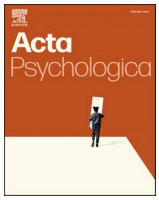
Acta Psychologica 246 (2024) 104236

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy



A standardized approach to measuring gender transparency in languages

Ana Rita Sá-Leite [a,b,*], Ian Craig Simpson [c,d], Isabel Fraga [b], Montserrat Comesaña [e]

[a] Institut für Romanische Sprachen und Literaturen, Goethe University Frankfurt, Germany
[b] Cognitive Processes & Behaviour Research Group, Department of Social Psychology, Basic Psychology, and Methodology, University of Santiago de Compostela, Spain
[c] Mind, Brain and Behavior Research Center (CIMCYC), University of Granada, Granada, Spain
[d] Department of Experimental Psychology, University of Granada, Granada, Spain
[e] Psycholinguistics Research Line, CIPsi, School of Psychology, University of Minho, Portugal

ARTICLE INFO

Keywords:
Grammatical gender
Gender transparency
Gender acquisition
Gender processing

ABSTRACT

Languages can express grammatical gender through different ortho-phonological regularities present in nouns (e.g., the cues "-o" and "-a" for the masculine and the feminine respectively in Italian, Portuguese, or Spanish). The term "gender transparency" was coined to describe these regularities (Bates et al., 1995). In gendered languages, we can hence distinguish between transparent nouns, i.e., those displaying form regularities; opaque nouns, i.e., those with ambiguous endings; and irregular nouns, i.e., those that display the typical form regularities but are associated with the opposite gender. Following a descriptive analysis of such regularities, languages have been recently classified according to their degree of gender transparency, which seems relevant in regard to gender acquisition and processing. Yet, there are certain inconsistencies in determining which languages are overall transparent and which are opaque. In particular, it is not clear whether some other complex regularities such as derivational suffixes are also "transparent" cues for gender, what really constitutes an "opaque" noun, or which role orthography and morphology have in transparency. Given the existing inconsistencies in classifying languages as transparent or opaque, this work introduces a proposal to assess gender transparency systematically. Our methodology adapts the standardized factors proposed by Audring (2019) to analyse the relative complexity of gender systems. Such factors are adapted to gender transparency on the basis of the literature on gender acquisition and processing. To support the feasibility of such a proposal, the concepts have been instantiated in a quantitative model to obtain for the first time an objective measure of gender transparency using European Portuguese and Dutch as instances of target languages. Our results coincide with the theoretically expected outcome: European Portuguese obtains a high value of gender transparency while Dutch obtains a moderately low one. Future adaptations of this model to the gender systems of other languages could allow the continuum of gender transparency to sustain robust predictions in studies on gender processing and acquisition.

1. Introduction

Grammatical gender is a lexical-syntactic feature present in languages that have a gender system (Corbett, 1991). Gender systems classify nouns according to two or more gender values (e.g., "masculine" and "feminine"). Nouns contain the gender information, and other words such as adjectives or determiners have to agree with the gender of the head noun by adapting their form and hence express this gender information (e.g., in Portuguese, 'table' [*mesa*] is feminine: *Uma mesa é barata* ['a table is cheap'], but 'car' [*carro*] is masculine, and hence, we say *Um carro é barato* ['a car is cheap']). In addition, a relationship can also exist between noun form and gender, as certain form-based regularities related to gender have been identified in a variety of languages (e.g., "-o" as in *carro* $_{\text{Masculine}}$ ['car'] is the common ending for the masculine in Italian, Portuguese, or Spanish, see Gudmundson, 2010; Harris, 1991; Sá-Leite, 2021). This link between noun form and gender is frequently referred to as "gender transparency" (Bates et al., 1995).

The concept of gender transparency has been defined on the basis of a nominal classification that distinguishes between transparent, opaque, and irregular nouns. Transparent nouns are those with endings correlating with specific gender values, as evidenced in Italian, Spanish, or Portuguese nouns ending in '-o' for masculine and '-a' for feminine (e.g., in Portuguese masculine nouns *carro* ['car'] and *livro* ['book'], and feminine nouns *mesa* ['table'] and *casa* ['house']). Conversely, opaque

* Corresponding author at: Department of Social Psychology, Basic Psychology, and Methodology, 15782, University of Santiago de Compostela, Spain.
E-mail addresses: saleited@rz.uni-frankfurt.de, anarita.saleite.dias@usc.es (A.R. Sá-Leite).

https://doi.org/10.1016/j.actpsy.2024.104236
Received 4 August 2023; Received in revised form 11 March 2024; Accepted 21 March 2024
Available online 12 April 2024

nouns have indeterminate gender correlations, like the "-e" ending in Italian, Portuguese, and Spanish (e.g., in Portuguese, *árvore* or *torre* [feminine, 'tree', 'tower'], and *pente* or *chocolate* [masculine, 'hair-brush', 'chocolate']). Irregular nouns, despite having transparent cues (e.g., "-o" and "-a"), belong to the opposite gender value, such as *rádio* (feminine, 'radio') or *problema* (masculine, 'problem') in Portuguese (Bates et al., 1995; Paolieri et al., 2010; Sá-Leite, Luna, et al., 2022). We can thus say that transparency is mainly defined by statistical correlations between gender values and nominal endings. These correlations evolved historically (Villalva, 2008) and have become cues for gender assignment. Importantly, based on this concept, languages are usually treated in the literature as either more gender transparent or opaque. Kupisch et al. (2018) recently proposed a continuum of transparency displaying a series of languages from the most to the least transparent: Spanish, Venetian, Italian, Russian, French, German, Norwegian, Swedish, and Dutch (see Fig. 1). In short, the higher the gender transparency, the easier it is to predict gender from the noun form. This continuum was later updated by Velnić (2020), who included Croatian. Indeed, this type of classification is not trivial: whether one language is more or less transparent seems to have consequences for the way gender is acquired and processed during lexical access (e.g., Gollan & Frost, 2001; Kupisch et al., 2022; Rodina & Westergaard, 2017; Sá-Leite et al., 2023).

However, we believe that the correlational basis usually underlying the definition of gender transparency leaves aside a series of aspects that seem fundamental to understanding the differences and similarities of the relationship between word form and gender across languages. While languages such as Spanish, Italian, Portuguese, and Russian are often labelled as transparent, and German and French as opaque, the criteria for these classifications remain ambiguous. In particular, languages like German and Norwegian might be deemed non-transparent due to their complex morphologies with numerous regularities and exceptions (Rodina & Westergaard, 2017; Salamoura & Williams, 2007; Spinner & Juffs, 2008). Meanwhile, a language such as Italian is gender transparent because it has a simple morphological system with two main cues that rarely fail to predict gender, "-o" and "-a" (Bates et al., 1995). Yet, we think that the basis on which the degree of gender transparency of languages is grounded can be considered in more detail across languages.

In this work, we intend to identify a series of aspects that might require a more fine-grained analysis across languages and to subsequently establish what we believe could be a more accurate view to determine the degree of gender transparency of a language. We will highlight the disregard of derivational suffixes as transparent cues in the so-called transparent languages in contrast to the so-called opaque ones, the unclear definition of opaqueness, and the often-overlooked role of orthography and morphology in gender transparency. A clarification of these points will be used to design a standardized method to quantitatively assess the degree of gender transparency of any language. Such a method is presented following the concept of relative complexity of gender systems by Audring (2019), and its quantification will be based mainly on European Portuguese, but also on Dutch as a means of establishing a comparison. Our aim is therefore to create a first approach based on a language whose available lexical databases include gender as a feature.

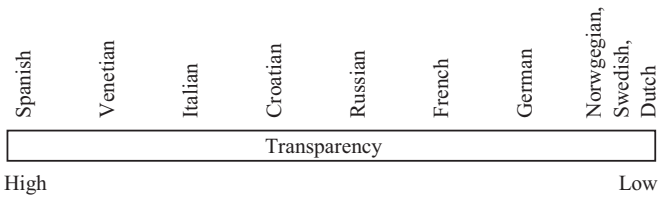


Fig. 1. Proposal of a continuum of gender transparency.
Note. Adapted from Kupisch et al. (2018) and Velnić (2020).

1.1. Gaps within the current concept of gender transparency across languages

1.1.1. Transparent nouns: the case of derivational suffixes

What falls into the category of transparent nouns seems to diverge across transparent and opaque languages. A quick inspection of the literature on gender acquisition, processing, and categorization with languages such as Italian, Portuguese, and Spanish shows that the transparent category is systematically reduced to nouns displaying gender cues "-o" and "-a" (e.g., Arias-Trejo & Alva, 2013; Arias-Trejo et al., 2013; Bates et al., 1995, 1996; Caffarra & Barber, 2015; Caffarra et al., 2014, 2015, 2017; De Martino et al., 2011, 2017; Garrido-Pozú, 2022; Hernandez et al., 2007, 2004; Lindsey & Gerken, 2012; Quiñones et al., 2018; Paolieri et al., 2010; Russo et al., 2021; Sá-Leite et al., 2021; Sá-Leite, Tomaz, et al., 2022; Urrutia et al., 2009). Nevertheless, there is a whole body of nouns in transparent languages that show more complex regularities, i.e., derivational suffixes and pseudosuffixes that strongly correlate with gender.

In Spanish, the suffix "*-aje*" is added to verbs to form nouns that express action, and those nouns are unequivocally masculine (*aprender* ➔ *aprendizaje* ['to learn' ➔ 'apprenticeship/learning']). Another example is the suffix "-ción", which is used to form nouns from verbs and is fully feminine (*competir* ➔ *competición* ['to compete' ➔ 'competition']). We can identify at least another 13 suffixes in Spanish that correlate either perfectly or near perfectly with the feminine value, and at least 6 with the masculine value (Murillo, 1999; Shadid, 1997). The same is the case with several Portuguese suffixes ("-agem", *hospedagem* ['accomodation'] or "-ção" *competição* ['competition'], which are feminine) and Italian ones ("-zione", which is feminine, *comunicazione* ['communication']; Oliphant, 1998; Rio-Torto et al., 2013; Sá-Leite, 2021). Importantly, many of these suffixes are included in other words as pseudosuffixes, showing the same strong gender correlation as in their suffix usage (e.g., "-ción" as a pseudossufix in *loción* $_{\text{Feminine}}$ ['lotion'], *nación* $_{\text{Feminine}}$ ['nation'], *facción* $_{\text{Feminine}}$ ['faction']). Despite demonstrating a high gender correlation, nouns featuring these suffixes or pseudosuffixes are nonetheless often omitted, or classified as opaque in gender processing studies. The examples are many, for instance in Spanish: *mensaje* ('message'), *estación* ('station'), *canción* ('song'), *situación* ('situation'), *acción* ('action'), *nación* ('nation'), *libertad* ('freedom'), *jabón* ('soap'), and a large etcetera, which can be seen in the stimuli list of various studies (e.g., Caffarra & Barber, 2015; Garrido-Pozú, 2022; Urrutia et al., 2009). A similar example can be seen in Italian. De Martino et al. (2017) examined grammatical gender and suffix transparency in processing Italian written nouns. Although a full stimulus list was not provided, one of the examples given is *stazione* ('station'), which is included as an opaque noun ending in "-e" even though the suffix "-zione" is feminine (Chini, 1995).

Other less transparent languages, such as German and French, also have multiple suffixes that correlate with gender. Yet, in contrast to what generally happens in transparent languages, researchers do not ignore the fact that they are strongly correlated with gender and include them within their list of transparent nouns, finding processing differences in comparison to opaque nouns (e.g., in German, among many others, "-heit" or "-keit" as feminine, "-ling" as masculine, or "-chen" as neuter; Heidolph et al., 1984; MacWhinney et al., 1989; in French, among others "-eur" and "-ie" are feminine suffixes, whereas "-eau" and "-ot" are masculine suffixes; see Ayoun, 2018; Desrochers et al., 1989; Hohlfeld, 2006; Szagun et al., 2007; Taft & Meunier, 1998).

In summary, in transparent languages transparent nouns are usually defined as ending in "-o" and "-a" while (pseudo)suffixes are ignored as gender cues. In contrast, in opaque languages, (pseudo)suffixes are generally considered to be gender cues. Therefore, the transparent category in transparent languages is not accurately defined as it is broader than usually assumed. It might be actually similar in size and complexity to the transparent category of other (more opaque) languages. The broadness of the category is even greater and the difference

with more opaque languages smaller once we dive more fully into what we understand by opaqueness, as seen in the next section.

1.1.2. Transparent nouns within the typical opaque category

Another gap that can be found in the current definition of gender transparency is related to the view of opaque nouns as those whose endings do not correlate with any gender. On the one hand, we would like to highlight the fact that a fifty-fifty situation between values and form is not so often observed (for a distributional analysis of nominal endings according to transparency in Portuguese, see Sá-Leite, 2021; for Spanish, see Clegg, 2011). It is hence unclear what the exact cut-off point is for an ending to be considered opaque within a stimuli list, and future studies should explore this further. On the other hand, what concerns us the most is that describing a series of endings as opaque within transparent languages simply because they are not "-o" and "-a" is not accurate. Yet, the literature regarding online gender processing shows a strong tendency in doing so (one exception is Afonso et al., 2014).

In languages like German or French, dozens of gender regularities have been identified. For instance, in German, more than 44 word-form regularities have been identified in monomorphemic and extended monomorphemic nouns, with numerous exceptions and nuances depending on lexical concreteness, sonority, and case, among other factors (e.g., Köpcke, 1982; Köpcke & Zubin, 1983). The number of regularities increases if we consider the huge proportion of polysyllabic words. In French, dozens of regularities have been identified as well, for instance in the final phonemes (e.g., "/m/" for the masculine; Ayoun, 2018), but also in other multiple combinations of phonemes and graphemes, suffixes and larger word units; importantly, when syllables come into play, the situation turns out to be even more intricate (e.g., /ɔ̃/, which is feminine in 70 % of cases, but whereas the ending /jɔ̃/ is feminine in 91 % of cases, /bɔ̃/ and /rɔ̃/ are masculine in approximately 87 % of the time; Matthews, 2010). An excessive number of gender-form regularities has been said to be a reason for a gender system to be defined by gender opacity (e.g., Hopp, 2013; Szagun et al., 2007). It supposedly contrasts with the simple system of transparent languages based on "-o" and "-a".

Yet, leaving suffixes and pseudosuffixes aside, a proper analysis of transparent languages reveals that many other gender correlations can be established with other non-(pseudo)suffixal endings, similar to German or French. Endings usually considered opaque may be regular for gender. For example, the literature seems to suggest that gender correlations (of above 0.8) can be established between endings "-l", "-n", "-z", "-r", and "-e" in Spanish, especially when considering their combination with other graphemes/phonemes (e.g., "-ie" [*especie* ('species')] is a feminine ending; see Clegg, 2011). Hence, not even the typical opaque ending "-e" can be defined as completely opaque.

1.1.3. Transparency may not only be phonological

Bates and colleagues referred exclusively to "phonological" transparency when describing this term. However, two aspects outside phonology seem critical to understand the overall degree of gender transparency of languages: orthography and morphology.

Regarding orthography, some studies have warned that in many languages form-based transparency should not be reduced to phonology for the literate population (Colé et al., 2003; Matthews, 2010). In languages that are not orthographically transparent (i.e., those whose grapheme-to-phoneme feedback is not consistent, Lété et al., 2008), gender correlations often depend on patterns of orthographic and phonemic relationships. For instance, Spanish is an orthographically transparent language in which the typical ending "-a" for the feminine is gender transparent from the view of both phonology (/a/) and orthography ("-a/-á"). However, in French, many mismatches between phonology and orthography can complicate the identification of the distributional patterns between nominal endings and gender when speakers acquire the ability to read and write (e.g., the orthographic sequences "-ée" (*idée*, 'idea') and "-er" (*papier*, 'paper') correspond to the phoneme /e/, although the former marks the feminine in 92 % of cases and the latter always marks the masculine).

Regarding morphology, the coincidence between ortho-phonological gender regularities and gender morphemes might be important in the salience of ortho-phonological cues and the overall degree of gender transparency of a language. This could actually explain, at least partially, why "-o" and "-a" have captured the attention of so many and, as we will see in the next section, seem to be cognitively "special".

In a study on gender assignment concerning word form in French, Boloh and Ibernon (2013) argue that during language acquisition, influence can occur between semantic gender and grammatical gender or vice-versa (as in Russian, Rodina, 2008), especially because in languages with a binary gender system such as French, children can make a connection between the two subclasses of grammatical gender (masculine vs. feminine) to the ones in natural gender (male vs. female). In languages such as Italian, Spanish, and Portuguese, this process might be facilitated because transparent cues coincide to a great extent with the morphemes for natural gender (in Portuguese, *casa* ['house']/*carro* ['car'] *menina* ['female kid']/*menino* ['male kid']). This leads us to think that such a coincidence might indeed constitute a link between both kinds of gender, whose entries may feed into each other. Indeed, as Corbett (1991) stated, all gender systems have some kind of semantic basis. Ultimately, this link could perhaps facilitate the overall acquisition of gender and further increase the relevance of the endings "-o" and "-a" for the system. Moreover, "-a" is the feminine gender morpheme for every element of agreement (e.g., *aquela, esta, uma* [Portuguese feminine for 'that', 'this', 'a']), and "-o" is present in many of them (e.g., *amarelo, o* [Portuguese masculine for 'yellow' and 'the']. The coincidence between ortho-phonological gender cues and gender morphemes is such that the presence of the "-a" and "-o" as overall morpho-phonological markers of gender is extremely high, even in single sentences, as can be seen in these examples from Italian, Portuguese, and Spanish:

> L**a** bambin**a** dell**a** famiglI**a** appena arrivat**a** è un**a** studentess**a** istruit**a**.
> Il figli**o** piccol**o** del vicin**o** al terz**o** pian**o** è un ragazz**o** rumoros**o**.
> **A** menin**a** pequenin**a** d**a** famíli**a** recém chegad**a** é um**a** alun**a** educad**a**.
> **O** filh**o** pequen**o** d**o** vizinh**o** d**o** terceir**o** andar é um menin**o** barulhent**o**.
> L**a** niñ**a** pequeñ**a** de l**a** famili**a** recién llegad**a** es un**a** alumn**a** educad**a**.
> El hij**o** pequeñ**o** del vecin**o** del tercer**o** es un niñ**o** ruidos**o**.
> 'The little (female) child from the newly arrived family is a polite student.'
> 'The youngest son of the neighbour on the third floor is a noisy boy.'

Indeed, the ease of detection, identification, and active establishment of distributional patterns across elements of gender agreement can be higher than in the case of other languages, such as German:

> Di**e** gro**ße** Dam**e** der neu angekommenen Famili**e** ist ein**e** gebildet**e** Student**in**.
> Der jüngste Sohn des Nachbarn im dritt**en** Stock ist ein laut**er** Junge.
> 'The great lady of the newly arrived family is an educated student.'
> 'The youngest son of the neighbour on the third floor is a noisy boy.'

1.1.4. Interim summary: gaps on the concept of gender transparency

As can be seen in the preceding discussion, the criteria used within the field to classify nouns as transparent or opaque across gender transparent vs. opaque languages, is not consistent. In Table 1 we summarize the identified inconsistencies. Transparent and opaque languages might hence be more similar in their distribution of gender regularities than initially thought. This could have promoted imprecise results in cognitive studies comparing both types of languages when it

Table 1
Summary of the inconsistencies in the concept of gender transparency across studies.

	Transparent languages	Opaque languages
Suffixes and pseudosuffixes that correlate with gender	They are considered opaque within experimental research	They are considered transparent within experimental research
Simple endings that correlate with gender	They are considered opaque when different from "-o" and "-a"	They are considered transparent
A clear cutoff point on the strength of the gender-form correlation for opaque nouns	Absent	Absent
Consideration of orthography and morphology within the general concept of gender transparency	Mostly absent	Mostly absent

Note. Here we establish a simplified dichotomy between transparent languages (Spanish, Italian, Portuguese) and other languages that may be moderately opaque or highly opaque (from French to Dutch).

comes to the role of gender regularities in gender processing. However, in the next section we will see that research still shows clear differences between both types of languages, in that the "-o" and "-a" of transparent languages seem to have a unique role in gender acquisition and processing. This might be related to the last gap we mentioned: the fact that we may have been dismissing the importance of orthography and morphology on the overall degree of gender transparency of languages.

1.2. The cognitively special role of cues "-o" and "-a"

The literature shows that "-o" and "-a" still display a special link with gender in comparison to other gender cues such as suffixes, pseudosuffixes, and non-suffixal endings. More specifically, these three types of endings do not seem to play the same facilitating role in gender acquisition and categorization as "-o" and "-a" do.

1.2.1. Gender acquisition

The endings "-o" and "-a" have been systematically shown to be critical to identify and acquire the statistical patterns of agreement associated with each gender value in Spanish, Italian, or Portuguese (agreement is correctly established earlier for transparent than for opaque or irregular nouns, e.g., Lew-Williams & Fernald, 2007; Lindsey & Gerken, 2012; Pérez-Pereira, 1991). Their central role in identifying and acquiring gender is such that even though they are cues, they are temporarily treated as rules by children, perhaps due to the coincidence between phonology and gender morphology as seen in the previous section. In this vein, a common mistake during this process of gender acquisition is the overgeneralization of the gender cues "-o" and "-a", to the point that children have been recorded in Portuguese and Spanish converting irregular and opaque nouns into transparent ones (e.g., in Spanish *un árbolo*, instead of *un árbol* masculine ['a tree'], or *un azoto*, instead of *un azote* masculine ['a spank']; in Portuguese, *um dio* masculine instead of *um dia* ['a day']; e.g., Fernández-Fuertes et al., 2016; Figueira, 2013; Hernández-Pina, 1984; Mariscal, 2008). Yet, this has not been registered for derivational regularities – children have not been reported generalizing (pseudo)suffixal endings (e.g., "-ción" or "-aje") or changing them for the sake of agreement. Thus, even though (pseudo)suffixal endings are statistically strong (or perfect) cues for gender in these languages, the cognitive role they play in gender acquisition does not seem to be the same as cues "-o" and "-a".

1.2.2. Gender processing

In terms of gender processing, a common way of studying the encoding of gender cues is through gender categorization tasks, in which the participants are presented with nouns on a screen and must classify them according to gender by pressing different keys, with their response times (RTs) being measured. On the one hand, transparent nouns show lower RTs in a variety of languages, with very consistent results found for Italian and Spanish, which use nouns ending in "-o" and "-a" (e.g., Agirre, 2016; Bates et al., 1995, 1996; Caffarra et al., 2017; Hernandez et al., 2004; Padovani et al., 2005). Results also show lower RTs for Italian nouns in plural, which use "-i" and "-e" for the masculine and feminine, respectively (see De Martino et al., 2011). In contrast, in categorization tasks conducted in French, the role of gender transparency has been shown to be more dependent on agreement and to be diminished due to the phenomenon of elision, i.e., the neutralization of the definite article when the noun starts with a vowel (e.g., instead of definite articles *la* Feminine or *le* Masculine, French speakers must use *l'arbre* Masculine ['the tree'] or *l'aiguille* Feminine ['the needle']; Ayoun, 2018; Desrochers & Paivio, 1990; Desrochers et al., 1989; Royle & Stine, 2013; Taft & Meunier, 1998; see also Desrochers & Brabant, 1995). Transparent nouns only show lower RTs in comparison to opaque and irregular nouns when definite articles have the discriminatory form between gender values. On the other hand, Afonso et al. (2014) considered derivational suffixes as transparent in a couple of categorization tasks in Spanish and still found processing advantages for nouns ending in "-o" and "-a" over nouns ending in other suffixal and pseudosuffixal regularities when comparing transparent vs. irregular nouns (for similar Italian results, see Oliphant, 1998). Taken together, it seems that even though (pseudo)suffixes that correlate with gender are as regular as "-o" and "-a" are, they are not processed in the same way.

1.2.3. Interim summary: the cognitively special role of "-o" and "-a"

Although there are actually many other regularities within transparent languages such as Spanish, Italian, or Portuguese, "-o" and "-a" still seem to have a special relationship with gender, playing a central role in gender acquisition and being the main cues of gender retrieval at least within gender categorization tasks during noun lexical access.

1.3. Defining the degree of gender transparency of languages

As seen up until now, one of the unresolved challenges in the study of gender transparency in languages is the inconsistency in its definition and application across different languages. Our aim now is to reach a consensus on how to analyse the degree of gender transparency. In sum, studies exploring the processing of gender and gender regularities in transparent languages either leave aside many endings that are actually regular for gender, or consider them opaque. Transparency is hence not precisely defined in the majority of studies exploring transparent languages when considering its correlational basis. Yet, perhaps a correlational basis for gender transparency is *not* ideal, since if it were, (pseudo)suffixal strongly correlated endings in Italian, Portuguese, or Spanish should occupy a central facilitating role within the studies on gender acquisition and categorization, and that is not the case. In this sense, for a more accurate understanding of how to measure the degree of gender transparency of languages, it seems relevant to acknowledge the impact of phoneme-grapheme transparency and the matching between ortho-phonological gender cues and gender morphemes.

We will follow the notion of the relative complexity of gender systems as defined by Audring (2019) to propose a standardized approach to the understanding of gender transparency. This theoretical basis will allow us to present the first quantitative approach for the measurement of the degree of gender transparency in languages. This quantitative approach should allow us to obtain values for transparency that are

theoretically and empirically supported by the descriptive continuum of gender transparency by Kupisch et al. (2018).

2. A proposal to standardise the definition of gender transparency

In this section, we propose basing the degree of gender transparency of every language on the notion of the relative complexity of a gender system as defined by Audring (2019) following four factors: frequency, perspicuity, consistency, and monofunctionality of the gender cues. These factors will be applied to the analysis of the ortho-phonological gender cues on the one hand, and gender morphemes on the other. They will be applied taking into consideration the most ideal or canonical situation for a gender system: that in which the degree of gender transparency, and hence of simplicity, reaches its highest possible value. A table summing up these factors is presented at the end of the text (see Table 2). As we believe transparency is linked to cognition, the four factors will be discussed considering not only their statistical description, but also their cognitive role in both gender acquisition and processing. For this, we will follow two empirically supported principles. First, the simpler the gender system, the earlier its acquisition (see Audring, 2019, p. 39). Second, we will follow the idea supported by many models of language processing by which the more the speaker is exposed to a certain language feature, the stronger the links between the mental representations related to that feature (Taft, 2023; also Parmentier et al., 2017; Perea et al., 2013). Hence, we will apply this to gender transparency. Importantly, note that this approach is designed to assess the degree of gender transparency considering language processing in the case of literate people. Otherwise, the degree of gender transparency of a language should be assessed for both young children and illiterate adults by excluding an assessment of orthography. Accordingly, note that the influence of any of the proposed factors for language acquisition should be considered in the light of morpho-phonological cues, but not orthographic cues. Orthographic cues are relevant for gender transparency by considering gender processing, as they modify and have a role in the organization and architecture of lexical access and hence in gender retrieval. In any case, we will jointly refer to ortho-phonological cues to simplify the text. Finally, as explained by Audring (2019), the four factors are unavoidably intertwined, and the same aspect may affect more than one factor.

2.1. Frequency

Frequency is related to the number of times a speaker is exposed to a particular cue. The higher the proportion of transparent nouns, the higher the children's frequency of exposure to these cues.

Regarding ortho-phonology, we must consider the proportion of transparent nouns in comparison to that of irregular nouns. For language acquisition, the higher the proportion of phonologically gender transparent nouns, the higher the number of times the child is exposed to these cues, and hence the ease of acquisition will be higher. In general, the higher the proportion of ortho-phonologically gender transparent nouns, the stronger the link between that cue and its gender value. As for irregular nouns, for the sake of transparency, the lower the proportion of irregular nouns, the higher the frequency by which transparent cues are indeed regular and signal their correlated gender value, and hence the higher the degree of predictability of these cues. Thus, the most transparent gender system should be the one that only has transparent nouns; if there are also non-transparent nouns (i.e., opaque and irregular nouns), it is better for the sake of transparency if those are opaque rather than irregular. This is because for opaque nouns the transparent cue is not being targeted (i.e., the cue is not being linked to the opposite gender).

The number of times a speaker is exposed to those transparent cues can also be tied to the number of gender values, considering of course other factors such as the proportion of words per gender value and the frequency of use of these words. Assuming an approximate equal distribution of words across genders, a higher number of gender values in a language might mean that the child will be exposed to fewer lexical entries of each gender (or one of the gender values), and consequently, there will be less exposure to the transparent cues for each gender value. Additionally, a greater number of regularities per gender value means that the child will be less exposed to each of them. Therefore, it follows that the canonical gender transparent language would have a bipartite gender system with a "one-to-one" mapping between form regularities and gender values (one regularity per value).

Importantly, since the basis of frequency is "opportunity for exposure", transparency must be analysed considering the frequency of use of each noun in the language: how many nouns are transparent within the most frequently used nouns in a language. Since we are also considering first language acquisition, frequency of use should be measured from two perspectives: (1) frequency of use and oral exposure of nouns for toddlers considering the phonological cues of gender (as in the MacArthur Communicative Development Inventory; Jackson-Maldonado et al., 1993), and (2) frequency of use in general considering the ortho-phonological cues of gender, mainly through materials based on written sources that reflect spoken production, such as databases of film subtitles (e.g., SUBTLEX; Keuleers et al., 2010; Soares et al., 2015).

In a broader sense, when addressing morphology, we should ideally have the highest possible number of elements of agreement displaying gender morphemes. This also applies to nouns with natural gender, which can use morphemes to express the biological or perceived sex of the referent (e.g., for the feminine, "-a" in Spanish and "-in" in German, as in *profesora* and *Lehrerin*, 'female teacher'). If these gender morphemes coincide with the ortho-phonological gender cues, the overall frequency of use and exposure to the cues (e.g., "-o" and "-a") will also rise. Importantly, during language acquisition, the higher the number of cases in which the phonological forms are related to the gender values, the easier will be the statistical learning of the gender-distributional patterns within the language. In line with this, an interesting aspect noted by Audring (2019) is that semantically motivated gender might be acquired later in a language, since there is lower frequency in the input. However, this delay might be reduced due to the coincidence between phonological gender cues and gender morphemes in natural gender (as in Portuguese, Italian and Spanish). In this sense, this coincidence might be one of the reasons for the fast acquisition of gender reported for languages like Portuguese, Italian, or Spanish.

2.2. Perspicuity

Perspicuity refers to the clarity of expression. In regards to gender transparency, it is related to the clarity and simplicity of the cues for gender; such cues are expected to work best when they are perspicuous and clear. So, the more perspicuous the cues are, the higher the degree of gender transparency. For instance, overall, the Spanish gender system is more perspicuous than that of German. The masculine value in Spanish is mostly represented by the one-letter/phoneme cue "-o". The masculine value in German is mostly represented by "-en", "-er", and "-el", three two-letter/phoneme cues, which are not only indicators of gender but also of number and case. In terms of gender transparency, the masculine value in Spanish is hence simpler and clearer (more perspicuous) than in German.

In terms of ortho-phonology, the clarity and simplicity of the cues may be affected by the mismatch between orthography and phonology: if the same phonological ending is represented by different orthographical forms, once writing is acquired, an extra grain of complexity may be added to the cognitive system. For instance, whereas in Spanish the final sounds /a/ and /o/ are represented exclusively by the orthographic endings "-a" and "-o", in French the ending /ɑ̃/ refers to the final orthographical units of "-a", "-and", "-ang", "-aon", "-ens", "-ends", "-empt", and "-ants", among others. Two alternative outcomes may be expected in this regard: i) on the one hand, the presence of multiple

graphemes representing a phoneme may help to disambiguate phonology. For instance, in French /e/ is a transparent cue for the feminine but corresponds to the orthographic sequences "-ée" and "-er", among others, with the former marking the feminine and the latter the masculine. These orthographic cues could in theory increase gender transparency by disambiguating the phonological cue; or ii) a higher number of regularities associated with each gender value may decrease the use of a form-based route to activate gender during lexical access. This is precisely what has been observed (see Gollan & Frost, 2001). In this sense, a lower number of regularities representing each gender node seem to increase the chances of the cognitive system to use word form to process gender. Hence, even though a greater number of graphemes could mean more cues to process gender and disambiguate between gender values in many instances, our system does not seem to actually benefit from them. Even though it is not clear when multiple cues are *too many*, in this method we will opt for considering that the higher the number of gender regularities associated with each gender value, the lower the exposure to each regularity individually, the lower the strength of the links between them and the gender nodes, and the lower the capability of the cognitive system to use them to process gender. Hence, the more gender orthographic or phonological regularities for each gender value, the lower the degree of gender transparency.

Related to this matching between phonology and orthography, an important concept here is that of syncretism. Syncretism refers to a linguistic phenomenon where distinct grammatical forms or categories in a language merge or overlap in certain contexts, resulting in a shared or identical form. Applied to gender transparency, if the same graphemes or phonemes are a cue for different functions, the gender transparency of this cue as a whole will be negatively affected. For instance, the ending "-e" in German is representative of both the feminine gender and the plural value (e.g., *Lampe* ['lamp'] is a feminine noun with a typical feminine ending; but *Tische* ['tables'] is a masculine noun in which "-e" marks the plural). Indeed, syncretism has been said to violate the principle of transparency (Audring, 2019), since it tends to create ambiguity and relative challenging situations for language learners (Storme, 2021; exception occurs when syncretism neutralizes a category, Finley, 2022). When applied to gender transparency, syncretism does not seem to promote the predictability and acquisition of gender or the use of word form to process gender later in life.

Syncretism in the agreement system of a language may also affect the capacity of children to identify the distributional patterns between ortho-phonological gender cues and the elements that go with the head noun. For instance, in Spanish, the seemingly masculine definite article *el* is used as feminine when the following feminine noun begins with a stressed "a" (e.g., ***alma*** ['soul'], ***arma*** ['gun'], ***águila*** ['eagle'], all start with "a" and have their first syllable as the stressed one [the diacritic mark is not necessary for disyllabic words]), and so is syncretic for the masculine and the feminine, diminishing its capability to predict the masculine value. In this sense, the existence of *cases* in a language tends to increase the complexity of the gender system and decrease its transparency, since it usually promotes syncretism. An element that is regular in one case may also be regular in another, creating a situation of irregularity for the element itself. For instance, the German definite article "der" is syncretic for nominative masculine (e.g., *der Hund bellt* ['the dog barks']), but also for genitive feminine (e.g., *das Auto der Familie ist alt* ['the car of the family is old']).

2.3. Consistency

The most transparent cues for gender are also the most consistent, and so an ideal cue has a unique form that consistently represents a particular gender value, and in an ideal transparent system, a particular gender value is consistently represented by the same cue. For less than perfect transparent languages (i.e. all real languages), the higher the consistency between a gender value and the cue/s, the higher the gender transparency of the language. Higher transparency would result in stronger connections between the gender values and their cues, and thus facilitate ease of acquisition of the cues as well as their processing. Paraphrasing Audring (2019), the validity of one cue is directly related to its predictive power. If a cue is associated with more than one gender value its predictive power decreases, decreasing the degree of transparency of the language. An example of this is the French phonological cue "-e": almost 50 % of the nouns displaying this cue are masculine, the other 50 % being feminine. Yet, a considerable proportion of all feminine nouns end in "-e", and "-e" is a feminine marker or agreement (e.g., masculine blue: *blau*, feminine blue: *blaue*). So, the feminine is represented by this cue, but this cue is also highly associated with the masculine value. This situation increases the inconsistency of this cue.

Importantly, when various gender cues represent the same gender value, a situation of inconsistency arises by which the gender value cannot be systematically represented by a unique cue. Therefore, more associations must be identified and learned and the strength of the gender cue-gender nodes connections is overall lower. In this sense, it is expected that the transparency of a "one-to-one" mapping situation (e.g., "-o" to masculine; Spanish, Italian, Portuguese) is higher than that of "various-to-one" mapping (e.g., non-palatal consonants to masculine; Russian).

A small number of opaque nouns also contribute to consistency. For example, Spanish contains a considerable proportion of masculine opaque nouns when compared to feminine opaque nouns, and hence the consistency of the masculine gender value with the cue "-o" is lower than that of the feminine gender value with the cue "-a". More specifically, around 60 % of masculine nouns end in "-o", while 77.5 % of feminine nouns end in "-a" (Teschner & Russell, 1984).

The matching between phonology and orthography is also relevant here, since the greater the number of orthographic representations of a certain phonological cue (as in French /ã/, represented by more than a dozen different orthographical units), the less the consistency of the orthographic form of the cue, and hence, in face of this inconsistency, the higher the difficulty for the speaker to generate the pattern of associations between this cue and gender.

In terms of morphology, the ideal situation is when gender morphemes coincide with the ortho-phonological cues for gender, since this creates the most consistent link between morpho-ortho-phonology and gender values and hence increases gender transparency. Conversely, the presence of gender morphemes that do not coincide with the ortho-phonological cues of gender diminishes the consistency. In Spanish, feminine agreement words are far more consistently represented, as they rely on the morpheme "-a", whereas the masculine relies on the morpheme "-o" but also on the morphemes "-e" and zero.[1]

2.4. Mono-functionality

Mono-functionality refers to the fact that transparent gender cues should display no other linguistic function than to statistically correlate with gender or to morphologically mark gender. Polyfunctionality, then,

[1] It is worth mentioning two fundamental concepts originally defined by Mills (1986), both of which are relevant for transparency as suggested by Caffarra et al. (2017): the concepts of availability and reliability, which involve the factors of frequency, perspicuity (syncretism), and consistency. Availability refers to the number of nouns displaying a certain form-function correspondence in a given language (i.e., transparent nouns minus non-transparent nouns). Reliability refers to the consistency of the relation between form and function in a given language (i.e., transparent nouns minus irregular nouns; Albright, 2002). The strength of a form-function mapping (i.e., cues for gender values) depends on both the availability and reliability of this mapping. Hence, when a probabilistic pattern covers a high number of items (high availability, large proportion of transparent nouns) and the number of exceptions is low (high reliability, low number of irregular nouns), it provides a strong distributional cue, which is considered to be influential in language acquisition and comprehension (MacWhinney, 1987, 2004; see also Culbertson et al., 2017).

increases the degree of relative complexity of the potential cues for gender and decreases their usefulness for gender acquisition and their relevance for gender processing.

Significant situations of polyfunctionality are those in which gender intertwines with other features such as number, case, and inflectional class. According to Audring (2019), any kind of polyfunctionality affects both clarity and consistency. In Italian, for instance, gender and *number* intertwine in the plural, as "-o" and "-a" are the usual endings for the masculine and feminine nouns in the singular, respectively, but "-i" and "-e" are used as markers for their respective plural forms. In contrast, Portuguese and Spanish add "-s" to form the plural, so any transparent noun with the "-o" to masculine or "-a" to feminine mapping in the singular will consistently maintain these cues in the plural, with an added "-s" (*casa* ['house'] to *casas* ['houses']). In this sense, an interesting idea to test the effect of polyfunctionality in gender transparency would be to make a direct comparison of gender acquisition in the plural between Spanish and Italian.

Even though *number* is relevant here, *case* seems to be the feature that may exponentially increase the relative complexity of the form assignment system. In fact, Eichler et al. (2013) points out that German gender is acquired later than Italian or Spanish gender, and attributes this not only to the cues being less transparent but to the influence of *case*.

Finally, polyfunctionality may also explain why derivational suffixes (e.g., in Spanish, "-ción" [feminine] or "-aje" [masculine]), which comply with the simple definition of "strong correlation between a nominal ending and a certain gender value", do not seem to have the same kind of role in gender acquisition and processing as the cues "-o" and "-a". Despite being cues for gender (although not simple ones), their function is something else: that of creating new words and assigning new functions to these words (e.g., to reflect action, to add a derogatory nuance to a noun, etc.). This may also explain why they are not overgeneralised in terms of gender – their function is something other than marking gender and hence children do not use these endings to make non-transparent nouns transparent when acquiring gender.

2.5. Definition of gender transparency and summary of the determining factors

We shall now attempt a more accurate definition of the degree of gender transparency of a language. We will define the degree of gender transparency of a language as the extent to which the form of nouns plays a role in the acquisition and processing of grammatical gender depending on the frequency and consistency of the associations between certain word patterns (of nouns themselves and agreement elements) with each gender value, as well as their perspicuity, syncretism, and mono-functionality. In Table 2, we present the summary of the above-mentioned determining factors of gender transparency.

3. A quantitative approach for the calculation of the degree of gender transparency

Up to this point, we have provided a definition of the concept of gender transparency based on a number of factors that we believe are relevant to turn this definition into a standardized one. The goal of this section is to instantiate the theoretical proposal in a model using real language. This was not done to present the model as *evidence* that the ideas exposed in this paper are correct; mainly because there is not enough behavioural data thus far that enable us to put it to test. Instead, our aim is twofold.

First, to demonstrate that it is possible to apply the concepts of this proposal to a real language and calculate a quantifiable, objective, plausible measurement for language transparency. However, given that this is the first such attempt to quantify transparency, the model, along with the concepts it is based on, will almost certainly require revision once new behavioural data become available. This leads to our second, and perhaps more important goal. At present, there is very little behavioural data in online noun gender retrieval (rather than gender acquisition) that directly addresses the concept of language transparency, and this would be due, in part, to the previously described inconsistency within the field with regard to the classification of nouns as transparent or opaque. By presenting a well-defined set of criteria, along with the associated model, we hope that researchers within the field will be inspired to think about gender transparency as a quantifiable entity, and design experiments that can verify, or falsify, the concepts we have presented.

Given our first goal, we needed to choose a language for which as much information as possible, relevant to the theoretical proposal, is currently available. Hence, we chose European Portuguese as the principal language to describe and demonstrate the model – to our knowledge, the available databases for European Portuguese provide more *relevant* information than that provided by databases of other languages. However, we also instantiated a second version of the model using Dutch. Although this second version of the model is more restricted owing to the reduced amount of relevant data contained in the Dutch

Table 2
Factors to operationalise the degree of gender transparency of languages.

Factor	Modality	Subfactors	Ideal scenario
Frequency	Ortho-phonology	Proportion of gender cues on nouns	Highest
		Number of gender cues per value	Lowest (one per gender value)
	Morphology	Proportion of gender morphemes in nouns and agreement elements	Highest
		Number of gender morphemes per value	Lowest (one per gender value)
	Both	Frequency of exposure and use during childhood	Highest possible number of nouns displaying phonological cues
		Overall frequency of exposure and use	Highest possible number of nouns displaying ortho-phonological cues
		Number of gender values	Lowest
Perspicuity	Ortho-phonology	Matching between phonology and orthography	Absolute (orthographically transparent language)
		Syncretism	Endings represent only gender
	Morphology	Syncretism (in agreement elements)	Agreement elements represent only one value
Consistency	Ortho-phonology	Proportion of irregular nouns	Lowest
		Proportion of opaque nouns	Lowest
		Mapping of gender values to gender cues	One-to-one
	Morphology	Coincidence between ortho-phonological gender cues and gender morphemes	Absolute
Monofunctionality	Ortho-phonology	Intertwining with other features	None
		Gender cues relation with other functions (e.g., derivation of words)	No

Note. The ideal scenario reflects the highest degree of gender transparency, which in terms of relative complexity will provide the highest degree of simplicity and ease of acquiring a gender system. Note that one unique situation may affect more than one factor (e.g., the existence of irregular nouns affects the proportion and frequency of use of gender cues, their syncretism, and consistency in terms of gender-to-value and value-to-gender linking; this had already been noted by Audring, 2019). Note also that for first language acquisition and illiterate speakers only phonological gender cues should be considered.

database, we thought it was important to create this second model for the following reason. In terms of gender, Portuguese is considered to be largely transparent whereas Dutch is considered to be generally opaque.[2] Consequently, by including two languages from opposite sides of the transparency continuum, we were able to provide evidence that the calculations underlying the model appear to be plausible.

The platform used for creating the model is Microsoft Excel, and this was chosen as its use is relatively widespread, and it can handle large amounts of data. It may also be more approachable for people with less technical backgrounds than, for example, Matlab. Although we use Portuguese and Dutch to describe the model, the design itself is meant to be generic and easily applied to other languages. Thus, anybody with a reasonable amount of experience with Excel should be able to take the examples presented in the supplementary materials and adapt them to other languages.

3.1. Overall strategy

One of the consequences that emerges from the four factors described in the first half of the present work is that all other things being equal, the more genders a language has, the less transparent it should be. Although it would be possible to design the model in such a way that, by default, a language with more genders receives a lower transparency score than a language with fewer genders, this would require several design decisions to be made a priori. For example, by how much is transparency reduced for each additional gender (10 %, 15 %, some other value)? Additionally, should this reduction be linear? That is, is the difference in transparency between a 2-gendered and 3-gendered language the same as the difference in transparency between a 5-gendered and a 6-gendered language? Or, is the second difference smaller because the *relative change* in the number of genders is different (a 50 % increase going from 2 to 3 genders, but only a 20 % increase going from 5 to 6 genders)? We believe that "hard wiring" values into the model for questions such as these without behavioural data as a guide would be unprincipled. Essentially, it would amount to a "guess and check" exercise – take an unmotivated design decision up front and then, at some future time when behavioural data becomes available, evaluate if the decision was correct. Hence, in creating our model, the overarching principle was to "let the data speak". Consequently, we designed the model to determine values for consistency, syncretism, etc., in a data-driven manner; that is, from the word characteristics inherent in the lexical databases themselves (Portuguese: P-PAL, Soares et al., 2014; Dutch: CELEX, Baayen et al., 1995). Languages with a higher number of gender values are still likely to come out less transparent as more values means more opportunities for exceptions to exist, although this is not guaranteed. A 3-gendered language with virtually no exceptions to the gender cues may be more transparent than a 2-gendered language which has many opaque and irregular gender cues.

Despite adopting a data-driven approach, due to a lack of empirical data, some a priori decisions about the model design were inevitable. Nevertheless, although not grounded in data, it was necessary to make these decisions in order to permit some components of the model to be computed, thus allowing for a fuller, more complete version of the model to be constructed. We will return to this point in the discussion.

[2] To understand the upcoming description of the model it is not necessary to know anything more about the Portuguese and Dutch gender systems beyond the fact that Portuguese is largely transparent and has two gender values (masculine and feminine) and Dutch is opaque and has also two gender values (common and neuter). Nevertheless, the interested reader can find a brief description of the gender system of European Portuguese, along with a description of the Portuguese and Dutch lexicons, upon which the model was applied, in the supplementary materials available on the Open Science Framework (OSF) at the following link: https://osf.io/74ezj/?view_only=e493c6cc2b3a4486b48c370d8575f2ed.

In summary, we have attempted to instantiate the factors and subfactors outlined in the first half of this present work using a design that is as free as possible from a priori decisions. Whether this design is correct will be answered by subsequent researchers who design experiments to test the transparency predictions made by this model for various languages. We next describe how each of the main factors has been implemented in the model.

3.2. General design principles

3.2.1. Frequency

When children are learning a language, multiple exposures to unambiguous gender cues should aid learning, as happens with other sublexical properties (Audring, 2019; Pierrehumbert, 2003; Treiman & Kessler, 2022). In contrast, multiple exposures to opaque and irregular gender cues should make learning gender cues more difficult (Kidd, 2012; Kupisch et al., 2002; Rodina & Westergaard, 2013). To capture this aspect, we take *word frequency* as a proxy for the number of times a child is exposed to each word. This is commonly done in reading studies as word frequency is believed to be the current best estimate of a child's exposure to words (Monster et al., 2022). In terms of the model, the consequence is that gender cues (and exceptions) that appear in high-frequency words should influence the calculation of transparency more (and should thus have larger weightings) compared with cues (and exceptions) found in low frequency words. To implement this, the calculation of all concepts in the accompanying spreadsheet that would likely be affected by the repeated exposure to words has been weighted by Zipf word frequency (van Heuven et al., 2014). However, to avoid needless repetition, and for the sake of simplicity, in the description of explicit calculations below, we omit any reference to word frequency. Anyway, the interested reader can verify exactly how frequency has been implemented by examining the formulas in the spreadsheet.

3.2.2. Consistency, perspicuity, syncretism and mono-functionality

Although frequency is quite easy to describe and implement in the model, the remaining theoretical concepts are often interrelated, their presence in the model is less obvious, and/or it is not possible to explicitly separate them in the calculations. For instance, simplicity can refer to one cue per gender as well as to the complexity of the cues (for example, single-letter endings versus multi-letter suffixes). In this first case, within the model, "many" rows in tables which represent cues for a specific gender would represent a more complex case than a table containing few rows. Again, for the sake of brevity, these implicit aspects are not described. Furthermore, perspicuity, syncretism, and simplicity will generally emerge from the data-driven design, rather than needing explicit a priori design decisions. To illustrate, if, for example, three different orthographical forms map to the same phonological ending, the perspicuity, that is, the clarity and simplicity of the phonological cue will be negatively impacted by this mismatch. What starts as a simple one-to-one relationship (the single phoneme related to a specific gender) when children are pre-readers will evolve into a more complex relationship once the child learns to read and discovers three written representations related to the one phoneme. As will be explained below, this added complexity will be captured naturally by the model because, although the phonological-gender transparency will be high in this case due to the one-to-one relationship, it will be combined with the corresponding one-to-many orthographic-gender relationship, and thus the overall transparency will be reduced. Hence, in many places in the following description, we only talk about consistency, but the other concepts are included in a less direct way within the model.

As stated in the first half of the present work, the most transparent cues for gender are also the most consistent. Therefore, in an ideal transparent system, a single orthographic cue and a single phonological cue would exist to consistently represent each particular gender value. That is, within each modality, the relationship between cue and gender would be strictly one-to-one. Hence, for the calculation of consistency,

the model essentially assesses how well the cues which exist in the language measure up to this ideal, with any departures from a one-to-one mapping reducing the consistency value. Fig. 2 (below) illustrates the steps required to evaluate consistency, but, in essence, they assess the following four associations:

i. Single-letter and single-phoneme word endings with gender, separately for nouns, adjectives, and determiners/pronouns.
ii. Multi-letter gender cues (derivational or not).
iii. The consistency between syntactic categories.
iv. The consistency between single- and multi-letter cues.

From this point, we present a more detailed, technical discussion of the model.[3]

3.3. The relationship between orthography and gender

As children learn to speak before they learn to read, the gender cues present in spoken language would influence children before the gender cues present in written language. Hence, it would seem natural to start with a description of how phonology is processed in our model. However, to instantiate the phonological-related aspects of our proposal, we would need access to the full phonological representations for each word included in the model. Unfortunately, the P-PAL database does not contain any phonological information in a coded (useable) form. However, Sá-Leite (2021) has produced a list of word-final phonemes for each Portuguese nominal ending in the P-PAL database, and we were able to use this data. Hence, while it was possible to include multi-lettered orthographic cues in the model, it was only possible to include single-phoneme cues, and only for nouns, as they were the only phonological cues available. Consequently, it was not possible to instantiate some of the theoretical steps encompassing the phonology-gender relationship due to this lack of data. Nevertheless, the theoretical steps required to incorporate phonology are, in general, identical to the theoretical steps that need to be considered for orthography. For this reason, we have chosen to begin by describing the relationship between orthography and gender as we believe that the reader will gain a better appreciation of the model by first describing this more complete relationship, rather than start by describing the partially implemented phonology-gender relationship.

3.3.1. The orthography-to-gender component

The first stage in determining the relationship between orthography and gender requires calculating the transparency value going *from* orthography *to* gender. Essentially, for each orthographic cue present in the language, a gender preference is determined, along with the strength of this preference. How these preferences are determined, along with how they are combined is described below. There are eight specific steps in this stage, and these are summarised in Fig. 2.

3.3.1.1. Single-letter word endings and gender. This first set of calculations is implemented for nouns, adjectives, and determiners/pronouns and is represented by the four boxes in the upper left corner of Fig. 2. The calculations are essentially identical for each syntactic class. The specific examples in the following discussion are all for nouns.[4] The following steps are performed:

i. Column E: The proportion of nouns that end in each of the 33 letters is determined. For example, 32.91 % of nouns end in "-a" while just 5.24 % of nouns end in "-r".
ii. Columns I & K: The "preferred gender" for each letter is determined, along with the strength of this association. For example, 89.47 % of nouns that end in "-a" are feminine while 95.36 % of nouns that end in "-r" are masculine. Thus, both of these orthographic endings have a clear, but not perfect, gender preference when found in nouns.
iii. Column J: The previous two factors are combined by simple multiplication. This is done to take into account both the strength of the association with the preferred gender, together with the frequency of the final letter in the corpus. As seen above, although the strength of the gender preference is similar for "-a" and "-r", given that "-a" is more frequent, its weighting is higher than that for "-r" (0.2945 vs 0.0499, respectively). Thus, the ending "-a" will contribute more to the transparency calculation than will "-r".
iv. Cell J37: The weights for all letters are summed to provide a final value for single-letter suffix consistency for nouns. This total is shown as Step 1 on the summary sheet and has a value of 78.95 %. That is, approximately 4 out of every 5 single-letter noun endings accurately predict gender.

The four steps just described are repeated for adjectives and determiners/pronouns, with the resultant transparency scores shown on the summary sheet as Steps 2 and 3. However, these three values represent the consistency scores within three different syntactic classes. As described in the introduction, consistency of cues is an important factor, with a lack of consistency negatively impacting transparency. Accordingly, a procedure is now required to both compare gender preferences between syntactic classes as well as combine them into a global consistency score for single-letter word suffixes.

There are essentially two tasks here, depending on whether the gender preference for nouns for each of the final letters agrees with its gender preference for the other syntactic categories. The calculations are essentially identical for each syntactic class. The specific examples in the following discussion are all for adjectives.[5]

a) For each final letter, if the gender preference for nouns agrees with the gender preference for adjectives, the strength of the noun preference is increased. This strengthening needs to take into account:
   i. The weighting of the gender preference for adjectives.
   ii. The weighting of the gender preference for nouns.
   iii. The proportion of the entire lexicon which is composed of adjectives, compared to nouns.

   To illustrate these three principles, the endings "-e" and "-l" have similar weightings for their gender preferences within adjectives (0.0691 vs 0.0932, respectively; cells E10 and E19). However, the noun weighting for "-e" is approximately 4 times stronger than the noun weighting for "-l" (0.0727 vs 0.0187, respectively; cells B10 and B19). Consequently, the adjustment made for "-e" needs to be larger than the adjustment made for "-l". Finally, the strengths of these adjustments need to take into account that there are far fewer adjectives in the lexicon, compared to nouns; just 24.8 % (cell H1). Thus, the influence of adjectives should be weaker than that of nouns. The final weighting for each letter is calculated in column F.
b) If, on the other hand, the gender preference for nouns differs from the gender preference for adjectives (for example, as is the case for the final letter "-m"; cell C20 vs cell G20), the strength of the noun preference should decrease. In these cases the strength of the noun

[3] Given the inherent complexity of some of the calculations, having the spreadsheet open while reading the description will greatly increase understanding of the model.

[4] The specific calculations can be found on sheet "1 – Noun Stats", in cell range A1:J37, sheet "2 – Adjective Stats", in cell range A1:J24, and sheet "3 – Det Stats", in cell range A1:13.

[5] The specific calculations can be found on sheet "4 - Adjustment1 Adj-Det-Pron".

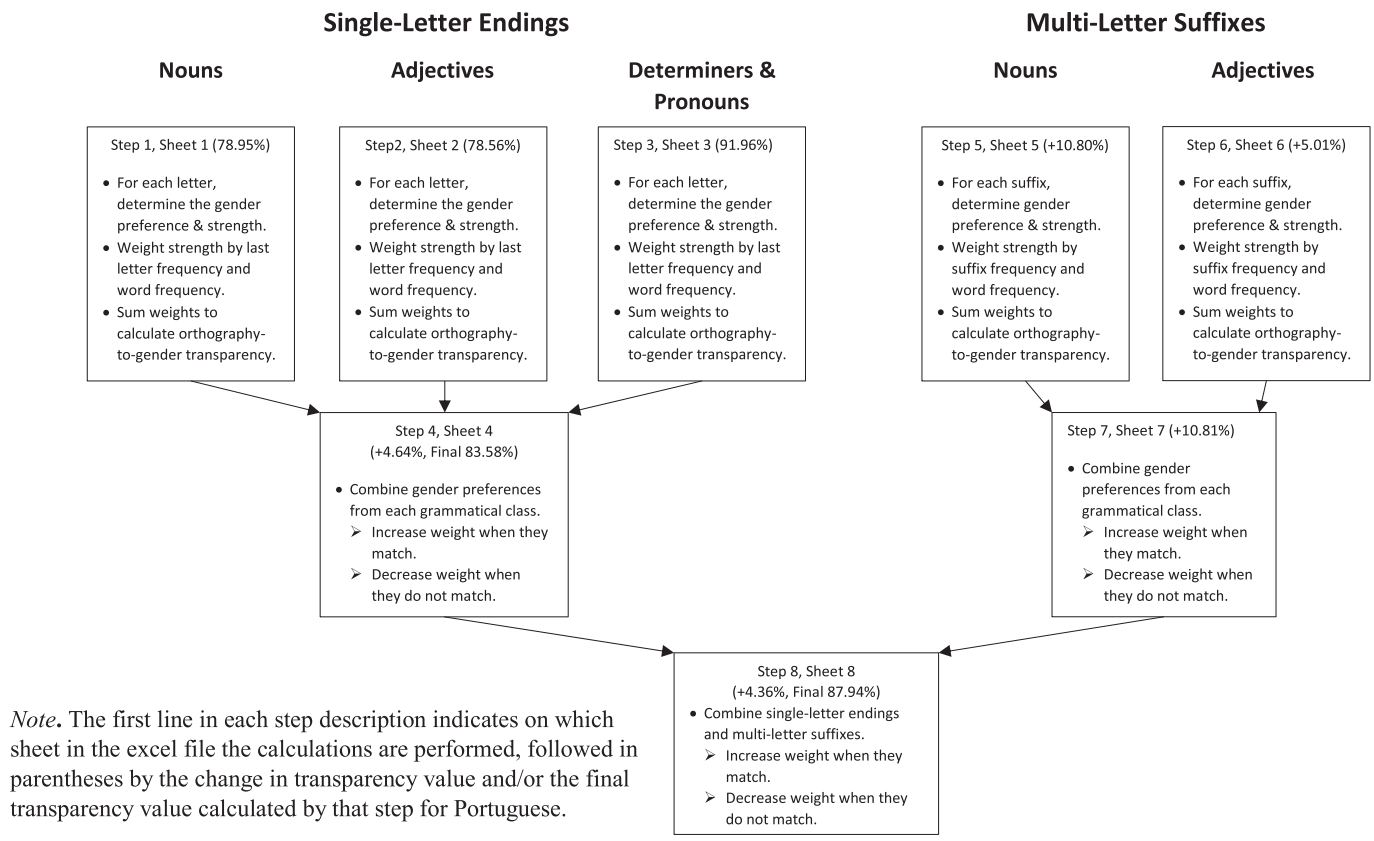


Fig. 2. Required steps to calculate the orthography-to-gender transparency value.

preference is reduced by the strength of the adjective preference, and this magnitude is calculated in the same way as the previous point (shown in column K). We note here that none of the final letters which have a conflict of gender preference between nouns and adjectives are very frequent, and consequently all of the resultant reductions in strength are quite weak.

The steps just described are repeated for determiners/pronouns, and these are combined with the adjustments calculated for adjectives, with the result shown in column S. The sum of this column represents the adjustment to final letter orthography-to-gender transparency, taking into account syntactic classes other than nouns. This value of +4.64 % appears in cell S36, as well as Step 4a on the summary sheet. This small positive value essentially means that, more often than not, the gender preferences for the final letters did not differ across syntactic categories, and so the transparency score increases. The final step is to simply apply this adjustment to the noun orthography-to-gender transparency value that was calculated in Step 1. This is step 4b on the summary sheet, and the value of 83.58 % represents the global orthography-to-gender consistency, taking into account the final letter of each word, as well as whether the gender preferences for adjectives, determiners, and pronouns agree with the preferences of nouns.

3.3.1.2. Multi-letter word endings and gender. Up until this point, the consistency calculations have only considered the final letter of words. However, as noted in the introduction when describing derivational suffixes, for many languages, including Portuguese, there are specific multi-letter endings that are strongly associated with a particular gender. For example, the four-letter ending "-ción" in Spanish is strongly associated with the feminine gender, as is "-heit" in German. Yet, while multi-letter endings are in some sense more complex than single-letter endings, if they represent a reliable predictor of gender they must be taken into account. In this section, we describe how this was achieved in our model.

While all single-letter endings were considered in the previous section (33 in total when the presence of a diacritic is considered), this is not practical for multi-letter endings. Within our database, just considering nouns, there are 336 unique 2-letter endings, 1.539 unique 3-letter endings, 4.907 unique 4-letter endings, and over 10.000 unique 5-letter endings. Furthermore, most of these combinations are very infrequent, and would not alter the final transparency calculation in any meaningful way. Thus, we needed to find a way to determine which multi-letter endings are relevant in terms of predicting word gender. Fortunately, this work has already been done for us. Based on Rio-Torto et al.'s (2013) analysis of noun formation through suffixation, Sá-Leite (2021) has collected all suffixes of European Portuguese and identified that all but two ("-nte" and "-vel") had a clear statistical gender preference. We used this list as the basis for assessing the influence of multi-letter gender cues in nouns. Similarly, a list of suffixes was created for adjectives (based on Rio-Torto et al.'s analysis). Although the specific multi-letter cues differ between syntactic classes, the calculations are identical for each and are represented by the three boxes in the upper right corner of Fig. 2. The following steps are performed, separately for suffixes of different lengths (i.e., for all 2-letter suffixes, for all 3-letter suffixes, etc.), although in the following discussion, all examples are taken from 2-letter noun suffixes.[6] We note at this point that steps i, ii, iii, and v are logically equivalent to the steps previously described for single-letter suffixes:

i. Column E: The proportion of nouns that end in each of the relevant 2-letter suffixes is determined. For example, 11.77 % of nouns end in "-ão".

[6] The specific calculations can be found on sheets "5 - Multi Letter Noun Endings" and "6 - Multi Letter Adj Endings".

ii. Columns I & L: The "preferred gender" for each 2-letter suffix is determined, along with the strength of this association. For example, 81.07 % of nouns that end in "-ão" are feminine.
iii. Column J: These two factors are combined by simple multiplication to create a weighted value for the 2-letter suffix gender preference. As per single-letter endings, this is done to take into account both the strength of the association with the preferred gender, together with the frequency of the multi-letter suffix in the corpus.
iv. Column K: As the multi-letter word endings are contributing to the interim transparency value already calculated for single-letter word endings, the multi-letter weights must be adjusted to ensure that scores greater than 100 % transparency are not possible. Hence, the weights are adjusted to consider the contribution to transparency already made by single-letter endings.
v. These weights are summed to provide a final value for 2-letter suffix consistency, giving a value of +3.95 %. Importantly, as only a small number of 2-letter suffixes were included (9 out of a possible 336 combinations), this value cannot be considered to be an absolute value for 2-letter suffix consistency. Instead, we conceptualise it as the increase in the consistency value calculated for single-letter endings, taking into account these important two-letter suffixes.

The final step is to simply add up the adjustments calculated for each suffix length, and this value of +10.80 % appears in cell D96, as well as Step 5 on the summary sheet. For the moment, this value can be thought of as the increase in the consistency value calculated for single-letter noun endings, considering all important multi-letter suffixes.

Once the multi-letter adjustments have been calculated for all available syntactic classes, the consistency between syntactic classes was compared. Penalties were applied when disagreements were found.[7] The final value of +10.81 % for multi-letter adjustment appears as Step 7 on the summary sheet. Again, the interim interpretation of this value is the increase in the consistency taking into account the gender preferences of the multi-letter suffixed, across all syntactic classes.

However, one last factor needs to be included before the above interim value can be finalised. All multi-letter suffixes contain final-letter endings. For example, the 5-letter suffix "-idade" contains the single-letter ending "-e", and each has an associated gender preference. Although in most instances the two gender preferences agree, this is not always the case. Examining the values calculated by the model reveals that the 3-letter noun suffix "-ção" is very strongly associated with the feminine gender; 96.48 % of nouns that end in this suffix are feminine. Additionally, nouns possessing this suffix represent 8.60 % of all nouns, so this particular suffix is quite influential. Nevertheless, as has been stated in the introduction, in Portuguese the single-letter ending "-o" is very strongly associated with the masculine gender; over 40 % of nouns end with "-o", and of these, 75.48 % are masculine. Hence, words ending in "-ção" contain two prominent, influential, but contradicting gender cues. Although it would be extremely interesting to follow children in a longitudinal study in order to determine how they learn about, and subsequently process these 2 cues, we are not aware of any study that has directly addressed this issue. Regardless, such conflict reduces consistency, and ultimately transparency, and the model must also account for this. This process is represented by the lowermost box in Fig. 2. As per previous evaluations of consistency, the model essentially compares the two gender preferences and penalises any instance where a conflict is found, with the size of the penalty being equal to the weight of the multi-letter suffix.[8] In practical terms, the contribution of the single-letter cue is reduced or increased by the strength of the multi-letter cue, depending on whether the two cues are in conflict or agreement. The result of these calculations, +4.36 %, can be conceptualised as the increase in consistency between word ending and gender once multi-letter suffixes are considered. Hence, this value can be added to the single-letter consistency score to derive a final, global, orthography-to-gender consistency. This value is shown as step 8b on the summary sheet, and has a value of 87.94 %. One way to conceptualise this is that, if one ignores phonology and just considers word final letters, along with multi-letter suffixes, almost 9 out of every 10 words contain valid cues to their gender.

3.3.2. *Bi-directional relationships*

Thus far we have considered the relationship between orthography and gender by only considering one side of the relationship. Essentially, we have taken multiple aspects of orthography and subjected them to various calculations in order to determine their preferred gender and the strength of these relationships. However, as stated at the outset, in a perfectly consistent language, each cue should be associated with one, and only one gender, and each gender should be predicted by one, and only one cue. Thus, we must now take the relevant calculations described up to this point and contemplate them in the reverse direction, that is, going *from* gender *to* orthography. However, the task is somewhat simpler in this direction. For example, given that the important multi-letter suffixes identified in previous steps represent a relatively small portion of the lexicon, when contemplating the gender to orthography relationships, only single-letter endings need to be considered, at least in Portuguese. The individual steps required to calculate the gender to orthography relationship are summarised in the right half of Table 3.

To better understand this process, Fig. 3 represents a reduced version of the Orthographic Noun table from the spreadsheet. The full table contains 33 rows, one for each letter that can appear in the final position of Portuguese nouns whereas Fig. 3 contains just 10 rows (5 letters with feminine preference and 5 letters with masculine preference). From Table 3, step 1 for the *gender to orthography* calculation required identifying the strongest relationship for each gender. From Fig. 3, columns J, K and L, we can see that the strongest relationship for the feminine gender is the value 0.2945, for the ending "-a", and that the strongest relationship for the masculine gender is 0.3066, for the ending "-o".

The second step in Table 3 mandates that these two values (now shown in columns M and N) are combined in some way. As was done when contemplating the relationship in the opposite (orthography-to-gender) direction, we simply sum these weighted values to give a total of 60.10 % (and this is step 9 on the summary sheet). Step 3 from Table 3 simply specifies that these calculations are repeated for other word classes, that is, in the present model, for adjectives, and determiners/pronouns (Steps 10 and 11 on the summary sheet). Hence, at this point we have a total of three consistency values for the *gender to orthography* relationship. Finally, step 4 from Table 3 specifies that these values need to be combined into a single value, and this is done by simply taking the weighted average (step 12 on the summary sheet). The averages are weighted by the proportions of each syntactic class in the lexicon. The reason behind this is to recognize that nouns, for example, are the most frequent syntactic class and should therefore exert a stronger influence on the calculation.

At this point we have two values for the relationship between orthography and gender, the value we just calculated which goes from gender to orthography (step 12) and a previous value which goes from orthography to gender (step 8b). In step 13 we simply take the average of these two values to derive a final value for consistency for the orthography-gender relationship, namely 73.79 %. Thus, taking into account all relevant aspects of the bi-directional relationship between orthography and gender, the model suggests that approximately 3 in every 4 Portuguese words have transparent orthographic cues that correctly predict gender.

[7] The specific calculations can be found on sheets sheet "7 - Adjustment2 Multi".
[8] The specific calculations can be found on sheet "8 - Adjustment3 Single vs Multi".

Table 3
Comparison of steps for orthography-to-gender and gender-to-orthography calculations.

Step	Orthography to gender	Gender to orthography
1	For each of the 93 relevant orthographic endings for nouns (33 single-letter endings, 9 two-letter suffixes, etc.), determine which of the possible genders (in this case, just masculine or feminine) was preferred.	For each of the 2 genders present in Portuguese, determine which single-letter noun ending it is most strongly associated with.
2	Taking into account conflicts and other factors, combine all of these values to determine a final consistency value.	Combine these 2 values in some way to determine a final consistency value.
3	Repeat steps 1 & 2 for words from other syntactic classes.	Repeat steps 1 & 2 for words from other syntactic classes.
4	Combine values from all syntactic classes.	Combine values from all syntactic classes.

	A	F	G	H	I	J	K	L	M	N
1		Orthography								
2	Final	Token Proportions					Reverse Consistency			
3	Letter	Masculine	Feminine	Total	Max Gender Proportion	Weighted Max Gender Prop	Gender Pref	Rank	Masculine	Feminine
4	a	0.1053	0.8947	1.0000	0.8947	0.2945	Feminine	1	-	0.2945
9	c	0.8530	0.1470	1.0000	0.8530	0.0001	Masculine	23	-	-
11	e	0.4370	0.5630	1.0000	0.5630	0.0727	Feminine	2	-	-
20	l	0.8021	0.1979	1.0000	0.8021	0.0187	Masculine	3	-	-
21	m	0.3566	0.6434	1.0000	0.6434	0.0107	Feminine	3	-	-
23	o	0.7548	0.2452	1.0000	0.7548	0.3066	Masculine	1	0.3066	-
27	r	0.9536	0.0464	1.0000	0.9536	0.0499	Masculine	2	-	-
28	s	0.8968	0.1032	1.0000	0.8968	0.0043	Masculine	4	-	-
32	v	0.4524	0.5476	1.0000	0.5476	0.0001	Feminine	7	-	-
36	z	0.4164	0.5836	1.0000	0.5836	0.0036	Feminine	4	-	-
37	Total					0.7895			0.3066	0.2945

Fig. 3. Reduced version of the noun orthography table taken from sheet "1 – Noun Stats".
Note. The purpose Fig. 3 is to clarify how the Bi-Directional Relationships are calculated. Accordingly, only 10 of the 33 rows, and 10 of the 14 columns from the full table are shown here for simplicity.

3.4. The relationship between phonology and gender

Having calculated a gender transparency value considering just orthography, we now turn to the contribution made by the phonology-gender relationship. Theoretically, the steps needed for phonology are identical to those needed for orthography, with phonemes substituted for letters. Additionally, given that our phonology data consists of just the final phoneme for each noun, rather than multi-phoneme endings for all syntactic classes, many of the steps described in the orthography section cannot be implemented in the present version of the model due to lack of data. For these two reasons, the subsequent description of the phonology-gender relationship will be necessarily brief compared with the description for orthography. However, we stress that this is not meant to downplay the importance of the phonology-gender relationship. Following the restrictions that result from having limited phonological data, Fig. 4 summarises the steps that have been implemented in the model.

In the first step, the equivalent process for orthography was to analyse the gender preference for the 33 possible letters that can appear at the end of a written Portuguese noun. Hence, for phonology, the model analyses the 35 possible phonemes that form the ends of spoken Portuguese nouns.[9] As per the orthography calculation, for each phoneme, the strength of the gender preference is weighted by the frequency of the phonemes in the final position of words, along with the frequencies of the words they appear in. For example, 32.91 % of nouns end in the phoneme /ɐ/, and of those, 89.47 % are feminine. Accordingly, this phoneme would influence the transparency calculation far more than, for example, the phoneme /g/ which is the final phoneme in just 0.21 % of nouns. The weighted transparency scores are summed to provide an overall phoneme-to-gender score in step 14a of 86.49 %.

Using a similar method for the gender-to-phoneme relationship the preferred phoneme and the strength of the relationship are determined for the two genders (/ɐ/ for feminine, /u/ for masculine), and their weighted values are combined in step 14b to provide an overall transparency value of 58.00 %.

Finally, the phoneme-to-gender and gender-to-phoneme values are combined in step 15 to determine the overall transparency for the phonology-gender relationship, which has a value of 72.24 %. Bearing in mind that we only have phonological information for nouns, this value indicates that for slightly less than three-quarters of Portuguese nouns, the final phoneme correctly indicates the gender.

3.5. Overall language transparency

Having derived a transparency value for the orthography-gender relationship and the phonology-gender relationship, the final step is to combine these into a global transparency value for the entire language. This is done in Step 16 by taking an average of these two numbers to give a value of 73.02 %. This value means that Portuguese is not completely transparent, but it infers that Portuguese does generally contain clear gender cues with only a relatively small number of ambiguous or irregular combinations. As would be expected, it falls on the transparent side of the supposed Transparency Continuum proposed by Kupisch et al. (2018), analogously to Spanish and Italian, which are similar Romance languages that have been qualitatively assessed as transparent in this proposal.

At this point, we note that step 16 is only relevant for skilled adult, literate language users, as is based on both orthographic and phonological information extracted from a database of adult word frequencies. It would not be relevant to calculate this overall value for individuals

[9] The specific calculations can be found on sheet "1 - Noun Stats", range A41:J79 and range K41:N79.

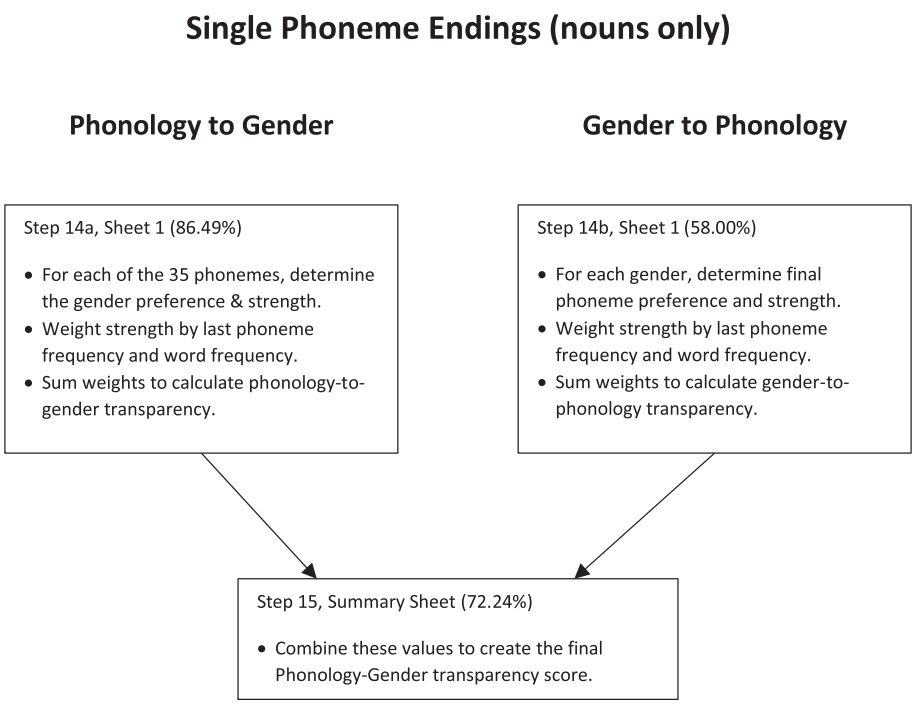


Fig. 4. Summary of steps for the phonology-gender calculations.

who were pre-readers, or who were illiterate, as these individuals would only have been exposed to phonological gender cues. However, given the modular, step-wise nature of the calculations in the model, it is nevertheless possible to extract a more relevant transparency value for non-typical examples. In the case of individuals who had not learned to read, the relevant gender transparency score would be calculated in step 15 as it is based purely on the relationship between phonology and gender.

3.6. An example using Dutch

In terms of gender, Portuguese is considered to be largely transparent. To provide additional evidence of the plausibility of the model, we decided to apply it to a language which is generally considered to be opaque. Dutch has a bipartite gender system with two gender values: common and neuter. The common gender value emerged in modern Dutch when the feminine and masculine gender values collapsed into one (Postma & van Reenen, 2009). Nowadays, masculine and feminine are almost only used for natural gender, similarly to English (pronouns; nouns referring to men/women; Audring, 2006; Hinskens et al., 2020). Relevant for our purposes, gender in Dutch is usually regarded as opaque (Unsworth, 2008). Usually, the only consistent regularity identified is that present for diminutive nouns, which are formed with the suffix "-chen" and are always neuter, regardless of the gender (common/neuter) of the noun they derive from. For this reason, instantiating the model in Dutch will provide a useful contrast to Portuguese.

To construct the model for Dutch, we used the CELEX Database (Baayen et al., 1995). Compared to the Portuguese database, CELEX contains limited data for our purposes; although this database contains a lot of information for many different syntactic classes, gender information is only available for nouns. Furthermore, we have no information regarding which multi-letter suffixes provide reliable gender cues in Dutch. However, given the modular nature of the model, this does not present an insurmountable problem. It is possible to perform a subset of calculations in order to provide a qualified value for language transparency whose scope is limited to the type of data that was present in the language database, rather than applying to the entire language. Hence, any conclusions drawn from our calculations for Dutch should be limited to just nouns. Table 4 summarises the values for transparency for Dutch, along with the equivalent values for Portuguese.

The overall transparency score for Dutch was 43.39 %. Given that this value was based on far less information than was used for Portuguese, rather than comparing the Dutch transparency score to the full, complete transparency score for Portuguese (top half of Table 4), we feel it would be more meaningful to make a like-for-like comparison. Hence, the bottom half of Table 4 includes a reduced set of Portuguese data, along with new adjusted totals based on this reduced data. Considering firstly orthography, going from orthography to gender we have 78.95 % for Portuguese versus 73.24 % for Dutch. This value for Dutch is unexpectedly high, given that Dutch is generally considered opaque for gender, and suggests that in approximately three-quarters of Dutch nouns, the final letter in correctly predicts gender. However, in the opposite direction, going from gender-to-orthography, there is a large dissociation between the two languages, with scores of 60.10 % for Portuguese versus 12.28 % for Dutch. This difference can be explained as follows. In Portuguese, a feminine noun is often spelled with a final "a" and a masculine noun is often spelled with a final "o". Hence, there are clear preferences, and consequently, the exceptions that do exist are relatively infrequent. In contrast, in Dutch, there is no clear final letter preference for either gender, with the strongest relationship for the common gender – just over 12 % of nouns with the common gender are spelled with a final "e". Hence, each gender is weakly associated with many letters, and these effectively act as exceptions, with the consequence that transparency is inevitably reduced.

We believe it would also be justified, and useful, to calculate averages; that is, overall orthography-gender values for nouns, thus giving 69.52 % for Portuguese and 42.76 % for Dutch. Therefore, if one considers the bidirectional nature of the orthographic-gender relationship, over two-thirds of Portuguese nouns contain final letters that correctly indicate gender, whereas this is the case in just over 40 % of Dutch words. Of interest, we also note that the value of 69.52 % for Portuguese is less in this reduced model compared with the corresponding value of 73.79 % for the full model. This reflects the fact that the transparency score for Portuguese in the full model benefits from a high degree of agreement in gender preference across syntactic classes as well as support from multi-letter suffixes.

Table 4
A comparison of transparency for Portuguese and Dutch using the full model (above) and results based just on nouns.

Model step	Portuguese		Dutch	
Comparison using full version of the model				
Orthography				
Orthography to Gender				
Step 1. Noun (Last letter)		78.95		73.24
Step 4a. Adjust. for Adj/Det/Pron (Last letter)	+4.64	83.58		
Step 8a. Final global multi-letter suffix adjust.	+4.36	87.94		73.24
Gender to Orthography				
Step 9. Noun (Last letter)	60.10		12.28	
Step 10. Adjective (Last letter)	57.58			
Step 11. Det/Pron (Last letter)	67.39			
Step 12. Final global (Last letter)		59.65		12.28
Step 13. Final global Orthography-Gender		73.79		42.76
Phonology				
Step 14a. Noun Phon-to-Gender (Last phoneme)	86.49		73.04	
Step 14b. Noun Gender-to-Phon (Last phoneme)	58.00		14.99	
Step *15. Final global Phonology-Gender*		72.74		44.01
Combined Transparency (Step 16)		73.02		43.39
Comparison using just nouns				
Orthography				
Step 1. Orthography-to-Gender (Last letter)	78.95		73.24	
Step 9. Gender-to-Orthography (Last letter)	60.10		12.28	
Final global Orthography-Gender Relationship		69.52		42.76
Phonology				
Step 14a. Phonology-to-Gender (Last phoneme)	86.49		73.04	
Step 14b. Gender-to-Phonology (Last phoneme)	58.00		14.99	
Final global Phonology-Gender Relationship		72.74		44.01
Combined Transparency		71.13		43.39

Note. Adj = Adjective; Adjust. = Adjustment; Det = Determiner; Phon = Phonology; Pron = Pronoun.

A similar pattern for Dutch was seen when analysing the phonology-gender relationship. When going from final phoneme to gender, the transparency score for Dutch nouns was reasonably high, being 73.04 %. However, going from gender to final phoneme, transparency is just 14.99 %. Again, this reflects the fact that, in general, many final phonemes appear to have a clear gender preference. In fact, examining the phonology table in the Dutch model reveals that two-thirds of phonemes have a preference for one of the two genders which exceeds 65 %. However, as occurred with letters, there is no clear phoneme preference for either gender, and, instead, the many weak relationships effectively act as exceptions, thus lowering the transparency score.

Finally, we calculate overall transparency scores for nouns in the two languages by averaging the orthography and phonology scores. The final values are 72.24 % for Portuguese and 44.01 % for Dutch. Indeed, Dutch is systematically regarded as an opaque language when it comes to gender (Unsworth, 2008), and qualitatively classified in the opaque side of transparency by Kupisch et al. (2018), and the values produced by the model support this classification. Interestingly, our model suggests that Dutch orthography and phonology may generally contain clear cues for gender, and that the apparent opaqueness of Dutch may largely be driven by the fact that each gender is associated with many final letters and many final phonemes.

4. Closing remarks

4.1. Limitations

In our opening comments, we stated that our goal was to let the data speak and create a model that was as free from a priori decisions as possible. However, this was only partially achieved. On the one hand, we were able to take advantage of the lexical information available in the Portuguese and Dutch databases. Thus, the effect of consistent gender cues, as well as the exceptions, on transparency were calculated from the data, rather than imposed on the model, with word frequency used as a proxy for word exposure. On the other hand, we are not aware of behavioural data which could be used to guide all of the design decisions we need to take. A perfect example of this is the final step in the calculation (step 16); the decision to derive the overall language transparency score by calculating the average of the orthography-gender relationship and the phonology-gender relationship. In effect, this is an a priori decision which does have consequences. By taking the average we are giving equal importance to orthographic and phonological gender cues. However, it may be the case that phonological cues are more salient given that, as previously noted, children generally learn to speak before they learn to read. Nevertheless, this is an open question that requires empirical data to be answered. Additionally, even if we had assumed that phonological cues should contribute more to the overall transparency of the language, this would have obliged us to include in the model a specific weighting in favour of phonology, yet this would be completely unmotivated due to a lack of behavioural data. Hence, we believe that calculating the average of the orthography-gender and the phonology-gender relationships (and hence assuming equal importance) is the lesser of two evils. Future research that collects relevant empirical evidence will inform us on how to improve the model's design.

Another potential limitation of the present work is that the word frequencies provided in the P-PAL database can be considered, for practical purposes, to be “adult exposure” word frequencies. However, the use of “adult frequencies” in and of itself does not invalidate the model. Instead, it just means that the final transparency value derived will be slightly different from the value that would be obtained if “toddler frequencies” were available. Furthermore, we hypothesise that there would be little difference in the final transparency value obtained if one substituted “adult frequencies” with “toddler frequencies”. The reason for this is that children tend to learn mostly high frequency words and the increase in lexicon size as one becomes an adult is largely made up of lower frequency words (Bonin et al., 2004; Zevin & Seidenberg, 2002). Thus, the lexical cues that children are exposed to would tend to influence the calculation of transparency to a greater extent than words which are only known to adults. Nevertheless, this hypothesis could be

tested if a database of child word frequencies in the Portuguese language were to become available.[10]

In this sense, we note that to fully test this model, new databases must be created that contain all of the information required to instantiate all of the concepts included here. To be fair, databases exist for many languages, but often some critical piece of information is lacking (quite often, the gender of each word, e.g., SUBTLEX-DE, Brysbaert et al., 2011; SUBTLEX-ES, Cuetos et al., 2011; EsPal, Duchon et al., 2013; dlexDB, Heister et al., 2011). Indeed, in the present work, word phonology in Portuguese was only available for nouns, and even then, only after combining two sources of lexical information (Sá-Leite, 2021; Soares et al., 2014). This means that we were unable to determine whether multi-phoneme endings always agree in their gender preference with that of the final phoneme, in which case transparency would increase, or whether there are many conflicts, such as the "-ção" $_{\text{Feminine}}$ vs "-o" $_{\text{Masculine}}$ case we saw in orthography, in which case, transparency would decrease.

Additionally, gender information was not available in Dutch for syntactic classes other than nouns. However, given that the overall transparency calculation requires many steps, it is possible to calculate transparency for specific aspects of a language, and this may still prove useful. As seen with the example using Dutch, although the database contained far less information than the Portuguese database, we were still able to make a direct, albeit limited comparison between the two languages; that is, solely based on the gender transparency calculations performed for nouns, and recognising that inconsistencies between grammatical categories were not taken into account. Depending on the specific research goals, this may still prove to be useful.

4.2. Possible enhancements to the model

The current version of the model includes word frequency but does not include other key variables known to influence lexical processing, such as age of acquisition and word length. Future versions of the model could include these features if behavioural data supported their inclusion. In any case, the decision was taken to not include many variables known to influence lexical processing for the following reasons. The goal of the model was to calculate an overall gender transparency value for an entire language. Thus, it is important to bear in mind that in the present work, we are concerned with how gender cues contained in the entire lexicon influence the learning of gender, rather than describing how gender is retrieved when individual words are processed. Consequently, we have not included variables known to be important in lexical processing, such as age of acquisition, as we do not believe them to be especially relevant to the present work.

Given that the two languages used to demonstrate the model have two genders, it is relevant to ask if the model could be applied to languages with more than two genders, for example, Swahili and other Bantu languages. The short answer is "yes". The model was designed to be scalable. In fact, we piloted the model on a small, invented language with three genders. The purpose was not to determine what typical transparency values might be for tri-gendered languages. Instead, it was merely to ensure that the formulas and steps proposed would still function adequately in a language with more than two genders. Thus, there is nothing inherent in the general structure of the model that would prevent it from being evaluated against a language with many genders.[11]

4.3. Possible uses for the model

Given that the model presented here instantiates the ideas presented in the first half of the present work, it essentially makes predictions, which, we believe, could be tested with the correct experimental design. The results could be used to refine both the model and the concepts upon which it is based. What might such experiments look like? We present two very basic (oversimplified) proposals here but believe that as researchers in the field begin to think of gender transparency as a quantifiable concept, many different experiments will be conceived.

4.3.1. Within language study

Given that the model calculates both the orthography-gender relationship and the phonology-gender relationship, it is possible that, for a particular language, one dimension is highly transparent while the other is not. For example, in a language with relatively few, but unambiguous phonetic cues accompanied by many orthographic cues where many exceptions abound, phonology may come out at 90 % transparent while orthography at only 50 % transparent (resulting in a global transparency value of 70 %). This suggests that the phonetic gender cues of this language would be easier and faster for children to learn, compared to the orthographic gender cues. Further, such a difference should be detectable. In this case, if children were presented with pseudowords which contained gender cues in both an oral and written form and were asked to use them in a sentence, we would expect more agreement between language elements and fewer errors in the children's orally produced sentences (with respect to gender agreement) compared to their written sentences.

Another possibility would be to study whether, within languages, transparency differs greatly across syntactic classes. As the model essentially calculates transparency for different grammatical classes (assuming a relevant database is available), the predictions of the model regarding these differences could be empirically evaluated.

A final possibility would be to determine if transparency varies between genders within the same language. If, for example, the model determines that the orthography-to-phonology transparency is high for gender-A, but low for gender-B, it potentially means that children would have more problems learning gender-B. It also suggests the possibility that educators may need to spend more time teaching children about reliable cues for gender-B compared to gender-A.

4.3.2. Cross-language study

Given that the model breaks down the concept of transparency into many steps, it may be the case that language A has a relatively low transparency score compared to language B, but that a specific aspect of language A was nevertheless more transparent than the corresponding aspect in Language B; for example, the transparency of multi-letter cues. Again, this observation would suggest experiments which could be carried out to determine if this predicted difference is observed between the speakers of each language.

4.3.3. Non-typical language users

Another possible advantage of the modular nature of the model is that it could make predictions regarding gender acquisition and use for non-typical language users. An obvious example is the fact that the orthography-gender and phonology-gender relationships are calculated separately. Thus, for example, individuals with little or no exposure to written words would rely almost exclusively on the gender cues in

[10] Current databases of language development in European Portuguese (e.g., FrePOP, Frota et al., 2010; PLEX5, Frota et al., 2012; CDI-European Portuguese, Frota et al., 2015; ESCOLEX, Soares et al., 2014) do not allow us to conduct such an analysis as they test older children that have already acquired gender, or do not provide information about gender, and future research might usefully fill this gap.

[11] We have included this tri-gendered version of the model in the supplementary materials (OSF link) as a guide for any researcher who might like to apply the model to languages with more than two genders.

spoken words. Hence, the values produced by the model using just phonological information could be of relevance to, for example, child pre-readers and illiterate adults. On the other hand, individuals with profound congenital or early acquired hearing loss will not have had the opportunity to be exposed to spoken language. Although modified phonics programs have been developed for deaf and hard-of-hearing (DHH) individuals (for example, Trezek et al., 2007), not all DHH individuals receive such programs. Furthermore, many DHH individuals become skilled readers despite having remarkably poor phonemic awareness (Miller & Clark, 2011). Consequently, it is plausible that the orthography-gender transparency values calculated by the model would be more relevant than the overall transparency scores for research involving gender processing in DHH individuals.

4.4. Final considerations

The present paper set out to achieve two goals. As noted in the introduction, an unresolved challenge in the study of gender transparency is the inconsistency in its definition and use in classifying different languages. Thus, our first goal was to propose a systematic method of assessing the degree of gender transparency of languages following the principle of complex relativity described by Audring (2019). This proposal is explained in detail in Section 2 of this paper. To our knowledge, this is the first attempt to develop a new way of assessing languages according to their degree of gender transparency. Our second goal was to instantiate the theoretical proposal in a model using two real languages which are generally considered to be very different in terms of gender transparency. The fact that the model produced plausible values for each language (when using only language elements available in both databases, 71 % for Portuguese vs 43 % for Dutch), serves to demonstrate that it is possible to apply the concept of this proposal to real languages and derive quantifiable values for transparency that are not far away from the theoretical assumptions made in the literature.

We believe that the achievement of these two goals represents a crucial step forward in the field. Firstly, we believe that the current proposal will allow the concept of gender transparency to be standardized across studies. Secondly, we hope that the model serves as a "call to arms" for researchers in the field, ultimately encouraging them to think about gender transparency as a quantifiable concept. In the first instance, we encourage our companions in the field to critique the model. Subsequently, we hope that other researchers will help to expand the model. Although we largely used Portuguese to illustrate the proposal, to fully validate the model, and to build a precise continuum of gender transparency, multiple languages should be analysed and compared to each other. The model was built to be both flexible and scalable. Accordingly, we encourage other researchers to embrace this project and take the first steps to the building of such a continuum. Nevertheless, there is currently a lack of suitable databases which can inform research in this area. Thus, another avenue of research that other authors may pursue is the expansion of existing databases, or the creation of new databases, which will allow further research along these lines. Ultimately, our motivation is that the present work inspires researchers who are interested in all aspects of language transparency and hopefully causes research in this area to flourish.

Open practices statement

The data and materials for the proposed procedure are available at https://osf.io/74ezj/?view_only=e493c6cc2b3a4486b48c370d8575f2ed.

PsycINFO codes

2720 Linguistics & Language & Speech
2200 Psychometrics & Statistics & Methodology
2240 Statistics & Mathematics
2340 Cognitive Processes

Funding

This project was funded by Consellería de Cultura, Educación e Ordenación Universitaria, Xunta de Galicia (ED481B-2022-041), by Axencia Galega de Innovación and Consellería de Economía, Industria e Innovación, Xunta de Galicia (ED431B 2019/2020; ED431B 2022/19) and also by Ministerio de Ciencia e Innovación (PID2019-110583GB-I00). It was partially supported by the Foundation for Science and Technology (FCT) through the Portuguese State Budget (Ref.: UIDB/PSI/01662/2020) and grant PD/BD/52396-2013. The participation of Ian Simpson was made possible by a grant from FEDER Junta de Andalucía-Consejería de Transformación Económica, Industria, Conocimiento y Universidades/ project E-SEJ-754-UGR20. The publication of this article was financed by the Goethe Universität Frankfurt through the Postdoc Förderprogramm Fokus A|B.

CRediT authorship contribution statement

Ana Rita Sá-Leite: Writing – review & editing, Writing – original draft, Supervision, Project administration, Investigation, Funding acquisition, Data curation, Conceptualization. **Ian Craig Simpson:** Writing – review & editing, Writing – original draft, Software, Methodology, Funding acquisition, Formal analysis. **Isabel Fraga:** Writing – review & editing, Investigation, Funding acquisition. **Montserrat Comesaña:** Writing – review & editing, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

We have shared an OSF link to our data in the manuscript.

Acknowledgments

We would like to thank two anonymous reviewers for their valuable comments during the review process.

References

Afonso, O., Domínguez, A., Alvarez, C. J., & Morales, D. (2014). Sublexical and lexico-syntactic factors in gender access in Spanish. *Journal of Psycholinguistic Research, 43*(1), 13–25. https://doi.org/10.1007/s10936-012-9236-0

Agirre, A. I. (2016). The processing of gender assignment in Spanish: Does canonicity have an effect on Basque speakers? *Spanish Journal of Applied Linguistics, 29*(2), 523–543. https://doi.org/10.1075/resla.29.2.06ima

Albright, A. (2002). Islands of reliability for regular morphology: Evidence from Italian. *Language, 78*(4), 684–709. https://doi.org/10.1353/lan.2003.0002

Arias-Trejo, N., & Alva, E. A. (2013). Early Spanish grammatical gender bootstrapping: Learning nouns through adjectives. *Developmental Psychology, 49*(7), 1308–1314. https://doi.org/10.1037/a0029621

Arias-Trejo, N., Falcón, A., & Alva, E. A. (2013). The gender puzzle: Toddlers' use of articles to access noun information. *Psicológica, 34*(1), 1–23. https://doi.org/10.17575/PSICOLOGIA.V34I1.1435

Audring, J. (2006). Pronominal gender in spoken Dutch. *Journal of Germanic Linguistics, 18*(2), 85–116. https://doi.org/10.1017/S1470542706000043

Audring, J. (2019). Canonical, complex, complicated? In F. Di Garbo, B. Olsson, & B. Wälchli (Eds.), *Grammatical gender and linguistic complexity: Volume I: General issues and specific studies (studies in diversity linguistics 26)* (pp. 15–52). Language Science Press.

Ayoun, D. (2018). Grammatical gender assignment in French: Dispelling the native speaker myth. *Journal of French Language Studies, 28*(1), 113–148. https://doi.org/10.1017/S095926951700014X

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *CELEX2 LDC96L14. Web download.* Philadelphia: Linguistic Data Consortium.

Bates, E., Devescovi, A., Hernandez, A., & Pizzamiglio, L. (1996). Gender priming in Italian. *Perception & Psychophysics, 58*(7), 992–1004. https://doi.org/10.3758/bf03206827

Bates, E., Devescovi, A., Pizzamiglio, L., D'Amico, S., & Hernandez, A. (1995). Gender and lexical access in Italian. *Perception & Psychophysics, 57*(6), 847–862. https://doi.org/10.3758/bf03206800

Boloh, Y., & Ibernon, L. (2013). Natural gender, phonological cues and the default grammatical gender in French children. *First Language, 33*(5), 449–468. https://doi.org/10.1177/0142723713499848

Bonin, P., Barry, C., Méot, A., & Chalard, M. (2004). The influence of age of acquisition in word reading and other tasks: A never ending story? *Journal of Memory and Language, 50*(4), 456–476. https://doi.org/10.1016/j.jml.2004.02.001

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology, 58*(5), 412–424.

Caffarra, S., Barber, H., Molinaro, N., & Carreiras, M. (2017). When the end matters: Influence of gender cues during agreement computation in bilinguals. *Language, Cognition, and Neuroscience, 32*(9), 1069–1085. https://doi.org/10.1080/23273798.2017.1283426

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

Chini, M. (1995). *Genere grammaticale e acquisizione. Aspetti della morfologia nominale dell'italiano L2*. Milano: Franco Angeli.

Clegg, J. H. (2011). A Frequency-based Analysis of the Norms for Spanish Noun Gender. *Hispania, 94*(2), 303–319. http://www.jstor.org/stable/23032172.

Colé, P., Pynte, J., & Andriamamonjy, P. (2003). Effect of grammatical gender on visual word recognition: Evidence from lexical decision and eye movement experiments. *Perception & Psychophysics, 65*, 407–419. https://doi.org/10.3758/BF03194572

Corbett, G. G. (1991). *Gender*. Cambridge University Press. https://doi.org/10.1017/CBO9781139166119

Cuetos, F., Glez-Nosti, M., Barbón, A., & Brysbaert, M. (2011). SUBTLEX-ESP: Spanish word frequencies based on film subtitles. *Psicológica, 32*(2), 133–143.

Culbertson, J., Gagliardi, A., & Smith, K. (2017). Competition between phonological and semantic cues in noun class learning. *Journal of Memory and Language, 92*, 343–358. https://doi.org/10.1016/j.jml.2016.08.001

De Martino, M., Bracco, G., & Laudanna, A. (2011). The activation of grammatical gender information in processing Italian nouns. *Language & Cognitive Processes, 26*(4–6), 745–776. https://doi.org/10.1080/01690965.2010.491977

De Martino, M., Bracco, G., Postiglione, F., & Laudanna, A. (2017). The influence of grammatical gender and suffix transparency in processing Italian written nouns. *The Mental Lexicon, 12*(1), 107–128. https://doi.org/10.1075/ml.12.1.05dem

Desrochers, A., & Brabant, M. (1995). Interaction entre facteurs phonologiques et sémantiques dans une épreuve de catégorisation lexicale [Interaction between phonological and semantic factors in a lexical categorization test]. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, 49*(2), 240–262. https://doi.org/10.1037/1196-1961.49.2.240

Desrochers, A., & Paivio, A. (1990). Le phonème initial des noms inanimés et son effet sur l'identification du genre grammatical [the initial phoneme of inanimate nouns and its effect on grammatical gender identification.]. *Canadian Journal of Psychology, 44*, 44–57.

Desrochers, A., Paivio, A., & Desrochers, S. (1989). L'effet de la fréquence d'usage des noms inanimés et de la valeur prédictive de leur terminaison sur l'identification du genre grammatical [the effect of frequency of use of inanimate nouns and the predictive value of their termination on the identification of grammatical gender]. *Revue Canadienne de Psychologie, 43*, 62–73.

Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. *Behavior Research Methods, 45*(4), 1246–1258. https://doi.org/10.3758/s13428-013-0326-1

Eichler, N., Jansen, V., & Müller, N. (2013). Gender acquisition in bilingual children: French–German, Italian–German, Spanish–German and Italian–French. *International Journal of Bilingualism, 17*(5), 550–572. https://doi.org/10.1177/1367006911435719

Fernández-Fuertes, R., Álvarez-de-la-Fuente, E., & Mujcinovic, S. (2016). The acquisition of grammatical gender in L1 bilingual Spanish. In A. A. de-la-Fuente, E. Valenzuela, & C. Martínez-Sanz (Eds.), *Language acquisition beyond parameters: Studies in honour of Juana M. Liceras* (pp. 237–279). https://doi.org/10.1075/sibil.51.10fue

Figueira, R. A. (2013). Marcas insólitas na aquisição do gênero gramatical: a propriedade reflexiva da linguagem na fala da criança. *Letras de Hoje, 36*(3), 313–320. Retrieved from https://revistaseletronicas.pucrs.br/ojs/index.php/fale/article/view/14582.

Finley, S. (2022). Learning complex morphological patterns: The role of syncretism and markedness. In A. D. Sims, A. Ussishkin, J. Parker, & S. Wray (Eds.), *Morphological diversity and linguistic cognition* (pp. 203–204). Cambridge University Press. https://doi.org/10.1017/9781108807951.

Frota, S., Butler, J., Correia, S., Severino, C., Vicente, S., & Vigário, M. (2015). *Questionários McArthur-Bates CDI para o Português Europeu: formas reduzidas*. Laboratório de Fonética, CLUL/FLUL. Retrieved from http://labfon.letras.ulisboa.pt/babylab/pt/CDI.

Frota, S., Correia, S., Severino, C., Cruz, M., Vigário, M., & Cortês, S. (2012). *PLEX5 – A production lexicon of child speech for European Portuguese / Um léxico Infantil para o Português Europeu*. Lisbon: Laboratório de Fonética. CLUL/FLUL. ISBN 978-989-95713-8-9. ISLRN 468-779-277-600-8.

Frota, S., Vigário, M., Martins, F., & Cruz, M. (2010). *FrePOP. Version 1.0. Laboratório de Fonética (CLUL)*. Faculdade de Letras da Universidade de Lisboa. Retrieved from http://frepop.fl.ul.pt/.

Garrido-Pozú, J. J. (2022). Predictive processing of grammatical gender: Using gender cues to facilitate processing in Spanish. *Lingua, 278*, Article 103416. https://doi.org/10.1016/j.lingua.2022.103416

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

Gudmundson, A. (2010). *L'acquisizione del genere grammaticale in italiano L2 Quali fattori possono influenzare il grado di accuratezza?*. Doctoral Dissertation. Stockholms Universitet. Obtained from https://www.diva-portal.org/smash/get/diva2:372636/FULLTEXT01.pdf.

Harris, J. W. (1991). The exponence of gender in Spanish. *Linguistic Inquiry, 22*(1), 27–62. https://doi.org/10.1007/978-94-011-2516-1_6

Heidolph, K. E., Flämig, W., & Motsch, W. (1984). *Grundzüge einer deutschen Grammatik [Basics of a German grammar]*. Akademie-Verlag, 2. Auflage.

Heister, J., Würzner, K.-M., Bubenzer, J., Pohl, E., Hanneforth, T., Geyken, A., & Kliegl, R. (2011). dlexDB—Eine lexikalische Datenbank für die psychologische und linguistische Forschung [dlexDB—A lexical database for the psychological and linguistic research]. *Psychologische Rundschau, 62*(1), 10–20.

Hernandez, A. E., Hofmann, J., & Kotz, S. A. (2007). Age of acquisition modulates neural activity for both regular and irregular syntactic functions. *NeuroImage, 36*(3), 912–923. https://doi.org/10.1016/j.neuroimage.2007.02.055

Hernandez, A. E., Kotz, S. A., Hofmann, J., Valentin, V. V., Dapretto, M., & Bookheimer, S. Y. (2004). The neural correlates of grammatical gender decisions in Spanish. *Neuroreport, 15*(5), 863–866. https://doi.org/10.1097/00001756-200404090-00026

Hernández-Pina, F. (1984). *Teorías psicolingüísticas y su aplicación a la adquisición del español como lengua materna*. Siglo XXI de España Editores, S.A.

Hinskens, F., van Hout, R., Muysken, P., & van Wijngaarden, A. (2020). Variation and change in grammatical gender marking: The case of Dutch ethnolects. *Linguistics, 59*(1), 75–100. https://doi.org/10.1515/ling-2020-0265

Hohlfeld, A. (2006). Accessing grammatical gender in German: The impact of gender-marking regularities. *Applied PsychoLinguistics, 27*(2), 127–142. https://doi.org/10.1017/S0142716406060218

Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research, 29*(1), 33–56. http://www.jstor.org/stable/43103911.

Jackson-Maldonado, D., Thal, D., Marchman, V. A., Bates, E., & Gutierrez-Clellen, V. (1993). Early lexical development in Spanish speaking infants and toddlers. *Journal of Child Language, 20*, 523–549. https://doi.org/10.1017/S0305000900008461

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods, 42*, 643–650. https://doi.org/10.3758/BRM.42.3.643

Kidd, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental Psychology, 48*(1), 171–184. https://doi.org/10.1037/a0025405

Köpcke, K.-M. (1982). *Untersuchungen zum Genussystem der deutschen Gegenwartssprache [investigations on the gender system of the contemporary German language]*. Tübingen.

Köpcke, K.-M., & Zubin, D. (1983). Die kognitive Organisation der Genuszuweisung zu den einsilbigen Nomen der deutschen Gegenwartssprache [the cognitive organization of gender allocation to monosyllabic nouns in contemporary German]. *Zeitschrift für germanistische Linguistik, 11*, 166–182.

Kupisch, T., Geiß, M., Mitrofanova, N., & Westergaard, M. (2018). *Gender cues in L1 Russian children acquiring German as an early L2*. Münster, Germany: EuroSLA 28 - Universität Münster.

Kupisch, T., Geiß, M., Mitrofanova, N., & Westergaard, M. (2022). Structural and phonological cues for gender assignment in monolingual and bilingual children acquiring German. Experiments with real and nonce words. *Glossa: a journal of general linguistics, 7*(1). https://doi.org/10.16995/glossa.5696

Kupisch, T., Müller, N., & Cantone, K. F. (2002). Gender in monolingual and bilingual first language acquisition: Comparing Italian and French. *Lingue e Linguaggio, 1*(1), 107–149. https://doi.org/10.1418/7559

Lété, B., Peereman, R., & Fayol, M. (2008). Consistency and word-frequency effects on spelling among first- to fifth-grade French children: A regression-based study. *Journal of Memory and Language, 58*(4), 952–977. https://doi.org/10.1016/j.jml.2008.01.001

Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science, 18*(3), 193–198. https://doi.org/10.1111/j.1467-9280.2007.01871.x

Lindsey, B. A., & Gerken, L. (2012). The role of morphophonological regularity in young Spanish-speaking children's production of gendered noun phrases. *Journal of Child Language, 39*(4), 753–776. https://doi.org/10.1017/S0305000911000250

MacWhinney, B. (1987). The competition model. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 249–308). Lawrence Erlbaum Associates, Inc.

MacWhinney, B. (2004). Parameters or cues? *Bilingualism: Language and Cognition, 7*(1), 35–36. https://doi.org/10.1017/S1366728904001233

MacWhinney, B., Leinbach, J., Taraban, R., & McDonald, J. (1989). Language learning: Cues or rules? *Journal of Memory and Language, 28*, 255–277. https://doi.org/10.1016/0749-596X(89)90033-8

Mariscal, S. (2008). Early acquisition of gender agreement in the Spanish noun phrase: Starting small. *Journal of Child Language, 35*, 1–29. https://doi.org/10.1017/S0305000908008908

Matthews, C. (2010). On the nature of phonological cues in the acquisition of French gender categories: Evidence from instance-based learning models. *Lingua, 120*, 879–900. https://doi.org/10.1016/J.LINGUA.2009.06.007

Miller, P., & Clark, M. D. (2011). Phonemic awareness is not necessary to become a skilled deaf reader. *Journal of Developmental and Physical Disabilities, 23*, 459–476. https://doi.org/10.1007/s10882-011-9246-0

Mills, A. E. (1986). *The acquisition of gender: A study of English and German*. Springer-Verlag. Retrieved from https://pure.uva.nl/ws/files/1894686/27859_TheAcquisitionofGender.pdf.

Monster, I., Tellings, A., Burk, W. J., Keuning, J., Segers, E., & Verhoeven, L. (2022). Word properties predicting children's word recognition. *Scientific Studies of Reading, 26*(5), 373–389. https://doi.org/10.1080/10888438.2021.2020795

Murillo, J. (1999). La marcación del género en los sustantivos del español: entre la flexión y la derivación. *Filología y Lingüística, 25*(1), 181–192.

Oliphant, K. (1998). Acquisition of grammatical gender in Italian as a foreign language. *The Canadian Modern Language Review, 54*(2), 239–262. https://doi.org/10.3138/cmlr.54.2.239

Padovani, R., Calandra-Buonaura, G., Cacciari, C., Benuzzi, F., & Nichelli, P. (2005). Grammatical gender in the brain: Evidence from an fMRI study on Italian. *Brain Research Bulletin, 65*(4), 301–308. https://doi.org/10.1016/j.brainresbull.2004.11.025

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

Parmentier, F., Comesaña, M., & Soares, A. P. (2017). Disentangling the effects of word frequency and contextual diversity on serial recall performance. *Quarterly Journal of Experimental Psychology, 70*(1), 1–17. https://doi.org/10.1080/17470218.2015.1105268

Perea, M., Soares, A. P., & Comesaña, M. (2013). Contextual diversity is a main determinant of word identification times in young readers. *Journal of Experimental Child Psychology, 116*(1), 37–44. https://doi.org/10.1016/j.jecp.2012.10.014

Pérez-Pereira, M. (1991). The acquisition of gender: What Spanish children tell us. *Journal of Child Language, 18*(3), 571–590. https://doi.org/10.1017/s0305000900011259

Pierrehumbert, J. B. (2003). Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech, 46*(2–3), 115–154. https://doi.org/10.1177/00238309030460020501

Postma, G., & van Reenen, P. (2009). De wisselwerking tussen spontane palatalisatie, Umlaut en genus [language-internal mechanisms of koineisation: 3-to-2 gender decay in the Northern Netherlands]. *Taal en tongval, 18*, 109–152. Retrieved from https://www.academia.edu/4132995/Language-internal_mechanisms_of_koineization-3-to-3_gender_decay_in_the_Northern_Netherlands.

Quiñones, I., Molinaro, N., Mancini, S., Hernández-Cabrera, J. A., Barber, H., & Carreiras, M. (2018). Tracing the interplay between syntactic and lexical features: fMRI evidence from agreement comprehension. *NeuroImage, 175*, 259–271. https://doi.org/10.1016/j.neuroimage.2018.03.069

Rio-Torto, G., Soares-Rodrigues, A., Pereira, I., Pereira, R., & Ribeiro, S. (2013). *Gramática derivacional do Português [derivational grammar of Portuguese]*. Imprensa da Universidade de Coimbra.

Rodina, Y. (2008). *Semantics and morphology: the acquisition of grammatical gender in Russian*. Doctoral dissertation. University of Tromso. Retrieved from https://munin.uit.no/handle/10037/2247.

Rodina, Y., & Westergaard, M. (2013). Two gender systems in one mind: The acquisition of grammatical gender in Norwegian–Russian bilinguals. In P. Siemund, I. Gogolin, M. E. Schulz, & J. Davydova (Eds.), *Multilingualism and language diversity in urban areas: Acquisition, identities, space, education* (pp. 95–126). John Benjamins. https://doi.org/10.1075/hsld.1.05rod.

Rodina, Y., & Westergaard, M. (2017). Grammatical gender in bilingual Norwegian–Russian acquisition: The role of input and transparency. *Bilingualism: Language and Cognition, 20*(1), 197–214. https://doi.org/10.1017/S1366728915000668

Royle, P., & Stine, I. (2013). The French noun phrase in preschool children with SLI: morphosyntactic and error analyses. *Journal of Child Language, 40*(5), 945–970. https://doi.org/10.1017/S0305000912000414

Russo, A. G., Esposito, F., Laudanna, A., Mancuso, A., Di Salle, F., Elia, A., & De Martino, M. (2021). The neural substrate of noun morphological inflection: A rapid event-related fMRI study in Italian. *Neuropsychologia, 151*, 107–699. https://doi.org/10.1016/j.neuropsychologia.2020.107699

Salamoura, A., & Williams, J. N. (2007). The representation of grammatical gender in the bilingual lexicon: Evidence from Greek and German. *Bilingualism: Language and Cognition, 10*(3), 257–275. https://doi.org/10.1017/S1366728907003069

Sá-Leite, A. R. (2021). *Representation and processing of grammatical gender: Analysing the gender congruency effect*. Doctoral dissertation. University of Santiago de Compostela. Retrieved from https://minerva.usc.es/xmlui/handle/10347/27486.

Sá-Leite, A. R., Flores, C., Eira, C., Haro, J., & Comesaña, M. (2023). Language balance rather than age of acquisition: A study on the cross-linguistic gender congruency effect in Portuguese–German bilinguals. *Bilingualism: Language and Cognition, 26*(5), 1079–1092. https://doi.org/10.1017/S1366728923000378

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*, Article 661175. https://doi.org/10.3389/fpsyg.2021.661175

Sá-Leite, A. R., Luna, K., Tomaz, Â., Fraga, I., & Comesaña, M. (2022). The mechanisms underlying grammatical gender selection in language production: A meta-analysis of the gender congruency effect. *Cognition, 224*, Article 105060. https://doi.org/10.1016/j.cognition.2022.105060

Sá-Leite, A. R., Tomaz, Â., Hernández-Cabrera, J. A., Fraga, I., Acuña-Fariña, C., & Comesaña, M. (2022). What a transparent Romance language with a Germanic gender-determiner mapping tells us about gender retrieval: Insights from European Portuguese. *Psicológica, 43*(2), 1–60. https://doi.org/10.20350/digitalCSIC/14777

Shadid, A. C. (1997). La formación del género en español. *Filología y Lingüística, 23*(2), 153–159. Retrieved from https://revistas.ucr.ac.cr/index.php/filyling/article/download/20410/20661/0.

Soares, A. P., Machado, J., Costa, A., Iriarte, Á., Simões, A., de Almeida, J. J., … 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

Soares, A. P., Medeiros, J. C., Simões, A., Machado, J., Costa, A., Iriarte, Á., … Comesaña, M. (2014). ESCOLEX: A grade-level lexical database from European Portuguese elementary to middle school textbooks. *Behavior Research Methods, 46*, 240–253. https://doi.org/10.3758/s13428-013-0350-1

Spinner, P., & Juffs, A. (2008). L2 grammatical gender in a complex morphological system: The case of German. *International Review of Applied Linguistics in Language Teaching, 46*(4), 315–348. https://doi.org/10.1515/IRAL.2008.014

Storme, B. (2021). Implicational generalizations in morphological syncretism: The role of communicative biases. *Journal of Linguistics, 58*(2), 381–421. https://doi.org/10.1017/s0022226721000207

Szagun, G., Stumper, B., Sondag, N., & Franik, M. (2007). The acquisition of gender marking by young German-speaking children: Evidence for learning guided by phonological regularities. *Journal of Child Language, 34*, 445–471. https://doi.org/10.1017/S0305000906007951

Taft, M. (2023). Localist lexical representation of polymorphemic words: The AUSTRAL model. In D. Crepaldi (Ed.), *Linguistic morphology in the mind and brain* (pp. 152–166). Routledge. https://doi.org/10.4324/9781003159759-11.

Taft, M., & Meunier, F. (1998). Lexical representation of gender: A quasiregular domain. *Journal of Psycholinguistic Research, 27*(1), 23–45. https://doi.org/10.1023/a:1023270723066

Teschner, R. V., & Russell, W. M. (1984). The gender patterns of Spanish nouns: An inverse dictionary-based analysis. *Hispanic Linguistics, 1*, 115–132.

Treiman, R., & Kessler, B. (2022). Statistical learning in word reading and spelling across languages and writing systems. *Scientific Studies of Reading, 26*(2), 139–149. https://doi.org/10.1080/10888438.2021.1920951

Trezek, B. J., Wang, Y., Woods, D. G., Gampp, T. L., & Paul, P. V. (2007). Using visual phonics to supplement beginning reading instruction for students who are deaf or hard of hearing. *Journal of Deaf Studies and Deaf Education, 12*(3), 373–384. https://doi.org/10.1093/deafed/enm014

Unsworth, S. (2008). Age and input in the acquisition of grammatical gender in Dutch. *Second Language Research, 24*(3), 365–395. https://doi.org/10.1177/0267658308090185

Urrutia, M., Domínguez, A., & Álvarez, C. J. (2009). Gender activation in transparent and opaque words. *Psicothema, 21*(1), 1–8.

van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology (2006), 67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

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

Villalva, A. (2008). *Morfologia do Português [morphology of the Portuguese language]*. Universidade Aberta.

Zevin, J. D., & Seidenberg, M. S. (2002). Age of acquisition effects in word reading and other tasks. *Journal of Memory and Language, 47*(1), 1–29. https://doi.org/10.1006/jmla.2001.2834