

Ladies last. Usage variation and directionality of arrangement of binomial expressions

As damas por último. A variação de uso e a direcionalidade do arranjo das expressões binomiais

Ivona Dömischová¹ 

Romina Leonor Toranzos² 

Konrad Szcześniak^{1,3} 

Myron Tsikandilakis⁴ 

¹ Palacký University, Olomouc, Czech Republic

² State University of Londrina, PR, Brazil

³ University of Silesia, Katowice, Poland

⁴ University of Nottingham, Nottingham, United Kingdom

Email: ivona.domischova@upol.cz

Email: rominaleonortoranzos@gmail.com

Email: konrad.szczeniak@us.edu.pl

Email: tsikandilakismyron@gmail.com

Editora-chefe

Marcia dos Santos
Machado Vieira

Recebido: 19/05/2025

Aceito: 30/06/2025

Como citar:

DÖMISCHOVÁ, Ivona;
TORANZOS, Romina;
SZCZEŚNIAK, Konrad;
TSIKANDILAKIS, Myron.
Ladies last. Usage variation and directionality of arrangement of binomial expressions. *Revista Diadorim*, v.26, n.3, e68476, 2024. doi: <https://doi.org/10.35520/diadorim.2024.v26n3a68476>

ABSTRACT:

This study investigates gendered binomial expressions in Portuguese and Spanish to examine whether linguistic forms reflect and respond to sociocultural shifts toward gender equality. Gendered pairs such as *pai e mãe* ('father and mother') *vs.* *mãe e pai* display some variation that is sociolinguistically meaningful, yet relatively limited. Using *corpus* data from the *Corpus del Español Web-Dialects* (Davies, 2016a) and the *Corpus do Português Web-Dialects* (Davies 2016b), we analyze frequencies and directionality patterns in a sample of gendered binomials. We reveal a consistent male-first bias, with only rare exceptions like *senhoras e senhores* ('ladies and gentlemen'). The degree of ordering flexibility depends

on frequency: more frequent binomials exhibit greater stability, making them less susceptible to reversal. Despite institutional efforts to reduce androcentrism in language, the masculine-default schema remains deeply entrenched, particularly in high-frequency expressions.

Keywords:

androcentrism; irreversible binomials; non-sexist language; variation.

RESUMO:

Este estudo investiga expressões binomiais com marca de género em português e espanhol, com o objetivo de examinar se as formas linguísticas refletem e respondem a mudanças socioculturais em direção à igualdade de género. Pares com marca de género, como *pai e mãe vs. mãe e pai*, revelam alguma variação sociolinguisticamente significativa, embora relativamente limitada. Com base em dados dos corpora *Corpus del Español Web-Dialects* (Davies, 2016a) e *Corpus do Português Web-Dialects* (Davies 2016b), analisamos frequências e padrões de ordenação numa amostra de binómios com marca de género. Revelamos uma preferência sistemática pela ordem masculina em primeiro lugar, com raras exceções como *senhoras e senhores*. O grau de flexibilidade na ordenação depende da frequência: os binómios mais frequentes tendem a ser mais estáveis, sendo menos suscetíveis de inversão. Apesar dos esforços institucionais para reduzir o androcentrismo na linguagem, o padrão de referência masculina continua profundamente enraizado, sobretudo nas expressões de alta frequência.

Palavras-chave:

androcentrismo, binómios irreversíveis, linguagem não-sexista, variação

Introduction

First analyzed by Malkiel (1959), irreversible binomials are expressions such as *black and white*, today studied under the general umbrella of formulaic language research as examples of formulaic sequences. Their formulaic status is reflected in the name itself, where the modifier “irreversible” suggests that only one form is used, and reversing the order (*e.g. white and black*) typically results in the expression appearing odd or unnatural. This irreversibility is a rather straightforward consequence of how the mind is believed to process formulaic sequences, a fact expressed in their definition: “a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar” (Wray, 2002, p. 9).

Quite simply, since a formulaic sequence is treated as a whole—rarely if ever analyzed—its form solidifies and resists changes. Thus, examples like *odds and ends*, *dribs and drabs*, or *hue and cry* are never really found in their reversed variants, because they are too opaque to allow analysis.

However, speakers do on occasion analyze irreversible binomials. People have the option to unpack an expression and treat the component words individually. This is true of transparent binomials, especially those composed of words denoting two distinct entities such as *mom and pop*, originally attested as *pop and mom* just a century ago. While most have a dominant variant, binomials consisting of gender opposites exhibit noticeable variation—an otherwise common feature of language use in general (Toranzos, 2024; Travassos, Vieira, 2019).

Binomials have attracted much research attention: Mollin (2013) traced the usage frequencies of binomials over two centuries and identified factors responsible for binomial orderings. Goldberg and Lee (2021) investigated the cognitive mechanisms of accessibility to account for the changes in the ordering of English male-female binomials (such as *mother and dad*). Szcześniak (2024) took advantage of the formal flexibility of binomials to test Schmidt's *Noticing Hypothesis*. Fonseca's (2007) study of English legalese binomials translated into Portuguese yielded a picture of almost complete irreversibility of binomials in legal documents. The present study expands on this line of work. It looks at Portuguese and Spanish binomials as a reflection of socially-driven language changes. Specifically, we will focus on how the push toward gender equality affects language use. We will do so by analyzing variation in the form of binomial pairs such as *pai e mãe* ('father and mother'), *marido e mulher* ('husband and wife'), and *rei e rainha* ('king and queen'), which are also attested in reversed variants *mãe e pai* ('mother and father'), *mulher e marido* ('wife and husband'), and *rainha e rei* ('queen and king'). The main interest of the present paper is to study usage frequencies exhibited by specific binomial expressions in Portuguese and Spanish.

Based on the available *corpus* data, we present two main findings. As a group, binomial expressions exhibit a very clear directionality of arrangement, in favor of male-first pairs. There are very few exceptions to this pattern. Second, usage variation is subject to frequency pressures, such that compared to the dominant variant (*e.g.* *pai e mãe*), the relative frequency of the reversed variant decreases. The more frequent a binomial expression, the less variation it is likely to exhibit.

Reversibility potential

Despite what their name suggests, irreversible binomials exhibit varying degrees of reversibility. While there are examples of binomials found almost exclusively in one dominant variant (*e.g.* *spic and span*), a large number of these conjoined phrases can be attested in both forms. A particularly interesting example is a group of binomials

including male-female pairs such as *filho e filha* ('son and daughter'). Usage patterns exhibited by gendered binomials represent an interesting source of information on the effects of societal changes on language use.

Male-first default

The social changes toward gender equality occurring throughout the 20th century were not without consequences for language use. Many expressions, which originally followed the male-first pattern, responded to the egalitarian spirit and started being attested increasingly more frequently in the reversed form. The variants *father and mother* and *mother and father*, shown in Figure 1, are among numerous such examples in English.

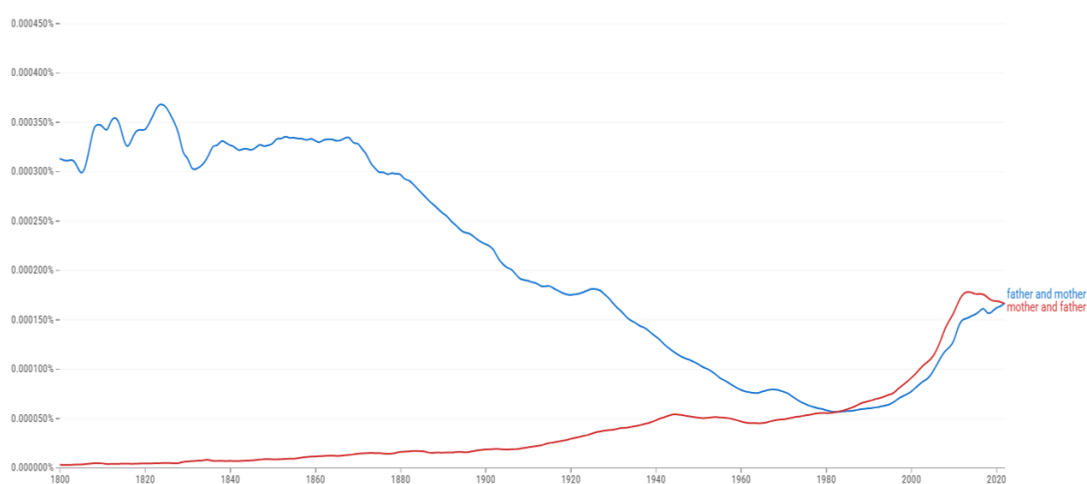


Figure 1. Diachronic shifts in preferred orderings from male- to female-first *mother and father*

Source: according to the Google Books N-grams corpus (GOOGLE, s.d.).

Goldberg and Lee (2021) identify several examples of binomials which have shifted from the male-first to the female-first ordering in a matter of decades in the twentieth century.

The tendency toward gender leveling is also visible in German, where some originally male-first binomials have over the decades become female-first. For example, *Mann und Frau* ('man and woman') is now attested as frequently as in the *Frau und Mann* variant (Figure 2). It is interesting to observe that German artist Till Lindemann's 2020 album is titled *Frau und Mann*, likely contributing to an increase in the frequency of the female-first variant and, consequently, its entrenchment in German speakers.

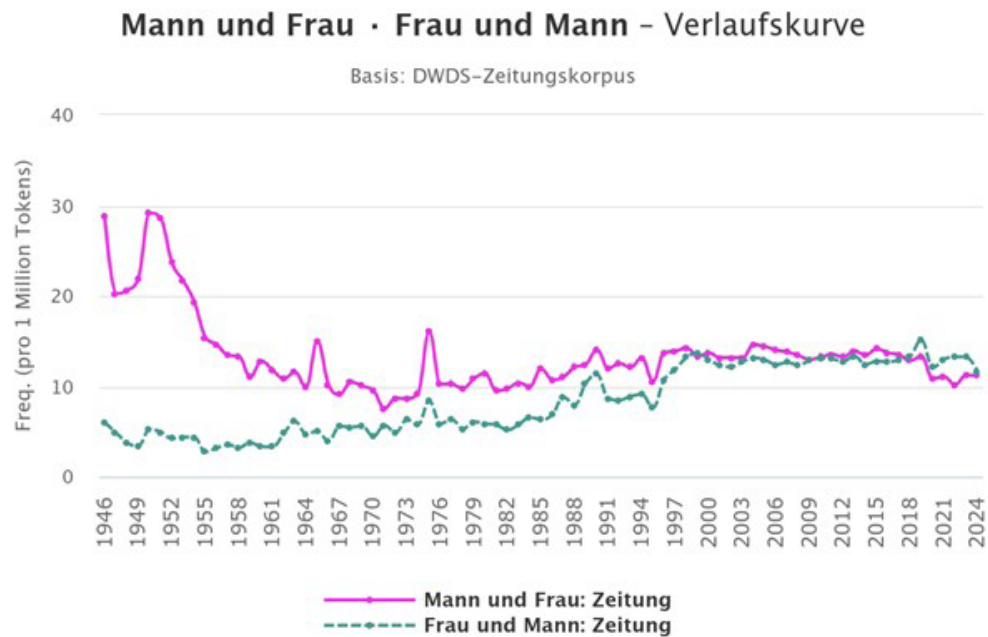


Figure 2. Frequency timeline for the German binomials *Mann und Frau* and *Frau und Mann*
Source: according to the Digitales Wörterbuch der deutschen Sprache DWDS corpus (DWDS, s.d.)

This equal spread can be seen in the use of *Tochter und Sohn* and *Sohn und Tochter* ('daughter and son', Figure 3) and in many other cases, the female-first variant becomes dominant, as in *Oma und Opa* ('granny and grandpa', Figure 4), *Mutter und Vater* ('mother and father', Figure 5), and *Damen und Herren* ('ladies and gentlemen', Figure 6).

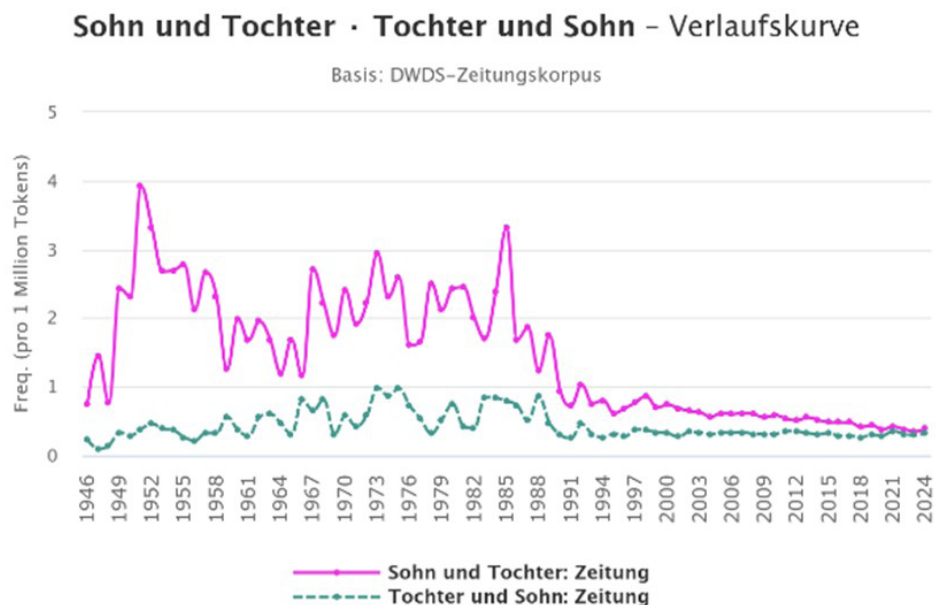


Figure 3. Frequency timeline for the German binomials *Sohn und Tochter* and *Tochter und Sohn*
Source: according to the Digitales Wörterbuch der deutschen Sprache DWDS corpus (DWDS, s.d.).

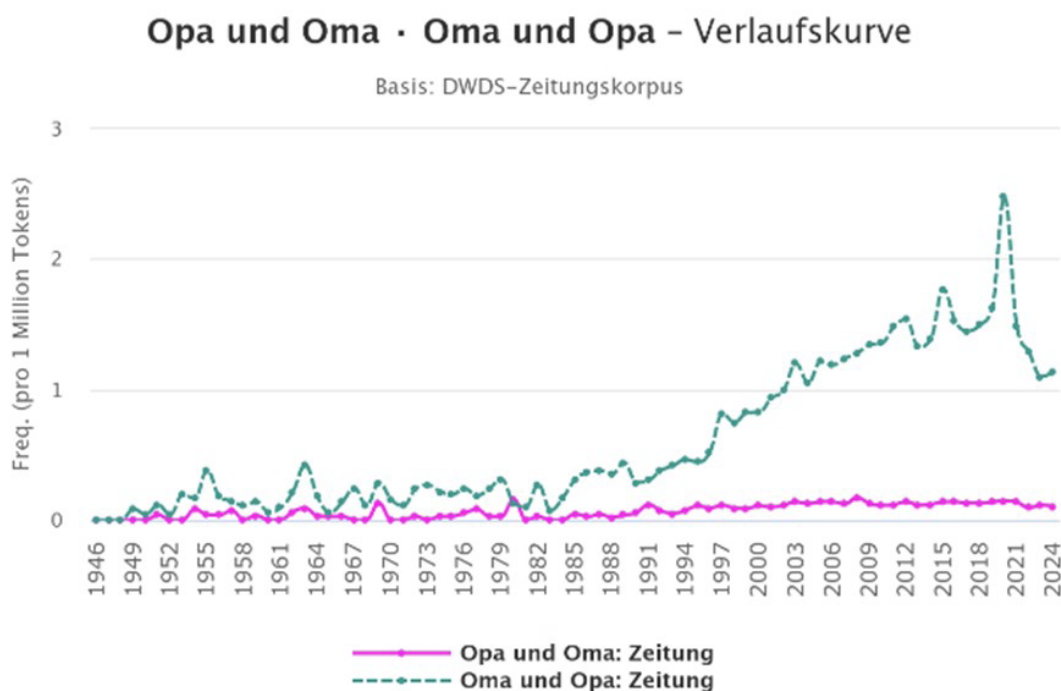


Figure 4. Frequency timeline for the German binomials *Opa und Oma* and *Oma und Opa*
Source: according to the Digitales Wörterbuch der deutschen Sprache DWDS *corpus* (DWDS, s.d.).

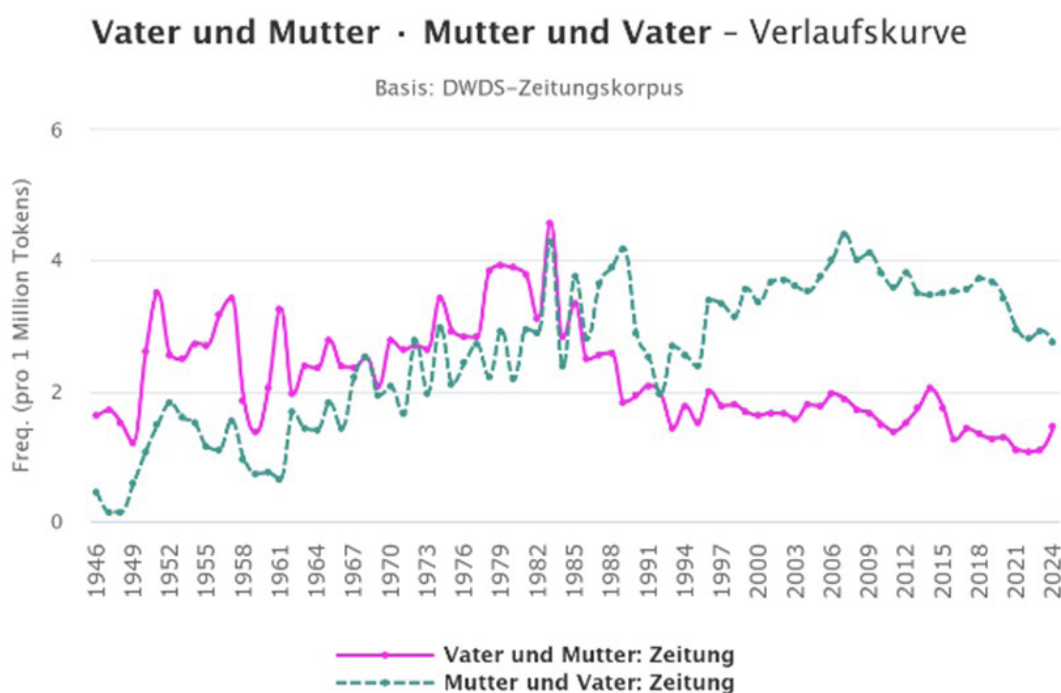


Figure 5. Frequency timeline for the German binomials *Vater und Mutter* and *Mutter und Vater*
Source: according to the Digitales Wörterbuch der deutschen Sprache DWDS *corpus* (DWDS, s.d.)

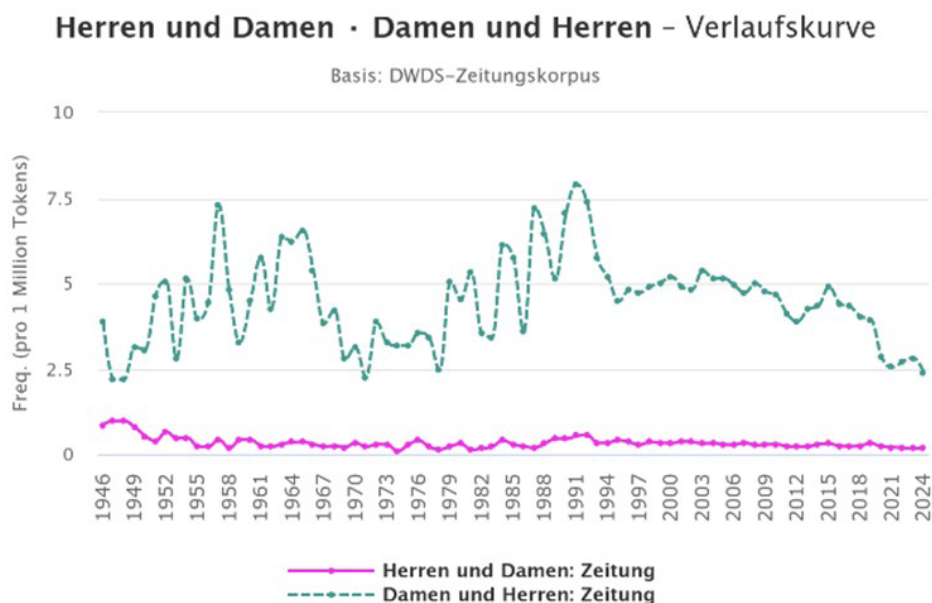


Figure 6. Frequency timeline for the German binomials *Herren und Damen* and *Damen und Herren*

Source: according to the Digitales Wörterbuch der deutschen Sprache DWDS *corpus* (DWDS, s.d.)

Interestingly, this tendency does not seem to occur in Spanish. As can be seen in Figure 7, while the phrase *hija e hijo* (‘daughter and son’) is attested and its frequency has increased in recent decades, the male-first variant *hijo e hija* (‘son and daughter’) predominates.

In fact, in both Portuguese and Spanish, the male-first variant is the dominant ordering for a great majority of binomials. The only significant exception is the phrase *senhoras e senhores* (‘ladies and gentlemen’) in Portuguese and the Spanish equivalent *señoras y señores*. Table 1 shows the frequencies for the binomials according to *Corpus del Español Web-Dialects* (2 billion words, Davies 2016a) and *Corpus do Português Web-Dialects* (1 billion words, Davies 2016b).

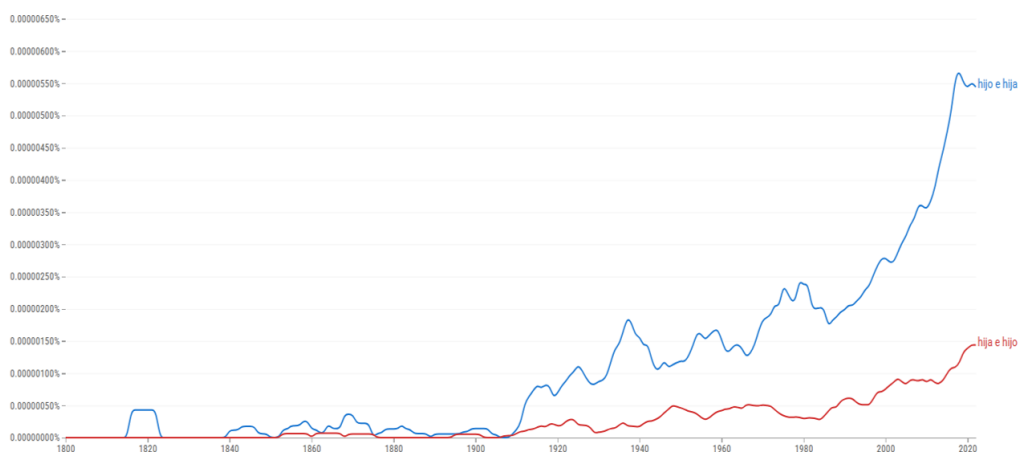


Figure 7. Diachronic changes in frequencies of *hijo e hija* (‘son and daughter’), which does not yield to *hija e hijo*

Source: according to the Google Books N-grams *corpus* (GOOGLE, s.d.).

Table 1. Raw frequencies of the Spanish and Portuguese binomial equivalent of *ladies and gentlemen* (*Corpus del Español Web-Dialects* and *Corpus do Português Web-Dialects*)

| Binomial | female-first freq | male-first freq | quotient ratio |
|----------------------------|--------------------------|------------------------|-----------------------|
| <i>señoras y señores</i> | 2547 | 964 | 2.64 |
| <i>senhoras e senhores</i> | 854 | 419 | 2.03 |

Source: Davies 2016a, 2016b.

A brief glance at Table 2 with the frequencies for some of the most common binomials in Spanish is enough to see that in every case, there is a clear preference for male-first variants.

Interestingly, this preference is also visible in what can be labeled “inclusive vocatives.” These are forms like *argentinos y argentinas*, used by politicians to address the nation, their constituents, or their public to convey a sense of egalitarian inclusiveness. However, while special care is taken to treat both sexes equally, again the male-first variant is the default, as is clear in Table 3. The only exception is *mexicanas y mexicanos*, but this represents the lowest degree of bias (quotient ratio of 1.43) among all the remaining pairs, where the preference toward the male-first variant can in some cases exceed quotient ratios of 10.

Table 2. Raw frequencies of the most common Spanish binomials (*Corpus del Español Web-Dialects*)

| Binomial | male-first freq | female-first freq | quotient ratio |
|----------------------------|------------------------|--------------------------|-----------------------|
| <i>hombres y mujeres</i> | 114416 | 30055 | 3.80 |
| <i>padres y madres</i> | 15482 | 7616 | 2.03 |
| <i>hijos e hijas</i> | 8683 | 2784 | 3.12 |
| <i>chicos y chicas</i> | 7773 | 1545 | 5.03 |
| <i>hombre y mujer</i> | 6133 | 667 | 9.19 |
| <i>hermanos y hermanas</i> | 5838 | 1626 | 3.59 |
| <i>padre y madre</i> | 3403 | 1229 | 2.77 |
| <i>amigos y amigas</i> | 2489 | 1507 | 1.65 |
| <i>reyes y reinas</i> | 541 | 104 | 5.20 |
| <i>hijo e hija</i> | 404 | 96 | 4.21 |
| <i>rey y reina</i> | 378 | 249 | 1.52 |
| <i>señor y señora</i> | 221 | 18 | 12.28 |
| <i>chico y chica</i> | 191 | 28 | 6.82 |

Source: Davies, 2016a.

Table 3. Raw frequencies of the most common Spanish binomials (*Corpus del Español Web-Dialects*)

| Binomial | dominant var | alt var | quotient ratio |
|-------------------------------------|---------------------|----------------|-----------------------|
| <i>argentinos y argentinas</i> | 673 | 109 | 6.17 |
| <i>bolivianos y bolivianas</i> | 443 | 219 | 2.02 |
| <i>brasileños y brasileñas</i> | 104 | 19 | 5.47 |
| <i>chilenos y chilenas</i> | 2320 | 998 | 2.32 |
| <i>colombianos y colombianas</i> | 815 | 220 | 3.70 |
| <i>españoles y españolas</i> | 317 | 29 | 10.93 |
| <i>hondureños y hondureñas</i> | 118 | 13 | 9.08 |
| <i>mexicanas y mexicanos</i> | 519 | 362 | 1.43 |
| <i>panameños y panameñas</i> | 167 | 17 | 9.82 |
| <i>paraguayos y paraguayas</i> | 163 | 12 | 13.58 |
| <i>peruanos y peruanas</i> | 728 | 187 | 3.89 |
| <i>uruguayos y uruguayas</i> | 105 | 41 | 2.56 |
| <i>venezolanos y venezolanas</i> | 2751 | 710 | 3.87 |

Source: Davies, 2016a.

A look at Portuguese gendered-pair binomials reveals the same strong bias toward male-dominant variants. Table 4 includes usage data for the most common examples. In the second column, the frequency of the dominant variant is given. This is followed by a percentage ratio, which reflects the share of frequencies of use of the dominant form relative to the alternate variant. Then the log-odds ratio is shown, which is a measure that reflects the relative preference (or dominance) of one variant over its reversed form. Again, in all cases except one (*senhoras e senhores*, ‘ladies and gentlemen’), the male-first form dominates. (Table 4 does not include the frequencies for reversed variants; these can be found in the Appendix.)

Some very infrequent uses are only found in the male-first variant. None of the following expressions have been found in its reversed (female-first) form:

candidato ou candidata, herói e heroína, primo e prima, herdeiros e herdeiras, neto e neta, herdeiro ou herdeira, autor e autora, aluno e aluna

And just like in Spanish, inclusive vocatives follow the male-first pattern (Table 5). There is one exception, *angolanas e angolanos*, and again it represents the lowest degree of preference (1.41), compared to the other pairs.

Table 4. Usage frequencies of the dominant variants of Portuguese binomials

| Binomial | Raw freq | % ratio | log | Binomial | Raw freq | % ratio | log |
|------------------------------------|-----------------|----------------|------------|----------------------------------|-----------------|----------------|------------|
| <i>homens e mulheres</i> | | | | | | | |
| <i>Adão e Eva</i> | 18401 | 88.05 | 2.88 | <i>avô e avó</i> | 45 | 80.36 | 2.03 |
| <i>pai e mãe</i> | 3319 | 98.43 | 5.97 | <i>avôs e avós</i> | 43 | 67.19 | 1.03 |
| <i>irmãos e irmãs</i> | 2885 | 82.31 | 2.22 | <i>namorados e namoradas</i> | 38 | 74.51 | 1.55 |
| <i>homem e mulher</i> | 2748 | 93.63 | 3.88 | <i>deputados e deputadas</i> | 38 | 84.44 | 2.44 |
| <i>meninos e meninas</i> | 2547 | 95.64 | 4.46 | <i>ator e atriz</i> | 37 | 84.09 | 2.4 |
| <i>amigos e amigas</i> | 1996 | 76.24 | 1.68 | <i>advogados e advogadas</i> | 36 | 85.71 | 2.58 |
| <i>pais e mães</i> | 1946 | 86.16 | 2.63 | <i>genros e noras</i> | 32 | 51.61 | 0.09 |
| <i>marido e mulher</i> | 1736 | 68.16 | 1.1 | <i>autores e autoras</i> | 31 | 83.78 | 2.37 |
| <i>filhos e filhas</i> | 1441 | 98.7 | 6.24 | <i>rapaz e rapariga</i> | 30 | 93.75 | 3.91 |
| <i>senhoras e senhores</i> | 1323 | 90.55 | 3.26 | <i>noivas e noivos</i> | 25 | 56.82 | 0.4 |
| <i>rapazes e raparigas</i> | 854 | 67.09 | 1.03 | <i>tio e tia</i> | 24 | 66.67 | 1 |
| <i>macho e fêmea</i> | 605 | 85.69 | 2.58 | <i>noivo e noiva</i> | 22 | 59.46 | 0.55 |
| <i>menino ou menina</i> | 521 | 97.57 | 5.32 | <i>dono ou dona</i> | 22 | 78.57 | 1.87 |
| <i>atores e atrizes</i> | 473 | 81.13 | 2.1 | <i>esposos e esposas</i> | 22 | 84.62 | 2.46 |
| <i>reis e rainhas</i> | 387 | 89.38 | 3.07 | <i>candidatos e candidatas</i> | 22 | 75.86 | 1.65 |
| <i>machos e fêmeas</i> | 323 | 97.88 | 5.53 | <i>príncipe e princesa</i> | 21 | 87.5 | 2.81 |
| <i>companheiros e companheiras</i> | 292 | 92.11 | 3.55 | <i>sogra e sogra</i> | 21 | 72.41 | 1.39 |
| <i>alunos e alunas</i> | 202 | 77.99 | 1.83 | <i>padrinho e madrinha</i> | 20 | 71.43 | 1.32 |
| <i>príncipes e princesas</i> | 194 | 87.39 | 2.79 | <i>netos e netas</i> | 20 | 76.92 | 1.74 |
| <i>professores e professoras</i> | 173 | 89.18 | 3.04 | <i>genro e nora</i> | 15 | 75 | 1.58 |
| <i>senhor e senhora</i> | 170 | 73.59 | 1.48 | <i>escritores e escritoras</i> | 15 | 88.24 | 2.91 |
| <i>irmão e irmã</i> | 165 | 95.93 | 4.56 | <i>cunhados e cunhadas</i> | 15 | 88.24 | 2.91 |
| <i>tios e tias</i> | 140 | 88.05 | 2.88 | <i>escritor ou escritora</i> | 11 | 91.67 | 3.46 |
| <i>namorado e namorada</i> | 139 | 78.98 | 1.91 | <i>vizinho ou vizinha</i> | 8 | 80 | 2 |
| <i>rei e rainha</i> | 123 | 96.09 | 4.62 | <i>vizinhos e vizinhas</i> | 7 | 70 | 1.22 |
| <i>heróis e heroínas</i> | 121 | 94.53 | 4.11 | <i>professora e professor</i> | 6 | 60 | 0.58 |
| <i>padrinhos e madrinhas</i> | 95 | 95 | 4.25 | <i>cunhado e cunhada</i> | 5 | 83.33 | 2.32 |
| <i>primos e primas</i> | 84 | 60 | 0.58 | <i>sogros e sogras</i> | 5 | 83.33 | 2.32 |
| <i>meninos ou meninas</i> | 76 | 84.44 | 2.44 | <i>Adãos e Evas</i> | 5 | 83.33 | 2.32 |
| <i>cantores e cantoras</i> | 73 | 81.11 | 2.1 | <i>companheiro e companheira</i> | 4 | 66.67 | 1 |
| <i>menino e menina</i> | 71 | 82.56 | 2.24 | <i>donos ou donas</i> | 4 | 80 | 2 |
| <i>esposo e esposa</i> | 62 | 72.09 | 1.37 | <i>cantor e cantora</i> | 2 | 66.67 | 1 |
| <i>filho e filha</i> | 62 | 81.58 | 2.15 | <i>advogado ou advogada</i> | 2 | 50 | 0 |
| <i>maridos e mulheres</i> | 55 | 78.57 | 1.87 | <i>deputado e deputada</i> | 2 | 66.67 | 1 |

Source: Data obtained from the *Corpus do Português Web-Dialects*
<https://www.corpusdoportugues.org/web-dial/>

Table 5. Raw frequencies of the most common Portuguese inclusive vocatives (*Corpus do Português Web-Dialects*)

| Binomial | dominant var | alt var | quotient ratio |
|-------------------------------------|---------------------|----------------|-----------------------|
| <i>brasileiros e brasileiras</i> | 342 | 85 | 4.02 |
| <i>negros e negras</i> | 136 | 56 | 2.43 |
| <i>portugueses e portuguesas</i> | 67 | 22 | 3.04 |
| <i>moçambicanos e moçambicanas</i> | 42 | 25 | 1.68 |
| <i>angolanas e angolanos</i> | 58 | 41 | 1.41 |
| <i>pretos e pretas</i> | 12 | 1 | 12 |
| <i>brancos e brancas</i> | 7 | 4 | 1.75 |

Source: Davies, 2016b.

Stability vs. variation

Recall that irreversible binomials have been claimed to exhibit strong stability in how consistently one order (A and B) appears over the reversed order (B and A). On the other hand, as the data above indicates, at least some variation occurs in many binomials. An important question to consider is whether the degree of stability is a random phenomenon or whether it depends on a discernible variable. One possible determinant of stability is the frequency of a binomial. It can be hypothesized that the more frequent a binomial, the more likely it is to develop a preference for one dominant variant.

This is, in fact, what can be concluded based on the usage data shown graphically in Figure 8. The chart presents the frequencies of binomials from Table 4 above juxtaposed with their log-odds ratios (last column in the Table). Briefly, the vertical axis shows logarithmic ratio values (0-6.24) for individual binomials, where the higher the value, the higher the magnitude (or strength) of preference for the dominant variant.

Before we discuss the data, a few notes on the data presentation are in order. First, the frequencies for the binomials in the present sample are spread over a wide range of values, with many phrases being quite rare (with only a couple of dozens of attestations) and others running in the thousands of attestations. The spread is quite asymmetrical, in that there are fewer highly frequent phrases than those in the lower-frequency ranges. To avoid a chart with most of the datapoints being concentrated in one corner and thus quite unavailable to any visual analysis, the chart has been divided into three binned frequency ranges: 1–100, 101–1000, and above 1000. The first range shows the group of the least frequent phrases, where most binomials are found. The second range is more compressed, and the last range shows the few most frequent binomials. The dataset contains one extreme outlier (*homens e mulheres*, 18401), which has been winsorized (Nicklin; Plonsky, 2020) to the frequency of the next datapoint (*Adão e Eva*, 3319).

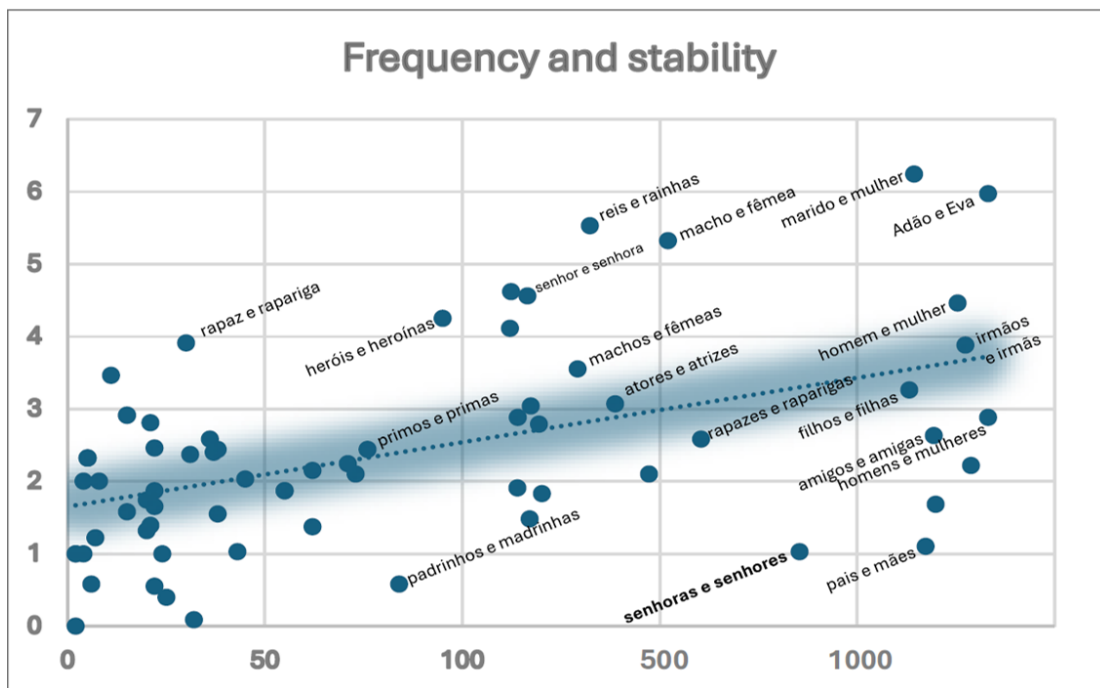


Figure 8. The frequency of a binomial and its stability expressed as a log-odds ratio (vertical axis). The horizontal axis shows frequencies in three logarithmic intervals: from 1-100, 100-1000, and 1000-above.

Source: Authors' own chart created based on data from the *Corpus do Português Web Dialects* <https://www.corpusdoportugues.org/web-dial/>

While the trendline (the dotted line) does not indicate a particularly strong association between frequency and stability, the tendency is clear enough. Among the less frequent binomials (binned range 1), only two exceed a log-odds ratio of 4 (and none approximates the ratio of 5). In the second range, two binomials go beyond a ratio of 5, and in the third binned range, two are around a ratio of 6.

These data can be explained in terms of frequency effects on processing. High-frequency expressions become more entrenched in memory, both for speakers and hearers. This leads to greater expectation and preference for a specific word order. Thus, highly frequent binomials like *Adão e Eva* become entrenched and are processed holistically. Additionally, familiarity reduces cognitive effort. Frequent binomials are processed more quickly and with less ambiguity when they follow a stable, conventional order. In other words, increasing frequency can be said to represent a way of consolidating a variant, making it relatively resistant to reversal.

This is consistent with Goldberg and Lee's (2021) account of variation in terms of cognitive accessibility. Briefly, the more frequently a form is encountered, the faster it tends to be accessed (p. 8). Thus, the choice between *grandfather and grandmother* vs *grandmother and grandfather* depends on a person's prior experience with the phrase. If *grandfather and grandmother* has been encountered more frequently, then it is a fairly straightforward prediction that precisely that male-first ordering will be

accessed and produced first. However, if some people have only rarely come across this binomial, then its ordering in their own production depends on the ordering of similar pairs such as *ma and pa* and *mother and daddy*. In other words, phrases such as *mother and daddy* set the tone for similar kinship binomials which switch their original male-first ordering to female first ordering. This shift in English occurred in the first half of the twentieth century, when the first binomials switched to the female-first ordering and started attracting other binomials in the related semantic field. Interestingly, not all binomials responded to the trend. Highly frequent phrases such as *brothers and sisters* and *men and women* have resisted the shift (Figures 9-10).

However, one should not ignore the variation, however minor, in these cases. The female-first variants of these binomials have been attested despite these binomials' high frequency (and therefore resistance to reversal). In the following section, we look at the factors behind the variation that has not resulted in reversal, either in English or in Portuguese. It can be speculated that the female-first ordering in examples like *women and men* could be due to three possible causes:

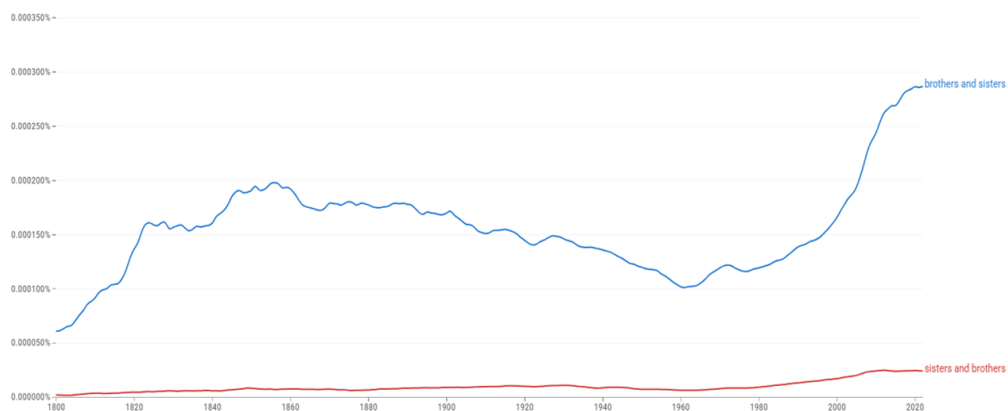


Figure 9. Diachronic changes in frequencies of *brothers and sisters*, which does not yield to *sisters and brothers*

Source: according to the Google Books N-grams corpus (GOOGLE, s.d.).

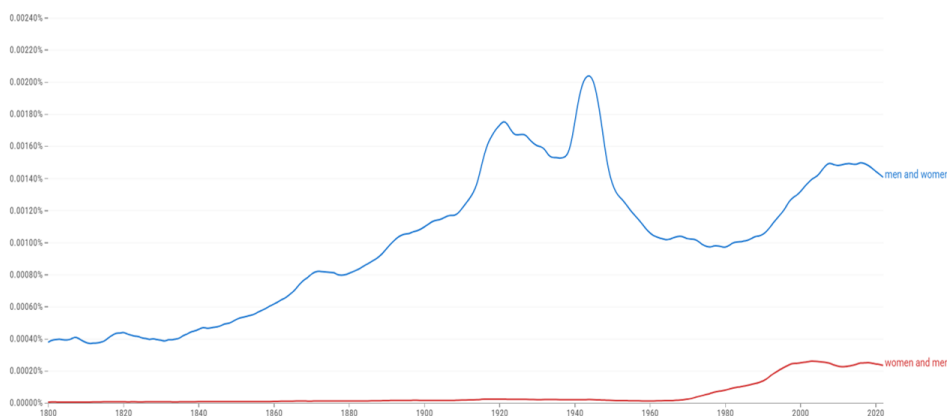


Figure 10 - Diachronic changes in frequencies of *men and women*, which does not yield to *women and men*

Source: according to the Google Books N-grams corpus (GOOGLE, s.d.).

- Attraction from the *mother and daddy* cluster
- Random variation
- Deliberate efforts toward linguistic inclusiveness

It is a safe assumption that the first two causes can be dismissed as these only tend to affect low-frequency binomials (whose male-first variant has not developed sufficient accessibility). Here, in binomials such as *men and women*, their entrenchment and accessibility are strong enough to resist reordering on analogy with the *mother and daddy* group. However, deliberate efforts to promote gender equality do seem like a sufficient pressure leading to variation in ordering. Such efforts are not without challenges or difficulties, to which we turn next.

Gender equality

Consider the following puzzle (found in many logic and reasoning tests):

Doctor Williams lives in Seattle and has a sister who lives in NY. But she doesn't have a brother. Is this scenario possible?

The answer is yes, but most people find it difficult to imagine how Dr. Williams' sister cannot have a brother. They automatically assume that Dr. Williams is a male. That is because gender-neutral titles like *doctor*, *engineer*, *officer*, or *professor* are traditionally associated with masculine referents. There is a masculine default that makes the choice of male-first forms natural and effortless. Any efforts to promote linguistic gender equality must contend with this tendency.

It is interesting to note that the binomial *he or she* can be found in documents published as early as in the 15th century. As expected, despite this binomial's inclusive nature and function, it too exhibits the usual androcentric ordering. The same is true of the Spanish *el y ella* and the Portuguese *ele e ela*. Indeed, it seems that any gender-inclusive binomials automatically assume the male-first order. This can be seen as an instantiation of a higher-level schema, which only sporadic expressions (e.g. *ladies and gentlemen*) escape.

Then there is the consolidation of ordering in more frequent binomials, which makes it particularly difficult to promote linguistic gender equality. While a low frequency expression remains plastic enough to lend itself to reversal (and infrequent binomials are indeed more prone to variation in ordering), attempts to modify highly consolidated binomials are faced with conservative resistance. This is especially true of traditional examples like *Adão e Eva* or the ritualistic *irmãos e irmãs / hermanos y hermanas* ('brothers and sisters').

Of course, efforts are made to reduce androcentric usage. Recent decades have seen numerous publications focusing on the importance of feminization:

Whether due to the growing awareness of the harmful effects of linguistic androcentrism and sexism, or due to the desire of women to name themselves—or at least to be represented on equal footing in public discourse—the fact is that the Spanish language seems to be undergoing a process of feminization (or, if preferred, of reduced androcentrism). Sea por la progresiva toma de conciencia de los efectos nocivos del androcentrismo y sexismo lingüísticos, sea por el deseo de las mujeres de auto-designarse o, al menos, de ser representadas en pie de igualdad en los discursos públicos, lo cierto es que la lengua española parece estar en proceso de feminización (o, si se prefiere, de menor androcentrismo) (Bartolomé; Sánchez-Seco; Araujo, 2009, p. 25).

In a similar spirit, UNESCO has released guidelines inviting

(...) the Director-General to adopt, in the drafting of all the Organization's working documents, a policy aimed at avoiding, as far as possible, the use of terms which refer explicitly or implicitly to only one sex, except where the nature of the document requires the use of such terms (UNESCO, 1987: 4).

A 1999 UNESCO document identifies a series of sexist uses and proposes solutions. Some of the examples are listed in Table 6.

Table 6. Current usage and proposed solutions to promote linguistic inclusivity

| Current usage | Possible solutions |
|-----------------------------------|--|
| <i>el hombre (o los hombres)</i> | <i>los hombres y las mujeres, (...)</i> |
| <i>hombre de negocios</i> | <i>los hombres y las mujeres de negocios, (...)</i> |
| <i>hombre de letras</i> | <i>los hombres y las mujeres de letras, los literatos y las literatas, (...)</i> |
| <i>los niños</i> | <i>los niños y las niñas</i> |
| <i>los muchachos, los jóvenes</i> | <i>los muchachos y las muchachas, los jóvenes y las jóvenes,</i> |
| <i>los profesores</i> | <i>los profesores y las profesoras, (...)</i> |
| <i>los electores</i> | <i>los electores y las electoras, (...)</i> |

However, note that even here, the desired gender equality efforts stop short of the female-first alternative. In not a single case of the proposed solutions is the reversed order recommended. Even in publications that do mention the reversed ordering as an alternative, as in *amigas y amigos* in the Papadopoulos (2022) quote below, the variant comes second after the male-first *amigos y amigas*.

Instead of using masculine plural forms to refer to mixed-gender groups (e.g., amigos), they have suggested various ways to represent both masculine and feminine forms together (e.g., amigos y amigas, amigas y amigos). En vez de usar formas plurales masculinas para referirse a grupos de género mixto (e.g. amigos), han sugerido de varias maneras representar las formas masculinas y femeninas juntas (e.g. amigas y amigas, amigas y amigos) (Papadopoulos, 2022, p. 33).

As outlined above, successful reversals are few. *Senhoras e senhores*, *señoras y señores*, and their Germanic equivalents such as *Damen und Herren* are products of decades of variation, which ended up yielding a female-first dominant variant. These developments are clearly inspired by the English *ladies and gentlemen*, which gained currency in the 18th century.

But the shift to *ladies and gentlemen* in English and to *senhoras e senhores* in Portuguese did not happen overnight. Such changes go against the inertia of deeply ingrained patterns of androcentric usage, which has been accepted as norm for a long time. In fact, it is important to note that the exclusion of feminine forms was initially even more extreme. Guerrero (2016) identifies the origins of current asymmetric usage styles in the early stages of formation of Spanish dialects, where agent nouns ending in *-or*, such as *trabajador* (“worker”), did not have a feminine form in the early stages of Castilian. She also quotes Zenenko’s (1983) observation that until the 16th century, manuscripts contain forms such as “mulher português” (‘Portuguese-MASC woman’).

Today, efforts to promote inclusive binomials or their more progressive female-first variants face conservative opposition. Documents issued by the *La Fundéu* (*Fundación del Español Urgente*) advise against forms like *los delegados y delegadas* (‘delegates-MASC and delegates-FEM’), arguing that *los delegados* (the masculine form) is a sufficient generic non-sexist form (Bartolomé; Sánchez-Seco; Araujo, 2009, p. 53). Similarly, in Portuguese, Oliveira (2018, p. 9) mentions “the surprise caused by President Sarney’s speech when he addressed the public using the binomial *brasileiras e brasileiros* (‘Brazilian women and Brazilian men,’) that is, reversing the established order in Brazilian society”.

Thus, a core tension lies in the opposition between two competing forces: the growing societal and institutional efforts to promote gender-inclusive or gender-reversed alternatives on the one hand and the deeply entrenched male-first default in binomial ordering, which underly conservative attitudes against change. The male-first preference is not merely habitual but reinforced by frequency effects and cognitive entrenchment, making these sequences resistant to change. At the same time, inclusive language campaigns—often driven by political, feminist, or organizational initiatives—aim to challenge this default by reversing traditional orderings or explicitly naming both genders. This results in a site of sociolinguistic conflict: while reformist impulses foster variation and innovation in binomial structures, the structural weight of established usage resists such change, especially in high-frequency expressions.

Conclusions

The analysis of binomial expressions in Portuguese and Spanish reveals a pervasive male-first ordering bias, confirming the persistence of linguistic androcentrism despite growing awareness and institutional efforts to promote gender-inclusive language. While some low-frequency expressions show a degree of variability, high-frequency binomials such as *pai e mãe* or *hombres y mujeres* exhibit considerable resistance to reversal. This resistance can be attributed to frequency-driven entrenchment, whereby familiar expressions become cognitively fixed and are less amenable to change. Notably, even explicitly inclusive forms tend to default to male-first order, a consequence of linguistic conventions overriding egalitarian intentions. The rare reversals observed—such as *senhoras e senhores*—are exceptions that prove the rule, typically associated with formulaic, ritualistic settings. Overall, the picture that emerges is one of the inertia of grammaticalized gender schemas which persist even among evident usage variation.

References

- BARTOLOMÉ, Mercedes Bengoechea. SÁNCHEZ-SECO, Fernando Centenera. ARAUJO, Verónica Gonzáles. *Efectos de las políticas lingüistas antisexistas y feminización del lenguaje*. Madrid: Instituto de la Mujer, 2009.
- DAVIES, Mark. *Corpus del Español: Web/Dialects*. 2016a. Available online at <http://www.corpusdelespanol.org/web-dial/>. Acesso em: 06 jun. 2025.
- DAVIES, Mark. *Corpus do Português: Web/Dialects*. 2016b. Available online at <http://www.corpusdoportugues.org/web-dial/>. Acesso em: 06 jun. 2025.
- DWDS – DIGITALER WÖRTERBUCH DER DEUTSCHEN SPRACHE. *Ressourcen*. Disponível em: <https://www.dwds.de/d/ressourcen>. Acesso em: 6 jun. 2025.
- FONSECA, Luciana Carvalho. 2007. 425 f. *A tradução de binômios nos contratos de 'common law' à luz da linguística de corpus*. Dissertação (Mestrado em Estudos Linguísticos e Literários em Inglês) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, 2007.
- GOLDBERG, Adele E. LEE, Crystal. Accessibility and Historical Change: An Emergent Cluster Led Uncles and Aunts to Become Aunts and Uncles. *Frontiers in Psychology*, v.12, p, 1-19, maio. 2021.
- GOOGLE. *Google Books Ngram Viewer*. Disponível em: <https://books.google.com/ngrams>. Acesso em: 6 jun. 2025.
- GUERRERO, María Márquez. Bases epistemológicas del debate sobre el sexismo lingüístico. *Arbor*, v. 192, n. 778, p.1-15, mar-abri. 2016.

- MALKIEL, Yakov. Studies in irreversible binomials. *Lingua* 8, p.113-160, 1959.
- MOLLIN, Sandra. Pathways of change in the diachronic development of binomial reversibility in Late Modern American English. *Journal of English Linguistics*, v. 41, n.2, 168–203, jun. 2013.
- NICKLIN, Christopher. PLONSKY, Luke. Outliers in L2 research in applied linguistics: A synthesis and data re-analysis. *Annual Review of Applied Linguistics*, v. 40, p. 26-55, jun. 2020.
- OLIVEIRA, Fernanda Silva de. Tim-Tim Por Tim-Tim: Descrevendo os Binômios do Português Brasileiro com Foco no Ensino de Segunda Língua. *Ensaio em Português como Segunda Língua ou Língua Estrangeira*, n.9, p.1-19, 2018.
- PAPADOPOULOS, Ben. Una breve historia del español no binario/A brief history of genderinclusive Spanish. *DEP. Deportate, esuli, profughe*, v. 48, n.1, p. 31-48, 2022.
- SZCZEŚNIAK, Konrad. The noticing hypothesis and formulaic language. Learnability of non-salient language forms. *Acta Psychologica* 248: 104372, 2024.
- TORANZOS, Romina Leonor. *Estudo Geolinguístico Sobre O Espanhol Da Terra Del Fuego Com Vista à Construção De Um Atlas*. Dissertation (PhD in Linguistics). Londrina: State University of Londrina, 2024.
- TRAVASSOS, Pâmela Fagundes. VIEIRA, Márcia Santos Machado. Uma análise construcionista da variação entre construções com verbo-suporte DAR no PB. *Soletras*, n.37: 272-298, 2019.
- UNESCO. *Guidelines on Gender-Neutral Language*. Paris: Unesco, 1987. https://eige.europa.eu/sites/default/files/unesco_guidelines_gender-neutral_language_0.pdf. Acesso em: 19 jun. 2025.
- UNESCO. *Recomendaciones para un uso no sexista del lenguaje*. Paris: Unesco, 1999. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000114950>. Acesso em: 06 jun. 2025.
- WRAY, Alison. *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press, 2002.
- ZENENKO, G. P. Acerca de la manifestación del género de los sustantivos y adjetivos como categoría gramatical en las lenguas romances ibéricas. *Verba*, n.10, p. 231-247, 1983.

APPENDIX

Frequencies of the most common personal binomials in Portuguese.

| Binomial | Raw frequency | Alt. var. | Quotient prop | % ratio | Log odds ratio |
|-----------------------------|----------------------|------------------|----------------------|----------------|-----------------------|
| homens e mulheres | 18401 | 2498 | 7.37 | 88.05 | 2.88 |
| Adão e Eva | 3319 | 53 | 62.62 | 98.43 | 5.97 |
| pai e mãe | 2885 | 620 | 4.65 | 82.31 | 2.22 |
| irmãos e irmãs | 2748 | 187 | 14.7 | 93.63 | 3.88 |
| homem e mulher | 2547 | 116 | 21.96 | 95.64 | 4.46 |
| meninos e meninas | 1996 | 622 | 3.21 | 76.24 | 1.68 |
| amigos e amigas | 1948 | 314 | 5.25 | 86.16% | 2.63 |
| pais e mães | 1736 | 811 | 2.14 | 68.16 | 1.1 |
| marido e mulher | 1441 | 19 | 75.84 | 98.7 | 6.24 |
| filhos e filhas | 1323 | 138 | 9.59 | 90.55 | 3.26 |
| senhoras e senhores | 854 | 419 | 2.04 | 67.09 | 1.03 |
| rapazes e raparigas | 605 | 101 | 5.99 | 85.69 | 2.58 |
| macho e fêmea | 521 | 13 | 40.08 | 97.57 | 5.32 |
| menino ou menina | 473 | 110 | 4.3 | 81.13 | 2.1 |
| atores e atrizes | 387 | 46 | 8.41 | 89.38 | 3.07 |
| reis e rainhas | 323 | 7 | 46.14 | 97.88 | 5.53 |
| machos e fêmeas | 292 | 25 | 11.68 | 92.11 | 3.55 |
| companheiros e companheiras | 202 | 57 | 3.54 | 77.99 | 1.83 |
| alunos e alunas | 194 | 28 | 6.93 | 87.39 | 2.79 |
| príncipes e princesas | 173 | 21 | 8.24 | 89.18 | 3.04 |
| professores e professoras | 170 | 61 | 2.79 | 73.59 | 1.48 |

APPENDIX

Cont.

| Binomial | Raw frequency | Alt. var. | Quotient prop | % ratio | Log odds ratio |
|-----------------------|----------------------|------------------|----------------------|----------------|-----------------------|
| senhor e senhora | 165 | 7 | 23.57 | 95.93 | 4.56 |
| irmão e irmã | 140 | 19 | 7.37 | 88.05 | 2.88 |
| tios e tias | 139 | 37 | 3.76 | 78.98 | 1.91 |
| namorado e namorada | 123 | 5 | 24.6 | 96.09 | 4.62 |
| rei e rainha | 121 | 7 | 17.29 | 94.53 | 4.11 |
| heróis e heroínas | 95 | 5 | 19 | 95 | 4.25 |
| padrinhos e madrinhas | 84 | 56 | 1.5 | 60 | 0.58 |
| primos e primas | 76 | 14 | 5.43 | 84.44 | 2.44 |
| meninos ou meninas | 73 | 17 | 4.29 | 81.11 | 2.1 |
| cantores e cantoras | 71 | 15 | 4.73 | 82.56 | 2.24 |
| menino e menina | 62 | 24 | 2.58 | 72.09 | 1.37 |
| esposo e esposa | 62 | 14 | 4.43 | 81.58 | 2.15 |
| filho e filha | 55 | 15 | 3.67 | 78.57 | 1.87 |
| maridos e mulheres | 55 | 15 | 3.67 | 78.57 | 1.87 |
| avô e avó | 45 | 11 | 4.09 | 80.36 | 2.03 |
| avôs e avós | 43 | 21 | 2.05 | 67.19 | 1.03 |
| namorados e namoradas | 38 | 13 | 2.92 | 74.51 | 1.55 |
| deputados e deputadas | 38 | 7 | 5.43 | 84.44 | 2.44 |
| ator e atriz | 37 | 7 | 5.29 | 84.09 | 2.4 |
| advogados e advogadas | 36 | 6 | 6 | 85.71 | 2.58 |
| genros e noras | 32 | 30 | 1.07 | 51.61 | 0.09 |
| autores e autoras | 31 | 6 | 5.17 | 83.78 | 2.37 |
| rapaz e rapariga | 30 | 2 | 15 | 93.75 | 3.91 |
| noivas e noivos | 25 | 19 | 1.31 | 56.82 | 0.4 |
| tio e tia | 24 | 12 | 2 | 66.67 | 1 |

APPENDIX

Cont.

| Binomial | Raw frequency | Alt. var. | Quotient prop | % ratio | Log odds ratio |
|----------------------------|----------------------|------------------|----------------------|----------------|-----------------------|
| noivo e noiva | 22 | 15 | 1.47 | 59.46 | 0.55 |
| dono ou dona | 22 | 6 | 3.67 | 78.57 | 1.87 |
| esposos e esposas | 22 | 4 | 5.5 | 84.62 | 2.46 |
| candidatos e candidatas | 22 | 7 | 3.14 | 75.86 | 1.65 |
| príncipe e princesa | 21 | 3 | 7 | 87.5 | 2.81 |
| sogro e sogra | 21 | 8 | 2.63 | 72.41 | 1.39 |
| padrinho e madrinha | 20 | 8 | 2.5 | 71.43 | 1.32 |
| netos e netas | 20 | 6 | 3.33 | 76.92 | 1.74 |
| genro e nora | 15 | 5 | 3 | 75 | 1.58 |
| escritores e escritoras | 15 | 2 | 7.5 | 88.24 | 2.91 |
| cunhados e cunhadas | 15 | 2 | 7.5 | 88.24 | 2.91 |
| escritor ou escritora | 11 | 1 | 11 | 91.67 | 3.46 |
| vizinho ou vizinha | 8 | 2 | 4 | 80 | 2 |
| vizinhos e vizinhas | 7 | 3 | 2.33 | 70 | 1.22 |
| professora e professor | 6 | 4 | 1.5 | 60 | 0.58 |
| cunhado e cunhada | 5 | 1 | 5 | 83.33 | 2.32 |
| sogros e sogras | 5 | 1 | 5 | 83.33 | 2.32 |
| Adãos e Evas | 5 | 1 | 5 | 83.33 | 2.32 |
| companheiro e companheira | 4 | 2 | 2 | 66.67 | 1 |
| donos ou donas | 4 | 1 | 4 | 80 | 2 |
| cantor e cantora | 2 | 1 | 2 | 66.67 | 1 |
| advogado ou advogada | 2 | 2 | 1 | 50 | 0 |
| <i>deputado e deputada</i> | 2 | 1 | 2 | 66.67 | 1 |

Source: Authors' own table created based on data from the Corpus do Português Web Dialects <https://www.corpusdoportugues.org/web-dial/>