DE GRUYTER MOUTON Linguistics 2021; 59(4): 1143–1177

Tanja Ackermann* and Christian Zimmer*

The sound of gender – correlations of name phonology and gender across languages

https://doi.org/10.1515/ling-2020-0027
Received February 19, 2020; accepted October 3, 2020; published online June 30, 2021

Abstract: Our article is dedicated to the relation of a given name's phonological structure and the gender of the referent. Phonology has been shown to play an important role with regard to gender marking on a name in some (Germanic) languages. For example, studies on English and on German have shown in detail that female and male names have significantly different phonological structures. However, little is known whether these phonological patterns are valid beyond (closely related) individual languages. This study, therefore, sets out to assess the relation of gender and the phonological structures of names across different languages/cultures. In order to do so, we analyzed a sample of popular given names from 13 countries. Our results indicate that there are both language/culture-overarching similarities between names used for people of the same gender and language/culture-specific correlations. Finally, our results are interpreted against the backdrop of conventional and synesthetic sound symbolism.

Keywords: gender marking; multifactorial analysis; name phonology; sound symbolism

1 Gender marking on names

Names are usually said to have no lexical meaning. However, there is one piece of information, which is quite commonly coded in personal names: the name bearer's gender (cf. Alford 1988: 66–68). Depending on the language/culture, different types of personal names are involved in the marking of gender, e.g., given names,

***Corresponding authors: Tanja Ackermann and Christian Zimmer**, Institut für Deutsche und Niederländische Philologie, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany, E-mail: tanja.ackermann@fu-berlin.de (T. Ackermann), christian.zimmer@fu-berlin.de (C. Zimmer).

the family name, and/or unofficial names.[1] Among these types of personal names gender marking is assigned most often to the given name (cf. Oelkers 2003: 134), which we focus on in the following. As we will outline in the next section, there are different ways to mark gender on a given name including specific phonological structures (Section 1.1). This will be exemplified by studies on German and English (Section 1.2). These studies show that – in these particular languages – there are strong correlations between semantic (e.g., femininity) and phonological properties (see below). For both German and English, it has been discussed how these correlations relate to the concept of sound symbolism (cf. Section 1.2). However, there has been little agreement yet: While some scholars interpret these correlations as instances of synesthetic sound symbolism (e.g., Cutler et al. 1990; Oelkers 2003, 2004; Pitcher et al. 2013; van de Weijer et al. 2020; Whissell 2001) others argue in favor of conventional sound symbolism (e.g., Cassidy et al. 1999; Hough 2000; Nübling 2018).[2] With our study, we want to add an important perspective to this discussion by considering a range of disparate languages/cultures. This promises to shed new light on this topic as one important controversy centers around the question whether we are dealing with arbitrary and language/culture-specific cues to gender (i.e., conventional sound symbolism; see, e.g., Cassidy et al. 1999: 378) or with non-arbitrary correlations, that are valid beyond individual languages/cultures (i.e., synesthetic sound symbolism; see, e.g., Oelkers 2003: 228). The basic questions we want to answer are: Are correlations of name phonology and gender language/culture-specific or are there language/culture-overarching similarities? Which phonological properties indicate (at least by tendency) gender? Hence, we add a contrastive perspective to the discussion on sound symbolism in the domain of onymic gender marking (and we investigate name giving practices in some countries which have not yet been studied in this respect, see below).

The remainder of this paper is structured as follows: After the short overviews of gender marking on names in general (Section 1.1) and by phonological means in particular (Section 1.2), we summarize two central positions on how phonological gender marking on names relates to the concept of sound symbolism (Section 2). Section 3 is dedicated to our own study: We describe the sample selection, the transcription, the annotation, and the multifactorial modelling of the data. The findings are discussed in Section 4 and Section 5 draws a brief conclusion.

---

1 Both purely linguistic and cultural aspects are highly relevant for name giving practices. Since these domains are intertwined and in many cases cannot be disentangled, we mostly use the notation with slash in the following: *language/culture*.

2 Both concepts will be outlined in Section 2.

1.1 Types of gender marking

According to Alford (1988: 66–68), gender is the most frequently marked information on proper names. Gender marking languages are traditionally assigned to one of three types – the semantic, the formal, or the conventional type:[3]

(i) Gender can be marked by the semantics of the name. In this case, desirable characteristics are assigned via a name according to gender stereotypes. E.g., the Turkish male name *Yılmaz* is etymologically rooted in an adjective meaning 'fearless' and thus fits a male gender stereotype. The same holds for female names such as *Gül* ('rose') which is, for example, associated with beauty and thereby connected to central components of the female gender stereotype. Hence, we are dealing with a case of 'doing gender'. Languages/Cultures belonging to this type of gender marking on names comprise, for example, Japanese and Chinese.

(ii) Gender can also be marked on the basis of formal means (e.g., suffixes). In languages/cultures belonging to this type (e.g., Italian, Ojibwa, or Garo), gender is overtly marked and can be deduced from the form of (a specific part of) a name. For example, Italian names ending in *-a* are (with some exceptions, such as *Luca*) female, while names ending in *-o* are male.

(iii) Finally, there can also be separate inventories in a language/culture (Alford 1988: 65–68). In this case, gender is associated with a name by convention. For example, one simply has to learn that certain names, which are commonly used in Germany (e.g., *Doris*), are used for females while others (including phonologically very similar ones) are used for males (e.g., *Boris*).

However, phonology also seems to play an important role – even in languages/cultures that are assigned to the conventional or the semantic type of gender marking (i/iii).[4] Studies on English (cf., e.g., Cassidy et al. 1999; Slater and Feinman 1985) and German (cf., e.g., Nübling 2009; Oelkers 2003) – languages/cultures which are usually assigned to type (iii) – have shown in detail that female

3 These distinctions are by no means clear-cut and should be understood as based on similarity to an ideal prototype. Many languages display several of these characteristics at the same time: For example, English is usually assigned to type (iii) but individual names such as *Grace* and *Rose* fit very well to type (i) (thanks to an anonymous reviewer for these examples!).

4 Given that languages can exhibit characteristics of different types (see above), it would therefore make sense to add phonological gender marking as a fourth type to Alford's (1988) classification.

and male names show significantly different phonological structures, which are discussed in the next section.[5]

1.2 Prosodic-phonological cues to gender in German (and English)

In her study on German given names Oelkers (2003) showed that female and male names differ significantly with regard to the prosodic-phonological features ‘number of syllable’, ‘main stress’, ‘portion of vowels/ consonants’, and ‘quality of stressed vowel’. The main results are that female names contain on average more syllables than male names, female names show initial stress less often than male names, for which initial stress is the dominant pattern, and – most importantly – female names prove to be more sonorous in general. This can partly be explained by two facts: female names have more vowels than male names and female names show significantly more final vowels on average. The final sound of a personal name has in general proved to be a structural position which particularly contributes to the gender differentiation of given names. A further result is that female names have more stressed front vowels than back vowels on average. According to these results, the German given name *Katharina* [ka.ta.ˈʁiː.na] can be classified as exhibiting predominately ‘female’ phonological structures (four syllables, penultimate stress, balanced portion of vowels and consonants, final vowel, stressed front vowel) while *Rolf* [ˈʁɔlf] exhibits predominantly ‘male’ structures (one syllable, higher portion of consonants, final obstruent, stressed back vowel). Table 1 summarizes these findings on prosodic-phonological differences between female and male names in German (for similar results obtained with other samples and a slightly different method, see Nübling 2009 and Nübling et al. 2015: 131–137) and lists possible phono-semantic explanations for these differences (we will comment on these explanations in Section 2).

In German, these phonological patterns seem to be so firmly established and associated with the respective gender that they can be transferred to other types of names, such as brand names, and are thus used for gender marketing.[6] In this area, the association of phonological properties and a certain gender is used for the marketing of products that have either a male or a female target group

5 For a study dedicated to the morphosyntactic means used for gender-marking on personal names in the languages of the world as well as to the integration of personal names into classificatory systems (namely gender and classifiers) see Handschuh (2019).

6 Recent studies have shown that male and female names approach each other phonologically (see Nübling 2018 on German; see Kürschner 2018 for similar findings on Icelandic). However, there still is a remarkable gap between both groups.

Table 1: Prosodic-phonological patterns of German given names and their relation to sound symbolism (cf. Oelkers 2003: 220).

Prosodic-Phonological Feature	Female Names	Male Names	Phonological 'meaning' (reported for female names)
Number of syllables	More syllables	Fewer syllables	Longer words = higher sonority; **Euphony** = female stereotype
Main stress	More non-initial stress	More initial stress	Deviation from unmarked pattern; **Exotic** = female stereotype
Portion of vowels/ consonants	More vowels	More consonants	Vowels = sonority/ soft sound structure;
Final sound	Vowel final	Consonant final	**Softness** = female stereotype
Quality of stressed vowel	More front vowels	Fewer front vowels	Front vowels = smaller size; **Small size** = female stereotype

(e.g., deodorants; cf. Ackermann 2011). Here, phonological properties are part of other marketing aspects such as the coloring of the products.[7]

Those correlations of (targeted) gender and phonological structure of a name cannot only be found in German. Cassidy et al. (1999) found very similar patterns for (product and) given names in English (see also Cutler et al. 1990; Fredrickson 2007; Pitcher et al. 2013; Slater and Feinman 1985; Whissell 2001; Wright et al. 2005; for an overview cf. Elsen 2016: 120–126).[8] Just as in German, English female names have more syllables than male names, show more often non-initial stress as well as a greater ratio of open to closed syllables on average and predominantly end in a final vowel. The results are depicted in Table 2.[9]

To date there has been little agreement in the literature on how these findings relate to the concepts of iconicity and sound symbolism: are we dealing with

7 Since findings regarding the meaning of a sound have implications for the field of consumer psychology there is a substantial body of literature on this topic (cf., e.g., Klink 2000; Shrum and Lowrey 2007; Yorkston and Menon 2004).

8 See Suire et al. (2019) and van de Weijer et al. (2020) for recent studies on phonological cues to gender in French and Chinese names. See also Cai and Zhao (2019) who investigated the extent to which native speakers of German and English can infer gender from name phonology in the Chinese variety Min, i.e., a variety the participants do not speak. (The similarity between the titles of the latter study and our paper is due to a coincidence. We learned about this paper after we had chosen the title for our study).

9 See also Slepian and Galinsky (2016) for a study on voiced vs. unvoiced initial phonemes as a cue to gender in American (and Indian) names.

Table 2: Prosodic-phonological patterns of English given names (cf. Cassidy et al. 1999).

Prosodic-Phon. feature	Female names	Male names
Number of syllables	Larger number	Smaller number
Main stress	Non-initial stress	Initial stress
Final sound	Vowel final	Consonant final
Ratio of open to closed syllables	Greater	–

synesthetic or with conventional sound symbolism? This controversy will be addressed in the next section.

2 Synesthetic versus conventional sound symbolism

According to Oelkers (2003), specific prosodic-phonological structures are not distributed randomly over female and male names. Instead, she interprets these differences phono-semantically. The main idea is that certain phonological properties (e.g., the vowel /i/) tend to be associated with semantic concepts (e.g., smallness), some of which are associated with gender stereotypes (e.g., femininity; sexual dimorphism plays a role here). Oelkers (2003) argues that phonological differences between male and female names reflect these associations of phonological property, semantic concept, and gender stereotype. Parents seem to "give names to children in a manner that is aurally (i.e., through sound) metaphorically congruent with gender stereotypes (i.e., shared beliefs about the traits of women and men)" (Slepian and Galinsky 2016: 512; see also Section 4).

Oelkers' (2003) main claims concerning specific phonological properties are briefly summarized in the following.[10] The finding that female names have more syllables is explained in such a way that longer words are perceived as more melodious than shorter words. Euphony again is connected to central components of the female gender stereotype (according to Oelkers 2003: 144). A similar explanation applies to the two factors 'higher portion of vowels' and 'vowel final': vowels make a name more sonorous, which is usually perceived as sounding softer, and softness fits better to the female gender stereotype than to

10 Note that some aspects mentioned by Oelkers are firmly established in the literature on phonosemantics (e.g., the association of pitch, size, and gender – see, e.g., Ohala 1983, 1984) while others are not (e.g., the association of word length, euphony, and gender).

the male one.[11] As far as the position of the main stress is concerned, exoticism is the decisive characteristic: Since initial stress is characteristic of Germanic languages, male names represent the unmarked case more often, which, according to Oelkers (2003: 160), can be associated with the male gender stereotype. Female names deviate from the unmarked stress pattern more often and thus correspond to the female gender stereotype due to their markedness and exoticism. Finally, the distribution of the quality of the stressed vowel can be associated with the famous size-sound symbolism (also known as *frequency code*, cf., e.g., Hinton et al. 1994: 10; Ohala 1983, 1984). Oelkers (2003: 227) argues that sounds that have been shown to be associated with bigger size (e.g., dark vowels such as /a/, /o/, /u/) are connected with masculinity (for an early study on the mapping of size and sound cf. Sapir 1929). This size-sound mapping is one of the best studied relations in the field of sound symbolism. For instance, Knoeferle et al. (2017) have shown in a recent study which acoustic cues best characterize 'large' and 'small' sounding phonemes. Consistent with the predictions, their experiment has shown that size judgements were indeed higher for sounds with a higher F1 (which reflects the progressive opening of the jaw) and simultaneously a lower F2 (which increases with vowel frontness) and for sounds with a longer duration.

In sum, Oelkers (2003: 227) claims that German female names are characterized by prosodic-phonological features that make them sound softer and more melodious compared to male names. Additionally, female names sound rather 'exotic' compared to typical German prosodic-phonological structures. Male given names, on the other hand, sound 'harder', and have rather unremarkable structures (Oelkers 2003: 199). Thus, the prosodic-phonological structures of given names do not only determine the gender of a name but also transport stereotypical ideas of gender (as they exist in many other areas of life) according to Oelkers (2003).

Oelkers (2003: 227–228) assumes a non-arbitrary mapping between phonetic properties of speech sounds and their meaning – namely gender in the case of personal names. Hence, she argues for a case of sound symbolism where semantic concepts – such as femininity – that have no audible characteristics are coded phonologically. This case of sound symbolism has been called *synesthetic sound symbolism* by Hinton et al. (1994: 4–5). In the following, we adopt this term (although the more common use of the term *synesthetic* relates to sensory experience).

11 For a detailed discussion of "soft/hard sounding names" and the according gender stereotypes, see Slepian and Galinsky (2016) and the literature cited therein. Using the example of voiced vs. unvoiced initial phonemes in American (and Indian) names, they examine experimentally how the perception of sounds interacts with gender stereotypes.

The hypothesis that ‘fe/male sounds’ exist implies that the prosodic-phonological patterns that we find in German given names must be valid beyond (closely related) individual languages if not universally (Oelkers 2003: 228). Accordingly, Oelkers (2003: 227–228) explicitly argues against Cassidy et al.’s (1999) claim that “phonological cues to gender appear to be language specific and psychologically arbitrary”. On the basis of data from countries where English is the majority language also Cutler et al. (1990), Whissell (2001), and Pitcher et al. (2013) reason in favor of synesthetic sound symbolism.[12]

By contrast, Nübling (2018) (and similarly also Cassidy et al. 1999 and Hough 2000), argues against synesthetic sound symbolism and states that no sound as such is male or female – it is only due to convention that some phonological patterns are associated with the information [± female] (i.e., conventional sound symbolism in the sense of Hinton et al. 1994). Furthermore, she claims that correlations between a name’s prosodic-phonological structure and the assigned gender can be explained from a diachronic perspective: In German *-a* (sometimes also *-e*) has attained the morphological status of a gender suffix by deriving female names from many male names for centuries (cf. Example (1)) – the other direction is blocked in German.

(1) a. /ˈmaɐ.tin/$_{\text{MALE NAME}}$ → / maɐ.ˈtiː.na/$_{\text{FEMALE NAME}}$
b. /ˈkʁɪs.tjan/$_{\text{MALE NAME}}$ → /kʁɪs.ˈtjaː.nə/$_{\text{FEMALE NAME}}$

Thus, we are dealing with reanalyzes at the morphology/phonology interface: Morphologically (by suffixes such as *-a*) caused phonological effects (open final syllables, more syllables, shift to non-initial stress, etc.) result in specific sound patterns and are associated with femininity. By contrast, the sound structures of morphologically unmodified names (closed final syllables, fewer syllables, initial stress, etc.) are associated with masculinity. Hence, Nübling (2018) interprets correlations of gender and phonological structure as language/culture-specific conventions.

Against this backdrop, a study which takes different languages/cultures into account suggests itself. This is where our study takes its point of departure. In the next sections, we compare the relation of gender and phonological structures of names across different languages/cultures. By doing so, we want to assess whether the observed correlations of name phonology and gender are valid beyond single languages/cultures.

12 These scholars do not use the term *synesthetic sound symbolism* but their claims are in line with Hinton et al.’s (1994: 4–5) definition.

3 Empirical investigation: phonological cues to gender in 13 languages/cultures

3.1 Sample selection

In order to fill the outlined research gap, we analyzed two samples of popular given names. Sample 1 consists of names from linguistically rather disparate countries, namely China, France, Germany, Hungary, Israel, Japan, Poland, and Turkey.[13] These countries differ not only concerning their majority language but also with regard to gender marking on given names. All three types introduced in Section 2 are covered, i.e., the conventional type (which predominates, e.g., in Germany), the semantic type (which predominates, e.g., in Turkey), and countries where gender is often marked by formal means such as final *-a* for female names (e.g., in Poland).[14] With this sample, we want to find out whether there are language/culture-*overarching* phonological cues to gender.[15]

Sample 2, by contrast, consists of names from European countries where an Indo-European language is primarily spoken, namely Bulgaria, Denmark, England, France, Germany, Poland, Romania, and Spain.[16] This sample allows us to study possible language-family/culture-*specific* correlations of name structure and gender.

For each country, the 30 most popular names per gender were selected on the basis of recent (maximum five years old) and trustworthy statistics on the naming of newborns, such as those made available, for example, by public authorities.[17] A

---

13 Why we could not create a wider and more balanced sample will be discussed below.

14 We decided to include countries belonging to the latter two types in order to test whether we find phonological cues to gender in these countries beyond the obvious semantic and formal means for gender marking. For example, van de Weijer et al. (2020: 2) show for Chinese names that they are "generally selected on the basis of positive semantic characteristics but must also be euphonic and perceived as beautiful and well-balanced in writing".

15 Needless to say, with our sample we can only disprove the hypothesis that a pattern is language-overarching. If we find similarities across all languages/cultures in our sample, it does of course not follow that these patterns are universal.

16 The Indo-European sample contains three Germanic, three Romance, but only two Slavic languages. This is because both samples must have the same number of names for statistical purposes (that means eight Indo-European languages). For Slavic languages there are name statistics from Poland, Bulgaria, and Russia, but the available Russian names have not been included in the sample because the official names differ too much from the names actually used (see below).

17 All statistics used here differentiate between only two genders (male and female). Appendix A lists the sources of these statistics. The relevant parts of the statistics have been adopted by us country was only selected for the study if such statistics were available – which excluded some interesting areas/countries.[18]

Table 3 gives an overview of the composition of the two samples. (See also the link to the data samples on Zenodo in the Supplement below).

For all countries, the official (and thus registered) full forms of the names were used. This is not without problems because it can be assumed that, for example, a *Mevlüt* is sometimes called *Mevo* in Turkish. The problem of the unofficial

Table 3: Composition of heterogeneous and Indo-European name sample.

SAMPLE 1 (HETEROGENEOUS)				SAMPLE 2 (INDO-EUROPEAN)			
Country	Majority language	Language Family	N	Country	Majority language	Language Family	N
China	Mandarin	Sino-Tibetan Sinitic	60	Bulgaria	Bulgarian	Slavic	60
France	French	Indo-European Romance	60	Poland	Polish	Slavic	60
Germany	German	Indo-European Germanic	60	Romania	Romanian	Romance	60
Hungary	Hungarian	Uralic Finno-Ugric	60	Spain	Spanish	Romance	60
Israel[19]	Hebrew	Afro-Asiatic Semitic	60	France	French	Romance	60
Japan	Japanese	Japonic	60	Denmark	Danish	Germanic	60
Poland	Polish	Indo-European Slavic	60	England	English	Germanic	60
Turkey	Turkish	Turkic Common Turkic	60	Germany	German	Germanic	60
Total			480				480

unchanged. This means that compound names have been included as a whole, even if parts of them also occurred in the list as simplex word (e.g., *Hira* and *Hira Nur* in the Turkish subsample). Furthermore, several graphematical variants of a name may appear in our sample (e.g., <Hira Nur> and <Hiranur>). In addition, our sample includes some unisex names (e.g., *Ariel* in the Israeli subsample; cf. Muchnik 2017). These were not treated separately. An in-depth analysis of unisex names from a contrastive perspective would certainly be very interesting but would go beyond the scope of this paper. Appendix B contains the sample with all (transliterated) given names.

18 For example, we would have liked to include sub-Saharan countries to make our sample wider and more balanced.

19 For Israel only names of Jewish newborns are available. Therefore, names of Muslim children are not included in the sample, although they are in part more frequent.

shortening or change of the name applies to many languages/cultures. However, for our study we decided to choose the official names since reliable and representative information on unofficial names in the selected countries can hardly be gathered. In addition, one can assume that official names have at least a certain relevance in the countries at hand. Only Russia, for which it is generally known that the full forms recorded in official statistics are exceptionally rarely in use and the unofficial names have little in common with the registered ones (e.g., *Alexander → Sascha; Maria → Mascha*), was not included in the sample for exactly this reason.[20]

After sample composition, all 780 names were transcribed with the help of native speakers and were annotated for a range of prosodic-phonological features, such as the number of syllables, the average sonority, et cetera.[21] The transcription of the names and the annotation of factors such as sonority are the subject of the next section.

3.2 Transcription and annotation of the data

First, a few essential procedures regarding transcription have to be addressed. In general, for all names narrow phonetic transcriptions were made based on speech recordings or native speaker competence. This was always based on what our informants (of whom all are linguistically well informed) perceived as default pronunciation. Affricates and diphthongs were always transcribed and counted as a sequence of two segments each (e.g., *James* [**dʒɛi**ms] = six sounds).[22] Hiatuses were only counted as two syllables if they appear non-contracted as in the German female name *Mia* [ˈmiː.a]. Names that are usually contracted were transcribed accordingly (e.g., *Jul**ia*** [ˈjuːl.**ja**]). This method is arguable (see, e.g., Nübling 2009, who consistently counts hiatuses as two syllables). However, since the transcription used here treats female and male names equally, possible biases between the genders should be minimal.

20 At least in some languages/cultures there are differences between official and unofficial names with regard to phonological gender marking. For example, Nübling (2017) shows in a study on German that phonological gender marking is reduced in nicknames that are derived from first names (such as *Konni* < *Konrad*) in comparison to official first names. Hence, it would be interesting to compare official and unofficial Russian names with regard to phonological gender marking.

21 Note that three countries/languages appear in both samples.

22 This seems to be quite uncontroversial from a phonetic perspective. At the same time, it is disputed whether (and which) affricates and diphthongs should be analyzed as one phoneme or as two phonemes. Our decision does not imply anything concerning this (phonological) question.

After phonological transcription all names were annotated for prosodic-phonological features that have been mentioned in the literature on the correlation of name phonology and gender (Cassidy et al. 1999; Cutler et al. 1990; Lieberson and Mikelson 1995; Nübling 2009; Oelkers 2003; Pitcher et al. 2013; Slater and Feinman 1985; Whissell 2001). One factor refers to the entire name structure, namely the number of syllables. The other factors – namely sonority and the vowel quality – refer to specific structural positions within the name, such as the final sound.

Let us first come to the factor vowel quality. As mentioned in Section 2, the quality of vowels can give rise to different associations with regard to the size of a denotatum. Applied to gender, this means that smaller-sounding vowels correspond to a female gender stereotype and should therefore occur more frequently in female names. As Knoeferle et al. (2017) have shown the opening of the jaw as well as frontness play a significant role for the size-judgement. Accordingly, we made a distinction between palatal (e.g., [i], [e], [y]) and non-palatal vowels (e.g., [a], [o], [u]). Unlike, e.g., Oelkers (2003), who only considers the quality of the stressed vowel, we counted all palatal and non-palatal vowels in one name, but not the final sound. We excluded the final sound here because the name final position has been shown to be especially prominent with regard to gender coding in many (Indo-European) languages. It is often the most sonorous, but also 'large' sounding vowel /a/, which indicates femininity (cf., e.g., Kürschner 2018: 306; Lieberson and Mikelson 1995: 935; Nübling 2018: 244; Oelkers 2003: 195). This structural position should therefore be considered separately in order to avoid the possibility that the final sound obscures other more subtle phonological cues to gender.

The second factor that needs to be explained in more detail is sonority. As Nübling (2009: 81) notes in her analysis of German given names, it is too simplistic to assume a dichotomy between vowels and consonants and to put them in opposition to each other. It is well known that a continuum extends between vowels, which are highly sonorous, and consonants. This can be shown with the sonority scale in Figure 1 (Nübling 2009: 81; cf. also, e.g., Neef 2002; Vennemann 1982).

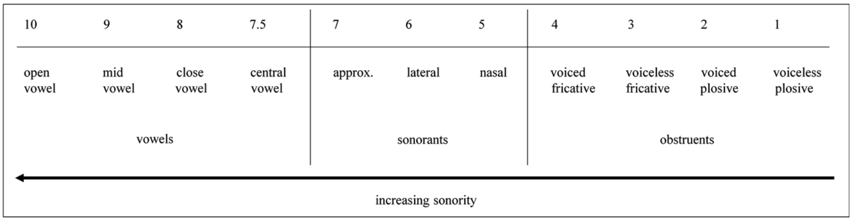

Figure 1: Sonority hierarchy and sonority values (cf. Nübling 2009: 81).

According to the sonority scale, vowels are most sonorous. Within the group of vowels open vowels (/a/) are more sonorous than mid vowels (e.g., /e/, /o/) and mid vowels are more sonorous than close vowels (e.g., /i/, /u/). In addition to considering such sonority differences within a group of sounds the scale also incorporates the fact that not only vowels at the top of the scale, but also consonantal sonorants (and voiced fricatives) are highly euphonic and give names a sound that is perceived as soft (cf. Nübling 2009: 81). The sonority decreases continuously towards the right end of the scale. Accordingly, voiceless plosives have the lowest sonority and thus the highest degree of consonantality. In order to operationalize the factor sonority, values – based on Nübling (2009) – have been assigned to all sounds. These values are shown in Figure 1 in the top row.[23] In accordance with this scale, numerical sonority values were assigned to every name as a whole and to special structural positions of the name, i.e., the initial and the final sound.[24]

Table 4 lists all coded factors and their possible characteristics as well as the predictions regarding a male or female gender cue (as discussed in Section 2).

Finally, it should be briefly mentioned why certain factors that have been proven to be relevant in individual languages are not (or: cannot be) taken into

Table 4: Relevant factors with possible characteristics and predictions regarding the gender cue.

Feature	Values	Gender cue
Number of syllables	1–5 syllables	Higher number of syllables = female gender cue
Number of non-palatal vowels (excluding final sound)	0–3	Higher number of non-palatal vowels = male gender cue
Sonority of first sound	1–10 points (on sonority scale)	Higher sonority = female gender cue
Average sonority (till final vowel)	1–10 points (on sonority scale)	Higher sonority = female gender cue
Sonority of final sound	1–10 points (on sonority scale)	Higher sonority = female gender cue

23 Note that many different scales have been proposed in the literature (cf., e.g., Neef 2002 for an overview). Most of these proposals share some basic assumptions, but they also differ in certain aspects, such as the concrete number of classes that should be distinguished. We have opted for a rather fine-grained classification here.

24 When calculating the average sonority, we excluded the final vowel (and all subsequent sounds) to avoid that the tendency of female names to end in /a/ obscures more subtle differences (see above).

account in our crosslinguistic study. These are stress position, the ratio of open versus closed syllables, and the number of consonant clusters. The main stress – which has been shown to be a significant factor in some Germanic languages – was not considered here due to different types of lexical stress realization and different stress positions in the languages of our sample. Thus, the stress types are not comparable and a derivation of hypotheses regarding gender cues seems questionable.[25]

The ratio of open and closed syllables in a name depends on the respective language: are we rather dealing with a syllable or with a word language (cf., e.g., Auer 1993; Caro Reina and Szczepaniak 2014)? German and English – languages in which the factor has been shown to be relevant – have traits of word languages, i.e., the prosodic domain of the phonological word is central. However, there are also syllable languages in our sample, i.e., languages in which the prosodic domain of the syllable is central, which, among other things, is reflected in an optimized CV-structure (e.g., Japanese). Therefore, it is a priori clear that the ratio of open and closed syllables cannot be relevant across all languages included.[26] The same holds for the number of consonant clusters per name.

In the next section, the crosslinguistic relevance of the factors discussed here will be examined in two multifactorial analyses.

3.3 Binary logistic regressions

3.3.1 Heterogeneous sample

In the previous section, various features were discussed that have been said to be typical for female or male names in German and English. As shown, there are some striking parallels between these two languages. However, this is not particularly surprising as English and German are geographically close and genetically related. In order to examine whether these factors are also influential when it comes to a sample of names from linguistically rather disparate countries, the data were analyzed with a binary logistic regression model.[27] This will be described in this section.

---

25 See, e.g., van de Weijer et al. (2020), who investigate the role of tonal sound symbolism in Chinese names.

26 Also, the sonority of the final sound depends to some extent on the language at hand: in syllable languages (such as Japanese) it is very likely that a name ends in a vowel (and thus in a comparably sonorous sound). However, there might still be systematic differences between the genders because of sonority differences within the group of vowels (cf. Figure 1).

27 The software *R* was used (R Core Team 2019) and the package *rms* (Harrell Jr 2019).

In the first multifactorial model the variables NUMBER OF SYLLABLES, NUMBER OF NON-PALATAL VOWELS (EXCLUDING FINAL SOUND), SONORITY OF FIRST SOUND, AVERAGE SONORITY (TILL FINAL VOWEL), and SONORITY OF FINAL SOUND were integrated. Hence, we started with a maximal model. Here (and in all following models), GENDER is the dependent variable. Independent variables that have no significant influence were removed from the model in a step-down procedure based on the Akaike Information Criterion (AIC). By this procedure, two out of five variables – i.e., NUMBER OF SYLLABLES and SONORITY OF FIRST SOUND – were removed as they do not improve the quality of the model. The specification of the final model is given in (2).

(2) GENDER ~ non-palatal vowels without final sound + average sonority till final V + sonority final sound

Table 5 gives an overview of the significance of the predictors. In addition, the effect sizes (log odds), standard errors, Wald Z, and p-values are listed. The concordance index shows that the model discriminates acceptably: $C = 0.784$ (cf., e.g., Hosmer and Lemeshow 2000: 162).[28]

In contrast to *p*-values, effect coefficients (here we used log odds) do not indicate whether an influence is significant, but its direction and its strength. Log odds are centered around 0 and reach from +Infinity to –Infinity. Thus, a coefficient of 0 would mean that there is no difference between the levels of the predictor with regard to the choice between female and male names. Coefficients greater than 0 mean that the probability that the respective name is a male name is greater than the probability of a female name.[29] For example, NUMBER OF NON-PALATAL VOWELS (EXCLUDING FINAL SOUND) has a positive coefficient and thus an

Table 5: Predictors in the minimal adequate logistic regression model for name gender (reference level = female).

	Coeff.	S.E.	Wald *Z*	Pr(>|*Z*|)
intercept	3.0712	0.6202	4.95	<0.0001
non.palatal_vowels_without_final_sound	0.5242	0.1491	3.52	0.0004
average_sonority_till_final_V	−0.2007	0.0956	−2.10	0.0357
sonority_final_sound	−0.3411	0.0412	−8.27	<0.0001

28 Nagelkerke pseudo-R^2 (which is another measure for the goodness of fit) = 0.300.

29 For an explanation of the relationship between log odds, odds ratios, and probabilities see Gries (2013: 299–301).

increase in the number of non-palatal vowels (excluding the final sound) increases the odds for a male name.[30]

Based on the included variables, the model predicts the gender of 341 names (i.e., 71 % of the tokens) correctly (cf. Table 6).

Table 6: Confusion matrix (heterogeneous sample).

Actual	Predicted		Correctly predicted gender
	f	m	
f	176	64	73.3%
m	75	165	68.8%
			71.0%

Altogether, the coefficients confirm our hypotheses: Female names have a more sonorous final sound, a higher average sonority (till final vowel), and fewer non-palatal vowels than male names. To test whether these significant predictors are relevant for all languages/cultures or whether the significance is based on high values in only some languages/cultures we checked possible interactions with the factor COUNTRY in a second step. In order to do so, we compared a first regression model where we included the factor COUNTRY but no interaction term with regression models respectively including interaction terms for COUNTRY*NUMBER OF NON-PALATAL VOWELS (EXCLUDING FINAL SOUND), COUNTRY*AVERAGE SONORITY (TILL FINAL VOWEL), and COUNTRY*SONORITY OF FINAL SOUND. Then, we compared the model with no interaction term with the models with an interaction term via ANOVA and tested whether the interaction term improved the model significantly. This was the case for the sonority-factors. Especially, the factor SONORITY OF FINAL SOUND is highly language/culture-dependent (3):

(3) ANOVA 1
Model 1: Gender ~ sonority_final_sound + non.palatal_vowels_without_final_sound + average_sonority_till_final_V + country
Model 2: Gender ~ sonority_final_sound * country + non.palatal_vowels_without_final_sound + average_sonority_till_final_V

Resid	Df	Resid. Dev	Df	Deviance	Pr(>Chi)
1	469	530.41			
2	462	453.09	7	77.319	**4.845e-14** ***

30 All VIFs are below 2. This means that there is no critical multicollineartity in our model. In addition, bootstrapping was used to test whether the model overfits the data, which is not the case.

As Table 7 shows, the gender differences regarding the sonority of the final sound are particularly high in European countries. Contrarily, in Non-European countries, such as China and Japan, the final sound seems to be irrelevant with regard to gender.[31]

In contrast to the sonority of the final sound, the factor NUMBER OF NON-PALATAL VOWELS (EXCLUDING FINAL SOUND) seems to be language-/culture-overarching as the comparison of the model with and without interaction term shows (4).

(4) ANOVA 2
Model 1: Gender ~ sonority_final_sound + non.palatal_ vowels_without_final_sound + average_sonority_till_final_V + country
Model 2: Gender ~ sonority_final_sound + non.palatal_ vowels_without_final_sound * country + average_sonority_till_final_V

Resid.	Df	Resid. Dev	Df	Deviance	Pr(>Chi)
1	469	530.41			
2	462	524.78	7	5.6339	**0.5831**

This result can be illustrated by the monofactorial analysis given in Table 8: In all countries in our sample male names have on average more non-palatal vowels (if one ignores the final sound) and the differences between the countries are much weaker compared to Table 7.

Table 7: Gender differences in the sonority of final sound per country.

	Median		
	Female	**Male**	**Differences**
Poland	10	4	6
Hungary	10	4.5	5.5
Germany	10	5	5
Turkey	8	4	4
Israel	9	5	4
France	9	6	3
Japan	9.5	9	0.5
China	5	5	0

31 For Japan, this can partly be explained with constraints on syllable structures – all Japanese names in our sample end in a vowel. However, a distinction would still be possible because of sonority differences within the group of vowels (see also Footnote 26).

Table 8: Gender differences in the number of non-palatal vowels per country.

	Mean		
	Female	**Male**	**Differences**
Turkey	0.6	1.2	0.6
Hungary	0.7	1.2	0.5
Germany	0.7	1.0	0.3
China	0.8	1.1	0.3
Israel	1.0	1.3	0.3
Japan	1.0	1.3	0.3
Poland	1.0	1.3	0.3
France	0.6	0.8	0.2

The results described in this section showed that there are three groups of variables:

(i) variables that did not have a significant effect at all (NUMBER OF SYLLABLES and SONORITY OF FIRST SOUND),

(ii) variables that have an effect but whose significance is based on only some countries in the sample (AVERAGE SONORITY TILL FINAL VOWEL and SONORITY OF FINAL SOUND), and

(iii) one variable that proved to be important across all countries: NUMBER OF NON-PALATAL VOWELS (EXCLUDING FINAL SOUND).

As there is only one variable in group (iii) it is not surprising that the model can only explain a moderate proportion of variance (Nagelkerke Pseudo-$R^2 = 0.300$) – the vast majority of variables is irrelevant for some countries in the sample. In the next section, we test whether the model quality improves if a linguistically more homogeneous set of countries is analyzed. In order to do so we compiled a sample of names from countries where an Indo-European language is the majority language. Comparing the resulting model and the model described in this section will clarify whether we are dealing with language family internal tendencies.

3.3.2 Indo-European sample

The analysis of the second sample mostly matches the procedure described in Section 3.3.1. To avoid repetitions, not all steps are explained in detail again. We started with a maximum model and removed all variables that have no significant influence in a step-down procedure based on the Akaike Information Criterion (AIC). Only SONORITY OF FIRST SOUND was excluded. The specification of the final model is given in (5):

(5) GENDER ~ number of syllables + non-palatal vowels without final sound + average sonority till final V + sonority final sound

Table 9 contains relevant values resulting from the regression analysis.

The model closely resembles the one described in Section 3.3.1. The only additional predictor that reached significance is NUMBER OF SYLLABLES. The coefficients of the other variables match their counterparts from the analysis of the other sample with regard to their direction. However, the sonority variables deviate more strongly from 0 which means that they have a stronger impact on the dependent variable. The additional variable and the higher impact of two variables are reflected in a better model quality, which is shown in Table 10.[32] According to Hosmer and Lemeshow (2000: 162) a *C* index above 0.8 indicates "excellent discrimination".[33]

Table 9: Predictors in the minimal adequate logistic regression model for name gender (reference level = female).

	Coef	S.E.	Wald Z	Pr(>|Z|)
Intercept	7.7929	0.9718	8.02	<0.0001
number_of_syllables	−0.4664	0.1857	−2.51	0.0120
non.palatal_vowels_without_final_sound	0.4998	0.1910	2.62	0.0089
average_sonority_till_final_V	−0.5368	0.1288	−4.17	<0.0001
sonority_final_sound	−0.5207	0.0565	−9.22	<0.0001

Table 10: Comparison of model qualities (Heterogeneous vs. Indo-European sample).

	Heterogeneous sample	Indo-European sample
Nagelkerke Pseudo-R^2	0.300	0.524
Index of discrimination (*C*)	0.784	0.885

32 Again, there is no critical multicollinearity: all VIFs < 2. Additionally, bootstrapping showed that the model does not overfit the data.

33 We tested all possible interactions with COUNTRY using the method described in Section 3.3.1. Only COUNTRY*SONORITY OF FINAL SOUND proved to be significant ($p < 0.001$***). This is because the effect is stronger for countries such as Poland, Romania, and Bulgaria where (almost) all female names end in *a*. Accordingly, Poland, Romania, and Bulgaria can be classified as formal gender marking languages (cf. Section 1.1). In the other countries there is a similar but weaker correlation.

Table 11: Confusion matrix (Indo-European sample).

Actual	Predicted		Correctly predicted gender
	f	m	
f	211	29	87.9%
m	61	179	74.5%
			81.3%

Given that both samples comprise the same number of names these differences demonstrate that our variables can explain the variance in the comparatively homogenous Indo-European sample better than the variance in the much more heterogeneous sample analyzed in the previous section. This can also be shown by the confusion matrix in Table 11: The gender of more names is predicted correctly compared to the model described in the previous section (cf. Table 6 above). The following section discusses what our empirical results imply with regard to sound symbolism in the domain of phonological gender marking on names.

4 Discussion – synesthetic versus conventional sound symbolism

As we have outlined in Section 1, there is no consensus in the literature whether correlations of phonological features and the name bearer's gender are language/culture-specific or crosslinguistically valid. Our results indicate that the majority of such correlations are not valid beyond related languages in geographically close countries. Using the example of the final sound's sonority, we have shown that there are important differences between the countries in our first sample. While the sonority of the final sound is a phonological cue to gender in some countries (e.g., Poland, Hungary, Germany), it is completely irrelevant in others (e.g., Japan, China). The way countries cluster in our analysis suggests that phonological gender marking strategies coincide with other cultural, geographical, and linguistic classifications. For example, China and Japan deviate from European countries with regard to the relevance of the final sound's sonority. The analysis of our second sample supports this idea: Our variables explain the variance in the

Indo-European sample much better than the variance in the heterogeneous sample. There are two explanations for this observation. Firstly, all countries in the second sample are not only closely related due to their majority language but they do also share cultural traditions and borrowing played and plays an important role – which results in overlapping/related onomasticons (cf., e.g., Gerhards 2010: 158). Hence, there are several etymologically related names in this sample. For instance, variants of *Alexander* (which is an extreme example) can be found in the sub-samples from Bulgaria, Denmark, England, Germany, Romania, and Spain (see Appendices A and B). Although these names are pronounced slightly differently in the respective countries, they make the sample more homogenous and this partly explains the difference between the two models. Secondly – and more importantly with regard to our research question – the names in the second sample are also more homogeneous with regard to the structure of the names in general and with regard to gender differences in particular. For example, final *a* is much more common for female names than for male. This holds irrespective of etymologically related names and applies also to female names without related variants in the sample (such as Bulgarian *Iwaila*).

Overall, it is not very surprising that our variables explain the variance in the Indo-European sample better than the variance in the heterogeneous sample since our variables are taken from studies on phonological gender marking on names in selected Indo-European languages. Still, this difference contradicts the idea that all or many correlations of phonological properties and the name bearer's gender discovered by, e.g., Oelkers (2003) for German, are non-arbitrary and potentially universal.

However, we discovered a solid correlation of gender and the name's sonority (regarding the final sound and the rest of name) in certain countries. In the Indo-European sample the number of syllables also had a significant impact. These correlations seem to be firmly established but linguistically and geographically limited. Therefore, they seem to qualify as cases of conventional sound symbolism. Originally, the diverging phonological properties of female and male names were motivated by, e.g., specific word/name formation patterns. For example, the formation of female names by -*a* suffixation led to numerous names with comparatively more syllables and a maximally sonorous final sound (cf. Section 2). Subsequently, the frequent co-occurrence of comparatively many syllables and/or a sonorous final sound on the one hand and the semantic property [+female] on the other hand resulted in associations of phonology and semantics in this domain. Since the word formation pattern that triggered this development traces to Latin and is only found in European languages these associations are not universal (cf. Cassidy et al. 1999: 362; Hough 2000: 6; Nübling 2009: 100, 2018: 242; Oelkers 2003: 144–145).

Though, our analyses also revealed that one correlation is valid across all countries in our samples: the number of non-palatal vowels (excluding the final sound). Thus, this variable qualifies as a candidate for synesthetic sound symbolism. Indeed, a corresponding explanation is obvious. One of the most robust findings in the literature on sound symbolism is the fact that non-palatal vowels are associated with bigger size (cf., e.g., Elsen 2017: 492) and bigger size is often associated with the masculine gender stereotype. Against this backdrop, it seems to be explicable that male names have more non-palatal vowels in word-initial and word-medial position than female names across all countries in our samples. The claim that synesthetic sound symbolism is at play here (as put forward, e.g., by Cutler et al. 1990; Oelkers 2003: 227–228; Pitcher et al. 2013; van de Weijer et al. 2020; Whissell 2001) therefore seems to be justifiable from a crosslinguistic perspective in this particular case. However, synesthetic sound symbolism has only an indirect influence (see also van de Weijer et al. 2020: 16). A matching gender stereotype is also necessary to establish an association of the phonological (e.g., word medial [i]) and the semantic property (e.g., female name bearer).[34] Figure 2 depicts this mechanism.

In contrast, conventional sound symbolism is based on the frequent co-occurrence of a phonological (e.g., word final [a]) and a semantic property (e.g., female name bearer) and does not involve a second semantic concept. This is depicted in Figure 3.

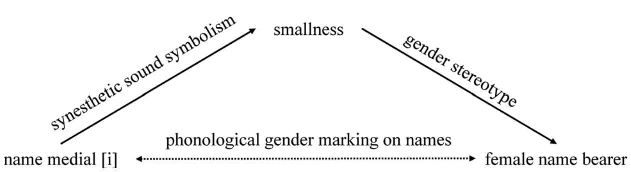

Figure 2: Phonological gender marking via sound symbolism and gender stereotypes.

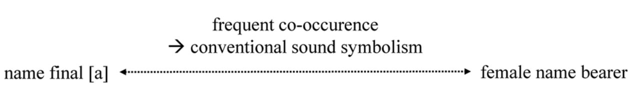

Figure 3: Phonological gender marking via conventional sound symbolism.

34 Differences between cultures with regard to gender stereotypes might also be a reason why certain correlations of name phonology and gender are limited to certain regions (see also Slepian and Galinsky 2016: 523).

Interestingly, synesthetic and conventional sound symbolism conflict in the Indo-European languages. This holds particularly for the sound [a]. On the one hand sonorous word final sounds (such as [a]) are associated with female names (due to word formation patterns; conventional sound symbolism) while on the other hand non-palatal vowels (such as [a]) are associated with male names (via association with bigger size; synesthetic sound symbolism). Hence, the connotation this sound triggers depends on its position in the name: word final position triggers a female association; word initial or word-medial position triggers a male association. In word final position the strong association of sonorous sound and female gender resulting from the specific word formation pattern overrides the general tendency to associate non-palatal vowels with bigger size and masculinity.

Both types of phonological gender marking seem to be important for name choices. Certain names are obviously perceived as more appropriate with regard to the gender of a child due to phonological properties. This reinforces correlations of gender and the phonology of the most popular names (which we analyzed here).

With regard to synesthetic sound symbolism the attractiveness of the associated meaning (e.g., smallness) also plays a role. While it might be the case that smaller sounding names are perceived as appropriate for women due to sexual size dimorphism, this does not necessarily mean that smallness is considered a desirable characteristic for women. This would diminish the attractiveness of the particular structure and would run contrary to a corresponding correlation of phonology and gender. As there are certainly cultures where this is the case it is likely that exceptions to the correlation of palatal vowels and feminine name bearers can be found (Pitcher et al. 2013: 5).

5 Conclusion

In our study, we addressed the correlation of name phonology and gender from a crosslinguistic perspective. Considering name giving practices in several (linguistically disparate) countries allows us to re-evaluate claims on how this topic relates to sound symbolism. Our empirical results yielded that the majority of the phonological patterns that have been shown to correlate with the gender of the name bearer in German and English (cf. Nübling 2009, 2012, 2018; Barry and Harper 1995; Cutler et al. 1990; Lieberson and Bell 1992; Oelkers 2003, 2004; Slater and Feinman 1985; Whissell 2001) are not associated with a certain gender universally. Instead, such correlations are limited to certain regions / language families. This substantiates Nübling's (2018) claim that no sound as such is male or female and that specific correlations can be explained with reference to the diachrony of individual languages or language families. Also, cultural aspects (which lead to partly shared / related onomasticons) play a role here.

Even so, we also detected a phonological variable that correlates with gender in all countries studied: the number of non-palatal vowels (excluding the final sound). Here, an explanation based on synesthetic sound symbolism seems appropriate which supports Oelkers' claim that correlations of name phonology and gender are not necessarily arbitrary. Thus, there is evidence for different kinds of sound symbolism in the domain of onymic gender marking. However, the instances of conventional sound symbolism outnumber synesthetic sound symbolism.

In addition to these results, our study has also revealed some directions for future research. For example, it would be particularly interesting to examine the relevance of non-palatal vowels experimentally. In such a study, one could ask participants with disparate linguistic backgrounds to assign "pseudonames" to female or male name bearers (cf. Cassidy at al. 1999 and Oelkers 2003 for similar studies with English and German speaking participants). Furthermore, it would be fruitful to analyze language families other than Indo-European more closely. Certainly, there are means of onymic gender marking that have not been considered here because it is obvious that they are language (family) specific. Another interesting topic would be the question how language and culture contact influence name structures and onymic gender marking. For example, it seems to be the case that Japanese names approach European patterns also in terms of phonological gender marking on names due to an increase of female names ending in *a*, which has been quite uncommon for Japanese female names in the past. In conclusion, there are still a number of open questions and issues that are worth addressing in the field of phonological gender marking on names. However, by adopting a crosslinguistic perspective, we hope to have shed some new light on this topic.

Acknowledgments: Many thanks are due to Beijia Chen, Semra Kızılkaya, Konrad Mazur, Malka Muchnik, Eva Meier, Takuma and Vera Melber, Nina Nikulova, Milena Osterloh, and Britta Stuhl for their help with the transcription of the names.
Supplement: The data samples are available for viewing at https://doi.org/10.5281/zenodo.4911448.

Appendix A

Table A1 contains information on the sources of our name samples for all countries considered here. If available, a URL is given (including the date the material was accessed). The year from which the statistics originate is also given.

Table A1: Sources of our name samples

Country	Year	Source	Access	Comment
Bulgaria	2017	<www.nsi.bg/en/content/15852/%D0%BF%D1%80%D0%B5%D1%81%D1%81%D1%8A%D0%BE%D0%B1%D1%89%D0%B5%D0%BD%D0%B8%D0%B5/names-bulgaria-2017-preliminary-data>	03/20/2019	–
China	2015	<https://www.qimingtong.com/article/0?fp=er#article	06/19/2018	based on a representative sample collected by the company *QiMing-Tong* in cooperation with the Tsinghua University Beijing
Denmark	2017	<www.dst.dk/en/Statistik/emner/befolkning-og-valg/navne/navne-til-nyfoedte#>	02/14/2019	–
England	2017	<https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths>	03/11/2019	only data from England were considered for our study
France	2016	<www.insee.fr/fr/statistiques/2540004>	07/09/2018	–
Germany	2017	<www.beliebte-vornamen.de/jahrgang/j2017/top-500-2017>	07/09/2018	based on statistics provided by selected regional register offices
Hungary	2016	<www.nyilvantarto.hu/hu/statisztikak>	07/24/2018	–
Israel	2016	No published data available. Data were provided by the Central Israeli Bureau of Statistics on request via e-mail.	–	–
Japan	2017	<www.meijiyasuda.co.jp/sp/enjoy/ranking/index.html#/year/2017y/2>	07/06/2018	based on statistics provided by an insurance company

Table A1: (continued)

Country	Year	Source	Access	Comment
Poland	2017	<www.gov.pl/web/cyfryzacja/statystyka-imion-za-2017-rok-podsumowanie>	02/21/2019	–
Romania	2009	<http://numedecopii.com/info.asp?id=nume-fete-2009>	02/14/2019	Data provided by the Romanian Ministry of Administration and Interior
Spain	2017	<www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177009&menu=resultados&secc=1254736195498&idp=1254734710990>	03/11/2019	–
Turkey	2014	Türkiye İstatistik Kurumu Matbaası [Turkish Statistical Institute]. 2014. İstatistiklerle Çocuk / Statistics On Child 2014. Ankara: Türkiye İstatistik Kurumu Matbaası. ISBN: 978-975-19-6341-3.	–	–

Appendix B

Female			Male		
Bulgaria					
01 Viktoriya	11 Elena	21 Stefani	31 Aleksandar	41 Kristiyan	51 Samuil
02 Mariya	12 Michaela	22 Magdalena	32 Georgi	42 Viktor	52 Aleks
03 Nikol	13 Teodora	23 Karina	33 Martin	43 Teodor	53 Hristo
04 Raya	14 Bozhidara	24 Anna	34 Iwan	44 Bozhidar	54 Atanas
05 Aleksandra	15 Gergana	25 Niya	35 Dimitar	45 Simeon	55 Iordan
06 Sofiya	16 Ema	26 Vanessa	36 Nikola	46 Stefan	56 David
07 Dariya	17 Siyana	27 Tzwetelina	37 Daniel	47 Petar	57 Ioan
08 Simona	18 Iwaila	28 Ewa	38 Boris	48 Iwailo	58 Wassil
09 Gebriela	19 Kalina	29 Darina	39 Kaloyan	49 Angel	59 Kristian
10 Ioana	20 Monika	30 Plamena	40 Nikolai	50 Michail	60 Borislav
China					
61 Zǐ·xuān	71 Jǐn·xuān	81 Mèng·qí	91 Hào·rán	101 Zǐ·muò	111 Jùn·jié
62 Zǐ·hán	72 Sī·hán	82 Yǜ·hán	92 Zǐ·xuān	102 Bó·wén	112 Zhì·yuǎn
63 Shī·hán	73 Kě·xīn	83 Zǐ·méng	93 Hào·xuān	103 Yī·nuò	113 Tiān·yòu
64 Kě·xīn	74 Zǐ·xuān	84 Zhǐ·xuān	94 Yǚ·xuān	104 Zǐ·hán	114 Míng·xuān
65 Yī·nuò	75 Ruò·xī	85 Yì·hán	95 Hào·yǚ	105 Zǐ·ruì	115 Zǐ·hán
66 Yǚ·xuān	76 Zǐ·xuān	86 Ruò·xuān	96 Zǐ·ruì	106 Ruì	116 Jùn·háo
67 Xīn·yí	77 Yǚ·hán	87 Yī·yī	97 Zǐ·xuān	107 Yǚ·zé	117 Hào·rán
68 Zǐ·hán	78 Yǚ·tóng	88 Ruò·xī	98 Hào·xuān	108 Míng·xuān	118 Yī·míng
69 Chén·xī	79 Xīn·yán	89 Yì·xīn	99 Jùn·xī	109 Yǚ·háng	119 Hào·yǚ
70 Zǐ·xuān	80 Ruò·xī	90 Shī·qí	100 Zǐ·háo	110 Zǐ·háo	120 Zǐ·chén

(continued)

Female			Male		
Denmark					
121 Ida	131 Laura	141 Aya	151 William	161 Valdemar	171 Alexander
122 Emma	132 Nora	142 Sofie	152 Noah	162 Elias	172 Villads
123 Sofia	133 Clara	143 Ellen	153 Oscar	163 Magnus	173 Christian
124 Ella	134 Karla	144 Lily	154 Lucas	164 Aksel	174 Johan
125 Freja	135 Isabella	145 Mathilde	155 Carl	165 Frederik	175 Adam
126 Josefine	136 Olivia	146 Maja	156 Victor	166 Felix	176 Arthur
127 Alma	137 Lærke	147 Frida	157 Oliver	167 Elliot	177 Liam
128 Alberte	138 Victoria	148 Emilie	158 Alfred	168 August	178 Theo
129 Anna	139 Mille	149 Marie	159 Malthe	169 Anton	179 Albert
130 Agnes	140 Luna	150 Esther	160 Emil	170 Nohr	180 Mikkel
England					
181 Olivia	191 Lily	201 Phoebe	211 Oliver	221 William	231 Alexander
182 Amelia	192 Sophia	202 Sienna	212 Harry	222 Henry	232 Edward
183 Isla	193 Grace	203 Evelyn	213 George	223 Alfie	233 Theo
184 Ava	194 Evie	204 Isabelle	214 Noah	224 Thomas	234 Isaac
185 Emily	195 Jessica	205 Ivy	215 Jack	225 Joshua	235 Lucas
186 Isabella	196 Sophie	206 Matilda	216 Jacob	226 Freddie	236 Ethan
187 Mia	197 Alice	207 Willow	217 Muhammad	227 James	237 Max
188 Poppy	198 Florence	208 Elsie	218 Leo	228 Arthur	238 Joseph
189 Ella	199 Daisy	209 Chloe	219 Oscar	229 Archie	239 Samuel
190 Charlotte	200 Freya	210 Scarlett	220 Charlie	230 Logan	240 Mohammed
France					
241 Emma	251 Camille	261 Julia	271 Gabriel	281 Nathan	291 Timéo
242 Louise	252 Léna	262 Romane	272 Jules	282 Arthur	292 Théo
243 Jade	253 Rose	263 Jeanne	273 Adam	283 Paul	293 Mohamed
244 Alice	254 Inès	264 Eva	274 Lucas	284 Nolan	294 Aaron

(continued)

Female			Male		
245 Chloé	255 Anna	265 Lou	275 Louis	285 Maël	295 Mathis
246 Mila	256 Sarah	266 Charlotte	276 Raphaël	286 Sacha	296 Axel
247 Léa	257 Zoé	267 Louna	277 Hugo	287 Tom	297 Antoine
248 Lina	258 Juliette	268 Mia	278 Léo	288 Noah	298 Victor
249 Manon	259 Ambre	269 Nina	279 Ethan	289 Enzo	299 Maxime
250 Lola	260 Lucie	270 Clara	280 Liam	290 Gabin	300 Clément
Germany					
301 Emma	311 Leni	321 Sophie	331 Ben	341 Henry	351 Liam
302 Hannah	312 Clara	322 Charlotte	332 Jonas	342 Maximilian	352 Moritz
303 Mia	313 Lena	323 Ida	333 Leon	343 Luca	353 Julian
304 Sofia	314 Luisa	324 Lilly	334 Paul	344 Oskar	354 Leo
305 Emilia	315 Leonie	325 Laura	335 Finn	345 Emil	355 David
306 Lina	316 Amelie	326 Maja	336 Noah	346 Anton	356 Alexander
307 Anna	317 Emily	327 Mathilda	337 Elias	347 Max	357 Milan
308 Marie	318 Johanna	328 Lara	338 Luis	348 Theo	358 Philipp
309 Mila	319 Ella	329 Frieda	339 Felix	349 Jakob	359 Niklas
310 Lea	320 Nele	330 Lia	340 Lukas	350 Matteo	360 Carl
Hungary					
361 Hanna	371 Nóra	381 Dóra	391 Bence	401 Zalán	411 Gergő
362 Anna	372 Maja	382 Sára	392 Máté	402 Áron	412 Benett
363 Jázmin	373 Fanni	383 Csenge	393 Levente	403 Balázs	413 Bálint
364 Zsófia	374 Laura	384 Petra	394 Dominik	404 Kristóf	414 Márk
365 Zoé	375 Dorina	385 Noémi	395 Marcell	405 Péter	415 Zoltán
366 Lili	376 Lilla	386 Eszter	396 Dávid	406 Botond	416 András
367 Boglárka	377 Gréta	387 Mira	397 Ádám	407 Olivér	417Attila
368 Luca	378 Izabella	388 Flóra	398 Noel	408 László	418 Márton
369 Emma	379 Viktória	389 Zselyke	399 Dániel	409 Zsombor	419 Benedek

(continued)

Female			Male		
370 Léna	380 Réka	390 Liza	400 Milán	410 Tamás	420 Gábor
Israel					
421 Tamar	431 Esther	441 Rachel	451 Noam	461 Yehuda	471 Raphael
422 Noa	432 Lia	442 Noya	452 David	462 Abraham	472 Shmuel
423 Avigail	433 Hanna	443 Miriam	453 Uri	463 Lavi	473 Ido
424 Maya	434 Rivka	444 Ema	454 Ariel	464 Israel	474 Michael
425 Yael	435 Roni	445 Ruth	455 Yosef	465 Itamar	475 Yitzhak
426 Adel	436 Romi	446 Yuval	456 Eitan	466 Omer	476 Chaim
427 Shira	437 Hodaya	447Alma	457 Daniel	467 Yonatan	477 Harel
428 Sarah	438 Ela	448 Halel	458 Itai	468 Yair	478 Shimon
429 Ayala	439 Michal	449 Ariel	459 Yehonatan	469 Yaakov	479 Amit
430 Talya	440 Chaya	450 Tahal	460 Moshe	470 Eliah	480 Alon
Japan					
481 Sakura	491 Honoka	501 Sara	511 Haruto	521 Yuusei	531 Yuuki
482 Yui	492 Riko	502 Hina	512 Souta	522 Kanata	532 Sora
483 Akari	493 Mio	503 Iroha	513 Yuuto	523 Souma	533 Yuito
484 Mei	494 Saki	504 Kaho	514 Haruki	524 Aoi	534 Itsuki
485 Hana	495 Miyu	505 Hinata	515 Riku	525 Kaito	535 Kouta
486 Sana	496 Ichika	506 Hikari	516 Sousuke	526 Hayato	536 Haruma
487 Rio	497 Riso	507 Haruka	517 Minato	527 Haru	537 Hiroto
488 Himari	498 Ema	508 Kanna	518 Aoto	528 Asaki	538 Ayato
489 Koaru	499 Tsumugi	509 Momoka	519 Hinata	529 Yuuma	539 Ryou
490 Aoi	500 Yuina	510 Yua	520 Kouki	530 Rikuto	540 Eito
Poland					
541 Julia	551 Oliwia	561 Nadia	571 Antoni	581 Adam	591 Bartosz
542 Zuzanna	552 Natalia	562 Marcelina	572 Jakub	582 Michał	592 Maksymilian
543 Zofia	553 Wiktoria	563 Gabriela	573 Jan	583 Marcel	593 Miłosz

(continued)

Female			Male		
544 Lena	554 Emilia	564 Michalina	574 Szymon	584 Stanisław	594 Tymon
545 Maja	555 Antonina	565 Kornelia	574 Franciszek	585 Wiktor	595 Oliwier
546 Hanna	556 Laura	566 Nikola	575 Filip	586 Piotr	596 Alan
547 Amelia	557 Pola	567 Helena	576 Aleksander	587 Igor	597 Ignacy
548 Alicja	558 Iga	568 Milena	578 Mikołaj	588 Leon	598 Tymoteusz
549 Maria	559 Anna	569 Martyna	579 Wojciech	589 Nikodem	599 Oskar
550 Aleksandra	560 Liliana	570 Jagoda	580 Kacper	590 Mateusz	600 Dawid
Romania					
601 Maria	611 Mihaela	621 Larisa	631 Andrei	641 Florin	651 Eduard
602 Andreea	612 Bianca	622 Sara	632 Alexandru	642 Darius	652 Mario
603 Elena	613 Georgiana	623 Daniela	633 Ionuț	643 Denis	653 Matei
604 Ioana	614 Nicoleta	624 Miruna	634 Gabriel	644 Constantin	654 Bogdan
605 Alexandra	615 Cristina	625 Roxana	635 Ștefan	645 Robert	655 Răzvan
606 Ana	616 Teodora	626 Valentina	636 Cristian	646 Adrian	656 Rareș
607 Denisa	617 Diana	627 Rebeca	637 Mihai	647 Ioan	657 Valentin
608 Gabriela	618 Alexia	628 Raluca	638 David	648 Sebastian	658 Cosmin
609 Ștefania	619 Ionela	629 Sofia	639 Daniel	649 Luca	659 Marius
610 Daria	620 Florentina	630 Mădălina	640 Marian	650 George	660 Nicolae
Spain					
661 Lucia	671 Carla	681 Olivia	691 Lucas	701 Leo	711 Carlos
662 Sofia	672 Sara	682 Elena	692 Hugo	702 David	712 Sergio
663 Maria	673 Noa	683 Adriana	693 Martin	703 Mario	713 Marc
664 Martina	674 Carmen	684 Laia	694 Daniel	704 Diego	714 Antonio
665 Paula	675 Claudia	685 Vega	695 Pablo	705 Javier	715 Bruno
666 Julia	676 Valentina	686 Vera	696 Alejandro	706 Enzo	716 Miguel
667 Daniela	677 Alma	687 Lola	697 Mateo	707 Izan	717 Gonzalo
668 Valeria	678 Ana	688 Irene	698 Adrian	708 Marcos	718 Jorge

(continued)

Female			Male		
669 Alba	679 Chloe	689 Jimena	699 Alvaro	709 Marco	719 Juan
670 Emma	680 Marta	690 Alejandra	700 Manuel	710 Alex	720 Angel
Turkey					
721 Zeynep	731 Meryem	741 Beren	751 Yusuf	761 Kerem	771 İbrahim
722 Elif	732 Nisanur	742 Ayşe	752 Berat	762 Ayaz	772 Ömer Faruk
723 Hiranur	733 Hira	743 Merve	753 Mustafa	763 Çınar	773 Umut
724 Yağmur	734 Belinay	744 Ada	754 Ömer	764 Ali	774 Furkan
725 Ecrin	735 Hira Nur	745 Defne	755 Ahmet	765 Enes	775 Mert
726 Zehra	736 Rabia	746 Fatma	756 Eymen	766 Muhammed Ali	776 Rüzgar
727 Azra	737 Ela	747 Asya	757 Muhammed	767 Yunus Emre	777 Hasan
728 Miray	738 Esma	748 Melek	758 Miraç	768 Hamza	778 Muhammed Emin
729 Nehir	739 Miray	749 Esmanur	759 Mehmet	769 Emirhan	779 Burak
730 Eylül	740 Sümeyye	750 Hatice	760 Emir	770 Hüseyin	780 Yiğit

References

Ackermann, Tanja. 2011. *Aloe vera* vs. *Click* – Zur phonologischen Kodierung von Geschlecht bei Warennamen (Deodorants). *Beiträge zur Namenforschung* 46. 1–50.

Alford, Richard D. 1988. *Naming and identity: A cross-cultural study of personal naming practices*. New Haven, CT: HRAF Press.

Auer, Peter. 1993. *Is a rhythm-based typology possible? A study of the role of prosody in phonological typology*. KontRI Working Paper 21. Universität Konstanz. Available at: http://paul.igl.uni-freiburg.de/auer/userfiles/downloads/Phonotypo_Kontri1.pdf.

Barry, Herbert & Aylene S. Harper. 1995. Increased choice of female phonetic attributes in first names. *Sex Roles* 32. 809–819.

Cai, Zhenguang G. & Nan Zhao. 2019. The sound of gender: Inferring the gender of names in a foreign language. *Journal of Cultural Cognitive Science* 3. 63–73.

Caro Reina, Javier & Renata Szczepaniak. 2014. Introduction: Syllable and word languages. In Javier Caro Reina & Renata Szczepaniak (eds.), *Syllable and word languages* (Linguae & litterae 40), 8–39. Berlin & Boston: De Gruyter.

Cassidy, Kimberly Wright, Michael H. Kelly & Lee'at J. Sharoni. 1999. Inferring gender from name phonology. *Journal of Experimental Psychology* 128. 362–381.

Cutler, Anne, James McQueen & Ken Robinson. 1990. Elizabeth and John: Sound patterns of men's and women's names. *Journal of Linguistics* 26. 471–482.

Elsen, Hilke. 2016. *Einführung in die Lautsymbolik*. Berlin: Schmidt.

Elsen, Hilke. 2017. The two meanings of sound symbolism. *Open Linguistics* 3. 491–499.

Fredrickson, Annie. 2007. Phonological cues to gender in sex-typed and unisex names. Available at: www. swarthmore.edu/SocSci/Linguistics/Papers/2007/fredrickson_annie.pdf.

Gerhards, Jürgen. 2010. *Die Moderne und ihre Vornamen: Eine Einladung in die Kultursoziologie*, 2nd edn. Wiesbaden: VS Verlag für Sozialwissenschaften.

Gries, Stefan Th. 2013. *Statistics for linguistics with R: A practical introduction*, 2nd edn. Berlin & Boston: De Gruyter Mouton.

Handschuh, Corinna. 2019. The classification of names: A crosslinguistic study of sex-specific forms, classifiers, and gender marking on personal names. *STUF – Language Typology and Universals* 72. 539–572.

Harrell Jr, E. Frank. 2019. rms: Regression modeling strategies. R package version 5.1-3. Available at: https://CRAN.R-project.org/package=rms.

Hinton, Leanne, Johanna Nichols & John J. Ohala. 1994. Introduction: Sound-symbolic processes. In Leanne Hinton, Johanna Nichols & John J. Ohala (eds.), *Sound symbolism*, 1–12. Cambridge: Cambridge University Press.

Hosmer, David W. & Lemeshow. Stanley. 2000. *Applied logistic regression: Hosmer/applied logistic regression*. Hoboken, NJ: Wiley.

Hough, Carole. 2000. Towards an explanation of phonetic differentiation in masculine and feminine personal names. *Journal of Linguistics* 36. 1–11.

Klink, Richard R. 2000. Creating brand names with meaning: The use of sound symbolism. *Marketing Letters* 11. 5–20.

Knoeferle, Klemens, Jixing Li, Emanuela Maggioni & Charles Spence. 2017. What drives sound symbolism? Different acoustic cues underlie sound-size and sound-shape mappings. *Scientific Reports* 7. 1–11.

Kürschner, Sebastian. 2018. Von *Helga* und *Björn* zu *Sara* und *Aron*: Der neueste Wandel isländischer Rufnamen im Vergleich mit dem Deutschen. In Seelow Hubert, Lena Rohrbach & Sebastian Kürschner (eds.), *Deutsch-isländische Beziehungen: Festschrift für Hubert Seelow zum 70. Geburtstag* (Berliner Beiträge zur Skandinavistik 24), 301–312. Berlin: Nordeuropa-Institut der Humboldt-Universität.

Lieberson, Stanley & Eleanor O. Bell. 1992. Children's first names: An empirical study of social taste. *American Journal of Sociology* 98. 511–554.

Lieberson, Stanley & Kelly S. Mikelson. 1995. Distinctive African American names: An experimental, historical, and linguistic analysis of innovation. *American Sociological Review* 60. 928–946.

Muchnik, Malka. 2017. Personal names in Modern Hebrew: A morphosyntactic and gender analysis. *Folia Linguistica* 51. 369–390.

Neef, Martin. 2002. Das Maß aller Dinge: Sonorität. In Michael Bommes, Christina Noack & Doris Tophinke (eds.), *Sprache als Form: Festschrift für Utz Maas: Zum 60. Geburtstag*, 32–48. Wiesbaden: Westdeutscher Verlag.

Nübling, Damaris. 2009. Von *Monika* zu *Mia*, von *Norbert* zu *Noah*: Zur Androgynisierung der Rufnamen seit 1945 auf prosodisch-phonologischer Ebene. *Beiträge zur Namenforschung* 44. 67–110.

Nübling, Damaris. 2012. Von *Elisabeth* zu *Lilly*, von *Klaus* zu *Nico*: Zur Androgynisierung und Infantilisierung der Rufnamen von 1945 bis 2008. In Susanne Günthner, Dagmar Hüpper & Constanze Spieß (eds.), *Genderlinguistik: Sprachliche Konstruktionen von Geschlechtsidentität* (Linguistik – Impulse und Tendenzen 45), 319–357. Berlin & Boston: de Gruyter.

Nübling, Damaris. 2017. Beziehung überschreibt Geschlecht. Zum Genderindex von Ruf- und Kosenamen. In Angelika Linke & Juliane Schröter (eds.), *Sprache und Beziehung* (Linguistik – Impulse & Tendenzen 69), 99–118. Berlin & Boston: de Gruyter.

Nübling, Damaris. 2018. *Luca* und *Noah* – Das phonologische Degendering von Jungennamen seit der Jahrtausendwende. In Damaris Nübling & Stefan Hirschauer (eds.), *Namen und Geschlechter: Studien zum onymischen un/doing Gender (Linguistik – Impulse & Tendenzen* 76), 239–270. Berlin & Boston: de Gruyter.

Nübling, Damaris, Fabian Fahlbusch & Rita Heuser. 2015. *Namen: Eine Einführung in die Onomastik*, 2nd edn. Tübingen: Narr.

Oelkers, Susanne. 2003. *Naming Gender. Empirische Untersuchungen zur phonologischen Struktur von Vornamen im Deutschen*. Frankfurt am Main: Lang.

Oelkers, Susanne. 2004. Der Fall *Luca:* Zur Männlichkeit und Weiblichkeit von Vornamen. *Bulletin suisse de linguistique appliquée* 80. 155–170.

Ohala, John J. 1983. Cross-language use of pitch: An ethological view. *Phonetica* 40. 1–18.

Ohala, John J. 1984. An ethological perspective on common cross-language utilization of F0 of voice. *Phonetica* 41. 1–16.

Pitcher, Benjamin J., Alex Mesoudi & Alan G. McElligott. 2013. Sex-biased sound symbolism in English-language first names. *PLoS ONE* 6. 1–6.

R Core Team. 2019. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Available at: https://www.R-project.org/.

Sapir, Edward. 1929. A study in phonetic sound symbolism. *Journal of Experimental Psychology* 12. 225–239.

Shrum, L. J. & Tina M. Lowrey. 2007. Sounds convey meaning: The implications of phonetic symbolism for brand name construction. In Tina M. Lowrey (ed.), *Psycho-linguistic phenomena in marketing communications*, 39–58. Mahwah, NJ: Lawrence Erlbaum.

Slater, Anne Saxon & Saul Feinman. 1985. Gender and the phonology of North American first names. *Sex Roles* 13. 429–440.

Slepian, Michael L. & Adam D. Galinsky. 2016. The voiced pronunciation of initial phonemes predicts the gender of names. *Journal of Personality and Social Psychology* 110. 509–527.

Suire, Alexandre, Alba Bossoms Mesa, Michel Raymond & Melissa Barkat-Defradas. 2019. Sex-biased sound symbolism in French first names. *Evolutionary Human Sciences* 1. 1–17.

Vennemann, Theo. 1982. Zur Silbenstruktur der deutschen Standardsprache. In Theo Vennemann (ed.), *Silben, Segmente, Akzente (Linguistische Arbeiten 126)*, 261–305. Tübingen: Niemeyer.

van de Weijer, Jeroen, Ryan G.-Y. Ren, Joost van de Weijer, Weiyun Wie & Yumeng Wang. 2020. Gender identification in Chinese names. *Lingua* 234. 1–18.

Whissell, Cynthia. 2001. Sound and emotion in given names. *Names* 49. 97–120.

Wright, Saundra K., Jennifer Hay & Tessa Bent. 2005. Ladies first? Phonology, frequency, and the naming conspiracy. *Linguistics* 43. 531–561.

Yorkston, Eric & Geeta Menon. 2004. A sound idea: Phonetic effects of brand names on consumer judgments. *Journal of Consumer Research* 31. 43–51.