List) + Target_Gender*GC*TC+ (TC|Participant) + (GC|Target)'

model.p='RT.p ~ Target_Gender*GC*TC + (TC|Participant) + (GC|Target)'

The coefficient of determination was $R^2 = 0.35$. The models were fit using the lme4 R package (Bates et al., 2015) as well as the LMerTest (Kuznetsova et al., 2016) and lsmeans packages (Lenth, 2016). This allowed us to contrast simple effects with differences in predicted least squares means from the estimated model. There was no averaging of the data prior to the analyses. The design was the same as in previous experiments except for the absence of the Transparency factor (i.e., it was a 2x2x2 design). The stepwise regression with RT as the outcome and the nine items' characteristics as predictors did not yield any significant predictive model. No significant effects were obtained.

Table 9. Means and standard errors (SE) as well as F and p values of the controlled variables per Target type for the 32 experimental stimuli in Experiment 3.

	Feminine	**Masculine**	***F* (*p*)**
FpM	1.05 (.18)	.76 (.11)	1.98 (.17)
PN	3.19 (1.03)	2.19 (.92)	.52 (.47)
ON	3.25 (1.01)	2.31 (1.09)	.40 (.53)
L	5.75 (.40)	5.88 (.40)	.05 (.83)
MLBF	2.45 (.14)	2.54 (.21)	.14 (.71)
VC	18067.06 (2590.53)	13137.20 (2216.83)	2.09 (.16)
GoD	5.81 (.23)	5.75 (.16)	.05 (.83)
I	5.97 (.08)	5.98 (.08)	.01 (.91)
C	6.27 (.53)	6.59 (.04)	4.13 (.05)
SF	4.63 (.34)	4.67 (.28)	.01 (.93)

Note. FpM = Frequency per million (Log_{10}); PN = Phonological number of neighbors (N); ON = Orthographic number of neighbors (N); L = Length (in letters); MLBF = Mean log bigram frequency. All obtained from the P-PAL database (Soares et al., 2018). VC = Visual complexity; GoD = Goodness-of-depiction. All obtained from the International Picture Naming Project (IPNP) database (Szekely et al. 2004). I = Imageability; C = Concreteness; SF = Subjective Frequency (obtained from the Minho Word Pool database, Soares et al., 2017).

Table 10. Means and standard errors (SE) as well as F and p values of the controlled variables for distractors per experimental condition of masculine targets in Experiment 3.

	F GCTC	**F GCTI**	**F GITC**	**F GITI**	**M GCTC**	**M GCTI**	**M GITC**	**M GITI**	***F* (*p*)**
FpM	1.03 (.20)	1.07 (.19)	1.05 (.19)	1.04 (.20)	1.01 (.15)	1.05 (.18)	1.04 (.16)	1.10 (.17)	.02 (1.00)
PN	2.19 (.85)	3.81 (1.11)	2.94 (.91)	2.63 (.86)	1.94 (.69)	2.50 (.97)	1.56 (.46)	1.94 (.59)	.73 (.65)
ON	2.19 (.67)	3.44 (1.08)	2.63 (.69)	2.19 (.57)	2.13 (.68)	2.88 (.92)	1.31 (.46)	2.44 (.85)	.66 (.70)
L	6.06 (.41)	5.94 (.29)	5.94 (.43)	6.25 (.28)	5.75 (.37)	6.06 (.37)	6.63 (.41)	6.13 (.34)	.50 (.83)
MLBF	2.49 (.13)	2.58 (.10)	2.56 (.12)	2.50 (.09)	2.23 (.16)	2.44 (.16)	2.50 (.08)	2.47 (.13)	.77 (.61)

OVTD	.14 (.03)	.09 (.02)	.13 (.04)	.06 (.03)	.09 (.03)	.11 (.02)	.14 (.03)	.09 (.02)	.92 (.50)

Note. FT and FO stand for Feminine Transparent and Feminine Opaque. GCTC, GCTI, GITC, and GITI stand for Gender Congruent and Transparency Congruent, Gender Congruent and Transparency Incongruent, Gender Incongruent and Transparency Congruent, and Gender Incongruent and Transparency Incongruent, respectively. FpM = Frequency per million (Log_{10}). Obtained from the SubtLex-PT database (Soares et al., 2015). PN = Phonological number of neighbors (N); ON = Orthographic number of neighbors (N); L = Length (in letters); MLBF = Mean log bigram frequency. All were obtained from P-Pal database. OVPN = Orthographic overlap between targets and distractors was obtained from the NIM database (Guasch et al., 2013).

Table 11. Raw and estimated mean response latencies (in milliseconds, standard deviations in parentheses) for Experiment 3 (bare nouns and opaque targets).

		Target type			
		Feminine opaque		**Masculine opaque**	
		Gender congruent	**Gender incongruent**	**Gender congruent**	**Gender incongruent**
Transparency congruency	Raw	807 (169.29)	822 (177.59)	797 (165.84)	792(143.95)
	Est.	810 (19.7)	824 (19.0)	799 (19.6)	796 (19.0)
difference (estimated)		14		-3	
Transparency incongruency	Raw	819 (167.20)	826 (172.41)	802 (159.91)	789 (157.04)
	Est.	822 (19.6)	829 (19.0)	807 (19.6)	792 (19.0)
difference (estimated)		7		-15	

Note. Response latencies and standard deviations for each target type considering its congruency or incongruency with the distractor both in terms of gender and transparency congruency.

Table 12. Model outcome including estimates, t values, degrees of freedom (DenDF) for F, F values, and p values for each factor and interaction for Experiment 3.

	Estimate	***t* value**	**DenDF *F***	***F* value**	***p***
Target Gender	-0.00044502	-0.910	30.0	0.8272	0.37033
Gender Congruency	-0.00002568	-0.267	30.2	0.0711	0.79151
Transparency Congruency	-0.00011610	-1.549	5656.3	24.000	0.12139
Target Gender: Gender Congruency	0.00016563	1.720	30.2	29.580	0.09567.
Target Gender: Transparency Congruency	0.00009400	1.254	5656.3	15.733	0.20978
Gender Congruency: Transparency Congruency	-0.00012949	-1.728	5656.1	29.855	0.08407.
Target Gender: Gender Congruency: Transparency Congruency	-0.00005759	-0.768	5656.1	0.5904	0.44230

Discussion

In the present study, we aimed to explore grammatical gender processing in European Portuguese through the PWI paradigm. In Experiment 1, we required participants to name the pictures using noun phrases formed by definite articles and nouns. In Experiments 2 and 3, we asked them to use bare nouns. Our design considered not only the gender congruency between the target and distractor, but also the specific gender of the target and the gender transparency of targets and distractors. In Experiment 1, we expected to confirm the tenets of the LSH (Miozzo & Caramazza, 1999) because European Portuguese has a one-to-one mapping between gender and definite articles; hence, we should obtain a classical GCE. In Experiment 2, since a genuine GCE had been obtained previously when testing bare nouns in European Portuguese (Sá-Leite et al., 2021), we expected to replicate this GCE with bare nouns.

The results showed a classical GCE in Experiment 1 in interaction with transparency, confirming our first hypothesis. Conversely, no effects of gender were obtained in Experiments 2 and 3. This ultimately supports the LSH, which understands the GCE as a result of competition between determiner forms that is only obtainable when determiners can be selected early during lexical access. Yet, the GCE was due to opaque target nouns, revealing that this effect is not as simple as the one obtained in opaque languages. To better explain the results, we shall propose an adaptation of the Dual-Route Model to language production by integrating the lexical and form-based routes in either WEAVER++ or the IN model (see Figures 2 and 3).

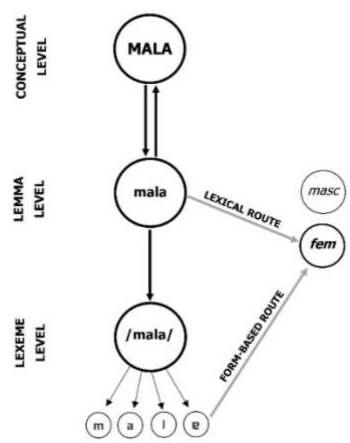

Figure 2. Representation of the routes of gender retrieval within WEAVER++.

Note. Model is simplified. Note that processing at the lemma level starts earlier than at the lexeme level, and hence, the lexical route is activated earlier than the form-based route. In fact, we are unsure about whether or not the activation of both routes overlaps in time, and within this model, it is possible that the form-based route has a confirmatory role for the selection already carried out by the lexical route.

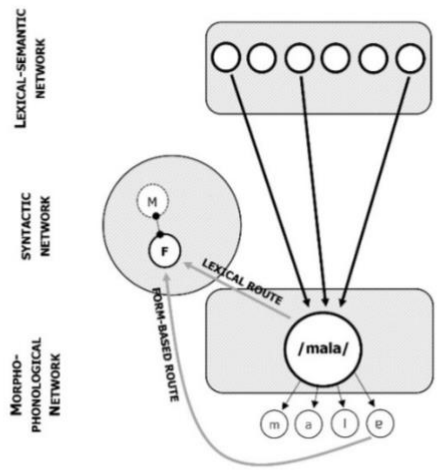

Figure 3. *Representation of the routes of gender retrieval within the IN model.*

Note. Model is simplified. Note that processing at the morpho-phonological network starts earlier than at the syntactic network; hence, the form-based route is expected to be activated earlier than the lexical route. In any case, we propose that the activation of both routes might overlap in time.

If we follow WEAVER++ and hence the idea that the processing of the syntactic features precedes that of the form, then the retrieval of gender takes place first through the lexical route. For transparency to have a role in the retrieval of gender in transparent nouns, we must either adopt an interactive view of our linguistic system, as proposed by Dell's account (1990), or use the self-monitoring internal speech system proposed by the authors of WEAVER++[7]. This way, after gender is activated through the lexical route, once the form of the noun is encoded, nominal endings could increase the activation of the node activated in a previous phase. Conversely, following the IN model, nouns' form would be processed in an earlier stage than gender, and hence the first source of gender retrieval would be the transparent cues present in the nominal endings of the nouns. The presence of certain specific cues of transparency ("-o" and "-a") would directly activate a specific gender value. The gender value activated by the form-based cue would also receive activation from the lexical route (i.e., the gender value associated with that noun in our memory). Note that either if we conceive noun form encoding starting before gender retrieval or gender retrieval starting before noun form encoding, transparent nouns always rely on a more accurate but costly way of gender retrieval, capturing most of the available cognitive resources. Indeed, as noted in the Introduction, participants make more classification and agreement errors with opaque than with transparent nouns in comprehension tasks, and empirical ERP evidence consistently shows that transparent nouns use a greater number of cognitive resources to retrieve gender (Caffarra & Barber, 2015; Caffarra et al., 2014). This is consistent with our results: transparent nouns, which are

[7] This system feeds a rightward incrementally generated phonological word back into the speech comprehension system (the lemma is shared between the speech comprehension and production systems).

scarce in opaque languages such as Dutch or German, are more protected from interference coming from distractors during gender retrieval. Hence, the GCE was observed for opaque nouns. In addition, a marginal interaction with the factor Transparency Congruency in this experiment revealed that distractors of a different transparency category than that of the target seemed to be interfering more than those of the same category. An inspection of the mean RTs reveals that this is especially true for opaque targets paired with transparent distractors (mean GCE of 37.5 ms). This makes sense: transparent distractor nouns consume more cognitive resources due to the noun-form having a role in gender retrieval and are not only less sensitive to interference but more prone to interfere. Particularly, transparent distractor nouns activate a certain gender node through the form-based route, in addition to the activation coming also from their own lexical route. However, opaque distractor nouns activate gender exclusively through the lexical route; hence, the amount of activation is lower and generates less interference in gender selection than a transparent distractor noun does. Ultimately, this transparency interference in the selection of the proper gender value is reflected in the selection of the correct determiner. Our results hence suggest that activation cascades from gender values to determiners. Indeed, activation cascading from features of definiteness, number, and gender during the selection of determiners has been consistently shown in the literature (this is conceived in the so-called *primed unitized activation hypothesis*; for more details, see Alario & Caramazza, 2002). Note that if gender selection were resolved before the determiner selection, we would not expect that the activation would come in different amounts and times from the different routes of gender retrieval to influence the competition between determiners of different gender. Interaction between both effects thus happens because both processes overlap in time due to a cascaded flow of activation. The differences in the amount of activation reaching one gender node or another are then reflected in the amount of activation distributed across determiner forms, hence creating different patterns of activation and competition depending on which and on how many routes are sending activation. Once gender selection finishes, different amounts of activation have already reached each determiner representation, influencing how the process of determiner selection takes place.

In Experiment 2, even though we did not find effects of gender congruency, opaque target nouns also benefited from being paired with other opaque distractors in comparison to transparent distractors. This might mean that the activation coming from different routes of gender is creating some kind of disruption in the processing of the target's gender. This suggests that gender is at least being activated. Yet, we could consider this effect of transparency as evidence for either of the next two ideas: (A) gender is activated but not selected, as WEAVER++ suggests, hence the absence of competitive effects of gender; (B) gender is being selected, but gender nodes do not bear experimentally observable effects, perhaps due to low levels of activation, or due to their being masked in the process of lexical access captured by the PWI paradigm (note that gender effects obtained in bilinguals, in which two languages and two lexical entries are systematically contributing to the activation of gender nodes, are small in size according to recent meta-analyses, Sá-Leite et al., 2020). It is also worth noting that, although this effect of transparency might be seen as a mere effect of form (i.e., of phonological facilitation of nominal endings), we do not think that this is the case. In this experiment, 36 of 64 of the distractors had the same nominal ending as their associated opaque target noun (mostly "-e"). Nevertheless, we need to bear in mind here that there is one condition with a perfect match of the coinciding final letters in targets and distractors: transparent targets paired with transparent distractors that are gender congruent (feminine targets ending in "-a" with distractors ending in "-a"; masculine targets ending in "-o" with distractors ending in "-o"). There were 32 gender congruent transparent targets paired with transparent distractors, and hence, 32/32 sharing the final letter with the distractor; conversely, there were 32 gender incongruent transparent targets paired with transparent distractors, and hence, 0/32 sharing the final letter with the distractor. If the coincidence between the final letter of the target and that of the distractor were of such relevance, we would have expected some kind of phonological facilitation effect for the former (transparent gender congruent) in comparison to the latter (transparent gender incongruent), yet none of the three experiments showed this, not even marginally[8]. It is hence more likely that the effect obtained in Experiment 2, in which opaque targets benefited from the presence of transparency congruent distractors in comparison with the presence of transparency incongruent distractors, is directly related to the routes of gender retrieval. In any case, note that this explanation is tentative. In fact, we could not replicate this effect of Experiment 2 in Experiment 3. There are two possible reasons underlying the absence of results in Experiment 3. On the one hand, it might be that, just like in Sá-Leite et al. (2021) second experiment with 25% of animate targets, the effect is rendered invisible. In our experiment we had 28% of animate targets. One of the explanations given by Sá-Leite et al. (2021) is that gender encoding is skipped in the absence of agreement (bare nouns) due to the semantic prioritization of animate nouns, this is, the more semantically richness shown by animate nouns that induces a deeper semantic processing and an increased usage of cognitive resources in comparison to inanimate nouns (Branigan et al., 2008). To preserve such resources and favor the production of the noun, our linguistic system would skip the processing of unnecessary characteristics, something in line with the tenets of the Independent Network model. In fact, the authors believe that the animates included in their experimental materials affected the way the rest of the stimuli were processed (i.e., if gender encoding is skipped for some targets, it ends up being skipped for every target,

[8] In this sense, there is evidence that the classic ortho-phonological facilitation effect in the PWI paradigm occurs based on the final phonemes of the stimuli pairs. Yet, this evidence is mainly obtained for the sharing of the final syllable or last two letters/phonemes when stress patterns are controlled (e.g., Ayora et al., 2011; Bürki et al., 2016; Meyer, 1996; Meyer & Schriefers, 1991; Wilshire et al., 2016). It is far from clear whether the sharing of only the final letter with mixed stressing patterns between targets and distractors could produce any kind of visible facilitation effect.

even inanimate nouns). This is because the effect of gender congruency is reduced and less significant when considering inanimate targets included in the lists with animate targets (in comparison to the effect of gender congruency restricted to the list with only inanimate targets). This would explain the absence of any type of effect related to gender retrieval (which includes transparency-related effects) in Experiment 3. On the other hand, since we actually increased the statistical power of Experiment 2 and we still failed to obtain the effect, the effect of transparency of Experiment 2 might simply be a fluke. Yet, it is worth noting here that recent studies using more sensitive and fine-grained techniques (Heim, Eickhoff, et al., 2009; Heim, Friederici, et al., 2009) found evidence of gender activation and competition/priming through haemodynamic responses registered with fMRI, using both a PWI paradigm and a primed picture-naming task with German participants. Hence, it would make sense to obtain effects of transparency related to the retrieval of gender, but RTs might not be sensitive enough to detect them consistently.

We believe that future research should replicate our study on this matter to confirm the GCE obtained in Experiment 1 and to understand what is going on in Experiment 2. The evidence coming from other measuring techniques, notably electrophysiological measures and fMRI, should help us to unveil how gender nodes are operating during lexical access. In this sense, these techniques should be quite informative regarding the time-course of determiner and gender selection in interaction with the different routes of gender retrieval.

Conclusions

In the present study, we conducted three PWI tasks in European Portuguese and obtained evidence of the classic GCE in a Romance language that has a one-to-one gender-form mapping. Our study ultimately supports the idea that the GCE reflects competition between determiners. Nevertheless, it also suggests that the retrieval of a determiner in a transparent language seems to be more complex than in opaque languages such as German or Dutch, as it depends on the flow of activation coming from two different routes of gender retrieval. This ultimately speaks in favor of the routes of gender retrieval proposed for language comprehension being also used in language production. Importantly, null results of gender congruency were obtained when participants named the pictures using bare nouns, although the transparency effect suggests that even though agreement is not present, activation is still flowing through gender nodes. Further replication of our GCE is necessary, as well as of studies featuring bare nouns and noun phrases with an appropriate sample size and perhaps more fine-grained techniques. This way, we would be able to properly answer the question of how gender is represented within the linguistic system and shed light on the mechanisms behind its retrieval.

Acknowledgments

We would like to thank Professor Ana Paula Soares and Professor Helena Oliveira for their help on an earlier version of this paper. We would also like to thank Reviewers Alba Casado and Sendy Caffarra for their insightful comments.

Funding: This work was supported by the Government of Spain, Ministry of Education and Vocational Training, through the Training program for Academic Staff (Ayudas para la Formación del Profesorado Universitario, FPU [FPU16/06983]); the Spanish Ministry of Science and Innovation [research project PID2019-110583GB-I00]; the Galician Government [grant for research groups ED431B 2022/2019]; and the Foundation for Science and Technology of Portugal [IF / 00784/2013 / CP1158 / CT0013]. Finally, the study has also been partially supported by the FCT and the Portuguese Ministry of Science, Technology and Higher Education through national funds and co-financed by FEDER through COMPETE2020 under the PT2020 Partnership Agreement [POCI-01-0145- FEDER-007653].

Conflict of interest

The authors do not have any conflict of interest to disclose.

Supplementary material

Appendix A contains other resources and data needed to replicate the study's results and claims.

References

Adamson, L., & Šereikaitė, M. (2019) Representing gender and defaults: Evidence from Lithuanian. *Glossa: A journal of General Linguistics*, *4*(1). https://doi.org/10.5334/gjgl.860

Alario, F. X., & Caramazza, A. (2002). The production of determiners: Evidence from French. *Cognition*, *82*, 179–223. https://doi.org/10.1016/S0010-0277(01)00158-5

Ayora, P., Peressotti, F., Alario, F.-X., Mulatti, C., Pluchino, P., Job, R., & Dell'Acqua, R. (2011). What phonological facilitation tells about semantic interference: A dual-task study. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00057

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Barreña, A. (1997). Desarrollo diferenciado de sistemas gramaticales en un niño vascoespañol bilingüe, en W. R. Glass & A. T. Pérez-Leroux (Eds.), *Contemporary Perspectives on the Acquisition of Spanish* (pp. 55-74). Cascadilla Press & Pennsylvania State University.

Bates, E., Devescovi, A., Pizzamiglio, L., D'Amico, S., & Hernandez, A. (1995). Gender and lexical access in Italian. *Perception & Psychophysics*, *57*(6), 847–847. https://doi.org/10.3758/BF03206800

Bates, D., Maechler, M, Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bi, Y., Wei, T., Janssen, N., & Han, Z. (2009). The contribution of orthography to spoken word production: evidence from Mandarin Chinese. *Psychonomic Bulletin & Review*, *16*(3), 555–560. https://doi.org/10.3758/PBR.16.3.555

Bordag, D., & Pechmann, T. (2008). Grammatical gender in speech production: Evidence from Czech. *Journal of Psycholinguistic Research*, *37*(2), 69–85. https://doi.org/10.1007/s10936-007-9060-0

Boutonnet, B., Athanasopoulos, P., & Thierry, G. (2012). Unconscious effects of grammatical gender during object categorisation. *Brain Research*, *1479*, 72–79. https://doi.org/10.1016/j.brainres.2012.08.044

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations (with discussion). *Journal of the Royal Statistical Society: Series B (Methodological)*, *26*(2), 211–252. https://doi.org/10.1111/j.2517-6161.1964.tb00553.x

Branigan, H. P., Pickering, M. J., & Tanaka, M. (2008). Contributions of animacy to grammatical function assignment and word order during production. *Lingua*, *118*(2), 172–189. https://doi.org/10.1016/j.lingua.2007.02.003

Bürki, A., Sadat, J., Dubarry, A.-S., & Alario, F.-X. (2016). Sequential processing during noun phrase production. *Cognition*, *146*, 90–99. https://doi.org/10.1016/j.cognition.2015.09.002

Bürki, A., Besana, T., Degiorgi, G., Gilbert, R., & Alario, F. -X. (2020). Representation and selection of determiners with phonological variants. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(7), 1287–1315. https://doi.org/10.1037/xlm0000643

Caffarra, S., & Barber, H. A. (2015). Does the ending matter? The role of gender-to-ending consistency in sentence reading. *Brain Research*, *1605*, 83–92. https://doi.org/10.1016/j.brainres.2015.02.018

Caffarra, S., Janssen, N., & Barber, H. A. (2014). Two sides of gender: ERP evidence for the presence of two routes during gender agreement processing. *Neuropsychologia*, *63*, 124–134. https://doi.org/10.1016/j.neuropsychologia.2014.08.016

Caffarra, S., Siyanova-Chanturia, A., Pesciarelli, F., Vespignani, F., & Cacciari, C. (2015). Is the noun ending a cue to grammatical gender processing? An ERP study on sentences in Italian. *Psychophysiology*, *52*(8), 1019–1030. https://doi.org/10.1111/psyp.12429

Caramazza, A. (1997). How Many Levels of Processing Are There in Lexical Access? *Cognitive Neuropsychology*, *14*(1), 177–208. https://doi.org/10.1080/026432997381664

Caramazza, A., & Miozzo, M. (1997). The relation between syntactic and phonological knowledge in lexical access: Evidence from the “tip-of-the-tongue” phenomenon. *Cognition*, *64*(3), 309–343. https://doi.org/10.1016/S0010-0277(97)00031-0

Caramazza, A., Miozzo, M., Costa, A., Schiller, N., & Alario, F.-X. (2001). A Cross-linguistic Investigation of Determiner Production. In E. Dupoux (Ed.), *Language, brain, and cognitive development: Essays in honor of Jacques Mehler* (pp. 209–226). MIT Press.

Casado, A., Palma, A., & Paolieri, D. (2021). The scope of grammatical gender in Spanish: Transference to the conceptual level. *Acta Psychologica*, *218*, 103361. https://doi.org/10.1016/j.actpsy.2021.103361

Casado, A., Palma, A., & Paolieri, D. (2018). The Influence of Sex Information on Gender Word Processing. *Journal of Psycholinguistic Research*, *47*(3), 557–583. https://doi.org/10.1007/s10936-017-9546-3

Corbett, G. G. (2013). Number of Genders. In M. S. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online. Leipzig: Max Planck Institute for Evolutionary Anthropology* (Available online at http://wals.info/chapter/30, Accessed on 2021-09-28).

Corbett, G. G. (2000). *Number*. Cambridge University Press. https://doi.org/10.1017/CBO9781139164344

Corrêa, L. M. S., Augusto, M. R. A., & Castro, A. (2011). Agreement and markedness in the ascription of gender to novel animate nouns by children acquiring Portuguese. Journal of Portuguese Linguistics, 10(1), 121–142. https://doi.org/10.5334/jpl.103

Costa, A., Sebastián-Gallés, N., Miozzo, M., & Caramazza, A. (1999). The gender congruity effect: Evidence from Spanish and Catalan. *Language and Cognitive Processes*, *14*(4), 381–391. https://doi.org/10.1080/016909699386275

Cubelli, R., Lotto, L., Paolieri, D., Girelli, M., & Job, R. (2005). Grammatical gender is selected in bare noun production: Evidence from the picture–word interference paradigm. *Journal of Memory and Language*, *53*(1), 42–59. https://doi.org/10.1016/j.jml.2005.02.007

Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes*, *5*(4), 313–349. https://doi.org/10.1080/01690969008407066

Finocchiaro, C., Alario, F. X., Schiller, N. O., Costa, A., Miozzo, M., & Caramazza, A. (2011). Gender congruency goes Europe: A cross-linguistic study of the gender congruency effect in Romance and Germanic languages. *Rivista di Linguistica*, *23*(2), 161–198.

Fodor, I. (1959). The origin of grammatical gender. *Lingua*, *8*. 186–214. https://doi.org/10.1016/0024-3841(59)90020-8

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods*, *35*(1), 116–124. https://doi.org/10.3758/bf03195503

Foucart, A., Branigan, H. P., & Bard, E. G. (2010). Determiner selection in romance languages: Evidence from French. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(6), 1414–1421. https://doi.org/10.1037/a0020432

Gollan, T. H., & Frost, R. (2001). Two routes to grammatical gender: Evidence from Hebrew. *Journal of Psycholinguistic Research*, *30*(6), 627–651. https://doi.org/10.1023/a:1014235223566

Guasch, M., Boada, R., Ferré, P., & Sánchez-Casas, R. (2013). NIM: A web-based Swiss army knife to select stimuli for psycholinguistic studies. *Behavior Research Methods*, *45*(3), 765–771. https://doi.org/10.3758/s13428-012-0296-8

Harris, J. (1991). The exponence of gender in Spanish. *Linguistic Inquiry*, *22*(1), 27–62.

Hartsuiker, R. J., & Kolk, H. H. J. (2001). Error monitoring in speech production: A computational test of the perceptual loop theory. *Cognitive Psychology*, *42*(2), 113–157. https://doi.org/10.1006/cogp.2000.0744

Heim, S., Eickhoff, S. B., Friederici, A. D., & Amunts, K. (2009). Left cytoarchitectonic area 44 supports selection in the mental lexicon during language production. *Brain Structure and Function*, *213*(4-5), 441–456. https://doi.org/10.1007/s00429-009-0213-9

Heim, S., Friederici, A. D., Schiller, N. O., Rüschemeyer, S. A., & Amunts, K. (2009). The determiner congruency effect in language production investigated with functional MRI. *Human brain mapping*, *30*(3), 928–940. https://doi.org/10.1002/hbm.20556

Ibrahim, M. H. (1973). *Grammatical Gender: Its Origin and Development*. De Gruyter Mouton. https://doi.org/10.1515/9783110905397

Jescheniak, J. D., Schriefers, H., & Lemhöfer, K. (2014). Selection of freestanding and bound gender-marking morphemes in speech production: A review. *Language, Cognition and Neuroscience*, *29*(6), 684–694. https://doi.org/10.1080/01690965.2012.654645

Kilarski, M. (2005). On grammatical gender as an arbitrary and redundant category. In D. A. Kibbee (Ed.), *History of Linguistics 2005: Selected papers from the Tenth International Conference on the History of the Language Sciences (ICHOLS X), 1–5 September 2005, Urbana-Champaign, Illinois* (pp. 24-36). John Benjamins Publishing Company. https://doi.org/10.1075/sihols.112.03kil

Kuchenbrandt, I. (2008). Cross-linguistic influences in the acquisition of grammatical gender? *Arbeiten zur Mehrsprachigkeit*, *86*, 1–28.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

La Heij, W., Mak, P., Sander, J., & Willeboorsde, E. (1998). The gender-congruency effect in picture word task. *Psychological Research*, *61*(3), 209–219. https://doi.org/10.1007/s004260050026

Lenth, R. V. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software*, *69*(1), 1–33. https://doi.org/10.18637/jss.v069.i01

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–75. https://doi.org/10.1017/S0140525X99001776

McGuire, P. K., Silbersweig, D. A., & Frith, C. D. (1996). Functional neuroanatomy of verbal self-monitoring. *Brain*, *119*(Pt. 3), 907–917. https://doi.org/10.1093/brain/119.3.907

Meyer, A. S. (1996). Lexical access in phrase and sentence production: results from picture–word interference experiments. *Journal of Memory and Language*, *35*(4), 477–496. https://doi.org/10.1006/jmla.1996.0026

Meyer, A. S., & Schriefers, H. (1991). Phonological facilitation in picture-word interference experiments: Effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(6), 1146–1160. https://doi.org/10.1037/0278-7393.17.6.1146

Miozzo, M., & Caramazza, A. (1999). The selection of determiners in noun phrase production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(4), 907–922. https://doi.org/10.1016/S0749-596X(02)00531-4

Miozzo, M., Costa, A., & Caramazza, A. (2002). The absence of a gender congruency effect in romance languages: A matter of stimulus onset asynchrony? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(2), 388–391. https://doi.org/10.1037/0278-7393.28.2.388

New, J., Cosmides, L., & Tooby, J. (2007). Category-specific attention for animals reflects ancestral priorities, not expertise. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(42), 16598–16603. https://doi.org/10.1073/pnas.0703913104

O'Rourke, P. (2007). The gender congruency effect in bare noun production in Spanish. *Coyote Papers: Working Papers in Linguistics, Linguistic Theory at the University of Arizona*, *15*, 66–89.

Paolieri, D., Lotto, L., Morales, L., Bajo, T., Cubelli, R., & Job, R. (2010). Grammatical gender processing in romance languages: Evidence from bare noun production in Italian and Spanish. *European Journal of Cognitive Psychology*, *22*(3), 335–347. https://doi.org/10.1080/09541440902916803

Paolieri, D., Lotto, L., Leoncini, D., Cubelli, R., & Job, R. (2011). Differential effects of grammatical gender and gender inflection in bare noun production. *The British Psychological Society*, *102*(1), 19–36. https://doi.org/10.1342/000712610X496536

Pinheiro, J. C., & Bates D. M. (2000). *Mixed-Effects Models in S and S-PLUS*. Springer-Verlag.

Protopapas, A. (2007). Check Vocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods*, *39*(4), 859–862. https://doi.org/10.3758/bf03192979

Roelofs, A. (2004). Error biases in spoken word planning and monitoring by aphasic and nonaphasic speakers: comment on Rapp and Goldrick (2000). *Psychological Review*, *111*(2), 573–570. https://doi.org/10.1037/0033-295X.111.2.561

Roelofs, A. (2008). Tracing attention and the activation flow of spoken word planning using eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(2), 353–368. https://doi.org/10.1037/0278-7393.34.2.353

Roelofs, A., Meyer, A. S., & Levelt, W. J. M. (1998). A case for the lemma/lexeme distinction in models of speaking: Comment on Caramazza and Miozzo (1997). *Cognition*, *69*(2), 219–230. https://doi.org/10.1016/s0010-0277(98)00056-0

Roelofs, A., & Piai, V. (2015). Aspects of competition in word production: reply to Mahon and Navarrete. *Cortex*, *64*, 420–424. https://doi.org/10.1016/j.cortex.2014.10.016

Sá-Leite, A. R., Fraga, I., & Comesaña, M. (2019). Grammatical gender processing in bilinguals: An analytic review. *Psychonomic Bulletin & Review*, *26*(4), 1148–1173. https://doi.org/10.3758/s13423-019-01596-8

Sá-Leite, A. R., Haro, J., Comesaña, M., & Fraga, I. (2021). Of beavers and tables: The role of animacy in the processing of grammatical gender within a picture-word interference task. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.661175

Sá-Leite, A. R., Luna, K., Fraga, I., & Comesaña, M. (2020). The gender congruency effect across languages in bilinguals: A meta-analysis. *Psychonomic Bulletin & Review, 27*(4), 677–693. https://doi.org/10.3758/s13423-019-01702-w

Schiller, N. O., & Caramazza, A. (2003). Grammatical feature selection in noun phrase production: Evidence from German and Dutch. *Journal of Memory and Language*, *48*(1), 169–194. https://doi.org/10.1016/S0749-596X(02)00508-9

Schiller, N. O., & Caramazza, A. (2006). Grammatical gender selection and the representation of morphemes: The production of Dutch diminutives. *Language and Cognitive Processes*, *21*(7-8), 945–973. https://doi.org/10.1080/01690960600824344

Schriefers, H. (1993). Syntactic processes in the production of noun phrases. *Journal of Experimental Psychology: Learning, memory, and Cognition*, *19*(4), 841–850. https://doi.org/10.1037/0278-7393.19.4.841

Schriefers, H., & Teruel, E. (2000). Grammatical gender in noun phrase production: The gender interference effect in German. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(6), 1368–1377. https://doi.org/10.1037/0278-7393.26.6.1368

Soares, A. P., Costa, A. S., Machado, J., Comesaña, M., & Oliveira, H. (2017). The Minho Word Pool: Norms for imageability, concreteness and subjective frequency for 3,800 Portuguese words. *Behavior Research Methods*, *49*(3), 1065–1081. https://doi.org/10.3758/s13428-016-0767-4

Soares, A. P., Iriarte, A., Almeida, J. J., Simões, A., Costa, A., Machado, J., França, P., Comesaña, M., Rauber, A., Rato, A., & Perea, M. (2018). Procura-PALavras (P-PAL): A web-based interface for a new European Portuguese lexical database. *Behavior Research Methods*, *50*(4), 1461–1481. https://doi.org/10.3758/s13428-018-1058-z

Soares, A. P., Machado, J., Costa, A., Iriarte, Á., Simões, A., de Almeida, J. J., Comesaña, M., & Perea, M. (2015). On the advantages of word frequency and contextual diversity measures extracted from subtitles: The case of Portuguese. *Quarterly Journal of Experimental Psychology*, *68*(4), 680–696. https://doi.org/10.1080/17470218.2014.964271

Starreveld, P., & La Heij, W. (2004). Phonological facilitation of grammatical gender retrieval. *Language and Cognitive Processes, 19*(6), 677–711. https://doi.org/10.1080/01690960444000061

Szekely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., Jacobsen, T., Arévalo, A., Vargha, A., & Bates, E. (2005). Timed action and object naming. *Cortex*, *41*(1), 7–25. https://doi.org/10.1016/s0010-9452(08)70174-6

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., Lu, C. C., Pechmann, T., Pléh, C., Wicha, N., Federmeier, K., Gerdjikova, I., Gutierrez, G., Hung, D., Hsu, J., Iyer, G., Kohnert, K., Mehotcheva, T., Orozco-Figueroa, A., Tzeng, A., ... & Bates, E. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language*, *51*(2), 247–250. https://doi.org/10.1016/j.jml.2004.03.002

van Berkum, J. J. A. (1997). Syntactic processes in speech production: The retrieval of grammatical gender. *Cognition*, *64*(2), 115–152. https://doi.org/10.1016/S0010-0277(97)00026-7

Vasishth, S. & Gelman, A. (2021). How to embrace variation and accept uncertainty in linguistic and psycholinguistic data analysis. *Linguistics*. https://doi.org/10.31234/osf.io/zcf8s

Velnić, M. (2020) Acquisition of a transparent gender system: A comparison of Italian and Croatian. *Frontiers*, *11*. https://doi.org/10.3389/fpsyg.2020.571674

Vigliocco, G., Lauer, M., Damian, M., & Levelt, W. J. (2002). Semantic and syntactic forces in noun phrase production. *Journal of experimental psychology. Learning, memory, and cognition*, *28*(1), 46–58. https://doi.org/10.1037//0278-7393.28.1.46

Wilshire, C., Singh, S., & Tattersall, C. (2016). Serial order in word form retrieval: New insights from the auditory picture–word interference task. *Psychonomic Bulletin & Review*, *23*(1), 299–305. https://doi.org/10.3758/s13423-015-0882-8

Appendix A - Materials for all the experiments are tabulated herein.

Table A1. Target picture nouns per category in Experiments 1 and 2.

Transparent Feminine nouns		**Transparent Masculine nouns**		**Opaque Feminine nouns**		**Opaque Masculine nouns**	
IPNP name	**European Portuguese Translation**	**IPNP name**	**European Portuguese Translation**	**IPNP name**	**European Portuguese Translation**	**IPNP name**	**European Portuguese Translation**
obj005anchor	âncora	obj011arm	braço	obj062bridge	ponte	obj106comb	pente
obj033bathtub	banheira	obj048boat	barco	obj084tape	cassete	obj138drum	tambor
obj053bottle	garrafa	obj050bone	osso	obj094chimney	chaminé	obj204helmet	capacete
obj066butter	manteiga	obj081car	carro	obj097city	cidade	obj261microphone	microfone
obj073camera	câmara	obj085castle	castelo	obj101cloud	nuvem	obj283nose	nariz
obj095church	igreja	obj090cheese	queijo	obj116cross	cruz	obj293paintbrush	pincel
obj104dime	moeda	obj096cigarette	cigarro	obj163flower	flor	obj311pencil	lápis
obj212hose	mangueira	obj180glass	copo	obj225key	chave	obj355ring	anel
obj237leg	perna	obj182globe	globo	obj241lettuce	alface	obj367rug	tapete
obj258mask	máscara	obj203helicopter	helicóptero	obj282net	rede	obj377saxophone	saxofone
obj295palmtree	palmeira	obj281nest	ninho	obj341pyramid	pirâmide	obj431sun	sol
obj339pumpkin	abóbora	obj327plate	prato	obj375sandwich	sanduíche	obj439tank	tanque
obj368ruler	régua	obj352refrigerator	frigorífico	obj418spoon	colher	obj459tomato	tomate
obj392shirt	camisa	obj427strawberry	morango	obj448tennisracket	raquete	obj465tractor	trator
obj491wallet	carteira	obj467train	comboio	obj469tree	árvore	obj484vest	colete
obj510window	janela	obj496watch	relógio	obj492walnut	noz	obj506whip	chicote

Table A2. Distractors per target picture in each experimental condition in Experiments 1 and 2.

Transparent Feminine nouns			**Transparent Masculine nouns**			**Opaque Feminine nouns**			**Opaque Masculine nouns**		
Target	**Distractor**	**(transl.)**	**Target**	**Distractor**	**(transl.)**	**Target**	**Distractor**	**(transl.)**	**Target**	**Distractor**	**(transl.)**
âncora	panela	pan	braço	praia	beach	ponte	feira	fair/market	pente	alga	seaweed
	lábio	lip		turno	shift		vinho	wine		cílio	eyelash
	laje	slab		novidade	novelty		fome	hunger		cárie	caries
	ringue	ring		romance	romance		arte	art		vime	wicker
banheira	compota	jam	barco	praça	square	cassete	esfera	sphere	tambor	pérola	pearl
	amianto	asbestos		dedo	finger		gasóleo	diesel oil		sarampo	measles
	estirpe	strain		tese	thesis		génese	genesis		vaidade	vanity
	trombone	trombone		parque	park		folclore	folklore		rodapé	footer
garrafa	novela	novella	osso	jóia	jewel	chaminé	asneira	swear word	capacete	paródia	parody
	cadeado	padlock		bingo	bingo		donativo	donation		alarido	fuss
	caridade	charity		lente	lens		aeronave	aircraft		avalanche	avalanche
	iene	yen		cume	peak		vinagre	vinegar		frenesim	frenzy
manteiga	reitoria	rectorate	carro	linha	line	cidade	semana	week	microfone	nicotina	nicotine
	estuário	estuary		golo	goal		período	period		labirinto	maze
	humidade	humidity		voz	voice		noite	night		imunidade	immunity
	suspense	thriller		som	sound		gabinete	cabinet		champanhe	champaca
câmara	zona	zone	castelo	avenida	avenue	nuvem	dieta	diet	nariz	órbita	orbit
	efeito	effect		pelouro	office		fígado	liver		alívio	relief
	frente	front		amizade	friendship		ênfase	emphasis		celulose	cellulose
	regime	regime		impasse	deadlock		milagre	miracle		sotaque	accent
igreja	língua	tongue	queijo	parcela	parcel	cruz	fúria	fury	pincel	tigela	bowl
	diálogo	dialogue		prólogo	prologue		lenço	scarf		biombo	screen
	sorte	luck		rapidez	speed		nave	ship		acidez	acidity
	crime	crime		enclave	enclave		iate	yacht		bigode	moustache
moeda	pintura	paint	cigarro	muralha	wall	flor	máfia	mafia	lápis	cebola	onion
	palco	stage		pátio	courtyard		duelo	duel		débito	debit
	greve	strike		interface	interface		maré	tide		cabine	cabin
	convite	invitation		apetite	appetite		arroz	rice		duche	shower
mangueira	cebolada	onion sauce	copo	lema	motto	chave	escola	school	anel	erva	herb
	estrondo	crash		pacto	pact		título	title		boneco	puppet
	anuidade	annuity		peste	plague		ordem	order		foice	scythe
	croquete	croquette		lume	fire		índice	index		boné	cap
perna	amora	blackberry	globo	lagoa	pond	alface	pancada	blow	tapete	pimenta	chili
	dardo	dart		ofício	profession		esófago	esophagus		joelho	knee

	raiz	root		índole	character		cirrose	cirrhosis		praxe	initiation ritual
	gene	gene		armazém	warehouse		naipe	suit (cards)		xadrez	chess
máscara	academia	academy	helicóptero	horta	vegetable garden	rede	cabeça	head	saxofone	cafeína	caffeine
	contágio	infection		santuário	sanctuary		preço	price		armário	wardrobe
	barbárie	barbarism		humildade	humility		atitude	attitude		entorse	sprain
	charme	charm		bronze	bronze		nome	name		balancé	swing
palmeira	carroça	wagon	ninho	rampa	ramp	pirâmide	galáxia	galaxy	sol	mesa	table
	ginásio	gym		mastro	mast		charuto	cigar		corpo	body
	grafite	graphite		claque	fan group		metrópole	metropolis		posse	possession
	foguete	rocket		cofre	safe box		chocolate	chocolate		golpe	hit
abóbora	gorjeta	tip	prato	selva	jungle	sanduíche	caderneta	notebook	tanque	agulha	needle
	placebo	placebo		cabelo	hair		vestíbulo	hall		casino	casino
	neurose	neurosis		saudade	yearning		aldrabice	fraud		estante	shelf
	fetiche	fetish		exame	exam		raspanete	telling-off		arame	wire
régua	geleia	jelly	frigorífico	anedota	joke	colher	magia	magic	tomate	comarca	county
	frasco	jar		convénio	pact/agreement		sismo	earthquake		feitiço	spell
	couve	cabbage		planície	lowlands		elite	elite		bondade	goodness
	truque	trick		envelope	envelope		latim	latin		mármore	marble
camisa	relva	grass	morango	trincheira	trench	raquete	alheira	Portuguese sausage	trator	bengala	walking stick
	defeito	defect		estojo	case		espargo	asparagus		fósforo	match
	diocese	diocese		sirene	siren		vigarice	swindle		diálise	dialysis
	ventre	belly		vértice	vertex		filete	fillet		caixote	crate
carteira	placa	board	comboio	fórmula	formula	árvore	lousa	chalkboard	colete	orelha	ear
	oceano	ocean		universo	universe		tornado	tornado		reflexo	reflection
	higiene	hygiene		entidade	entity		tear	loom		lealdade	loyalty
	impacte	impact		jardim	garden		bazar	bazaar		alaúde	lute
janela	estátua	statue	relógio	piscina	pool	noz	família	family	chicote	amêndoa	almond
	navio	ship		cérebro	brain		outubro	october		hálito	breath
	síntese	synthesis		gravidez	pregnancy		imagem	image		marquise	marquee
	pacote	package		bilhete	ticket		futebol	soccer		berlinde	marble (toy)

Note. From top to bottom of the distractors columns: transparent feminine nouns, transparent masculine nouns, opaque feminine nouns, and opaque masculine nouns.

Table A3. Experimental target picture nouns per category in Experiment 3

Opaque Feminine nouns		**Opaque Masculine nouns**	
IPNP name	**European Portuguese Translation**	**IPNP name**	**European Portuguese Translation**
obj062bridge	ponte	obj017baby	bebé
obj084tape	cassete	obj106comb	pente
obj094chimney	chaminé	obj138drum	tambor
obj097city	cidade	obj145elephant	elefante
obj101cloud	nuvem	obj199hat	chapéu
obj116cross	cruz	obj204helmet	capacete
obj163flower	flor	obj226king	rei
obj241lettuce	alface	obj256man	homem
obj282net	rede	obj261microphone	microfone
obj341pyramid	pirâmide	obj283nose	nariz
obj375sandwich	sanduíche	obj338priest	padre
obj418spoon	colher	obj408snail	caracol
obj448tennisracket	raquete	obj431sun	sol
obj469tree	árvore	obj459tomato	tomate
obj492walnut	noz	obj475turkey	peru
obj515woman	mulher	obj157fish	peixe

Table A4. Experimental distractors per target picture in each experimental condition in Experiment 3.

Feminine target nouns			Masculine target nouns		
Target	**Distractor**	**transl.**	**Target**	**Distractor**	**transl.**
ponte	feira	fair	bebé	letra	letter
	vinho	wine		eixo	axis
	fome	hunger		raiz	root
	sangue	blood		lazer	recreation
cassete	terapia	therapy	pente	lapa	cave/limp
	ginásio	gym		talo	stalk
	génese	genesis		vide	grapevine
	penálti	penalty		rim	kidney
chaminé	asneira	swear word	tambor	gaivota	seagull
	aquário	aquarium		autismo	autism
	aeronave	aircraft		osmose	osmosis
	vinagre	vinegar		clip	clip
cidade	semana	week	elefante	inércia	inertia
	exemplo	example		centeio	rye
	noite	night		lealdade	loyalty
	nível	level		tamboril	monkfish
nuvem	dieta	diet	chapéu	fortuna	fortune
	figado	liver		oceano	ocean
	ênfase	emphasis		diocese	diocese
	farol	lighthouse		aluguer	rental
cruz	fúria	fury	capacete	paródia	parody
	quilo	kilo		alarido	fuss
	nave	ship		avalanche	avalanche
	réu	defendant		frenesim	frenzy
flor	máfia	mafia	rei	praia	beach
	duelo	duel		estilo	style
	maré	tide		frase	phrase
	edil	edile		som	sound
alface	gaiola	cage	homem	câmara	camera

	umbigo	belly button		mundo	world
	mascote	pet		crise	crisis
	delfim	dolphin		clube	club
rede	banda	band	microfone	astronomia	astronomy
	preço	price		labirinto	maze
	atitude	attitude		imunidade	immunity
	mar	sea		calcanhar	heel
pirâmide	galáxia	galaxy	nariz	placa	board
	charuto	cigar		convento	convent
	metrópole	metropolis		elite	elite
	chocolate	chocolate		hóquei	hockey
sanduíche	caderneta	notebook	padre	tarefa	assignment
	vestíbulo	hall		metro	subway
	aldrabice	fraud		défice	deficit
	expositor	exhibitor		convite	invitation
colher	magia	magic	caracol	alcova	shelter
	sismo	earthquake		granizo	hail
	elite	elite		benesse	bless
	recital	recital		abdómen	abdomen
raquete	alheira	Portuguese sausage	sol	mesa	table
	canário	canary		corpo	body
	necrose	necrosis		arte	art
	coiote	coyote		casal	couple
árvore	norma	norm	tomate	estátua	statue
	estágio	internship		ozono	ozone
	fraude	fraud		garagem	garage
	cartaz	placard		fervor	fervour
noz	pinça	tweezers	peru	úlcera	ulcer
	veado	deer		cílio	eyelash
	variz	varicose		cárie	caries
	bazar	bazaar		funil	funnel
mulher	família	family	peixe	teoría	theory

outubro	October	inverno	winter
imagem	image	lente	lens
futebol	soccer	café	coffee

Note. From top to bottom of the distractors columns: transparent feminine nouns, transparent masculine nouns, opaque feminine nouns, and opaque masculine nouns.