https://doi.org/10.24303/lakdoi.2024.32.1.93

Vowel Phonology of Gendered Names in Linguistically Comparable Countries: Korea, Bangladesh, Britain, and the United States*

Jong-mi Kim[1], U-ri Go[1] & Sharon N. Obasi[2**]
([1]Kangwon National University & [2]University of Nebraska at Kearney)

Kim, Jong-mi; Go, U-ri & Obasi, Sharon N. (2024). Phonology of gendered names in linguistically comparable countries: Korea, Bangladesh, Britain, and the United States. *The Linguistic Association of Korea Journal, 32*(1), 93-112. Prior investigations of the phonology of names have revealed both shared and distinct phonological patterns. In this study, we conducted a comprehensive comparison of given names from linguistically diverse and comparable countries, namely Bangladesh, South Korea, Britain, and the United States. We aimed to determine the commonalities and variations in the attribution of gender to names based on phonological features. The findings elucidated a prevalent phonological pattern associated with gender, specifically centered on final vowel letters *a, e,* and *i*, across all four countries. However, notable disparities in both frequency and pattern emerged, reflecting different familial relationships: unrelated for English and Korean, related for English and Bengali, and identical for British and American English. This cross-linguistic evidence of the gender specificity of final vowel letters in names indicates that certain aspects of name phonology are universally gender-specific.

* A website hosting the complete dataset used in this study can be found at https://github.com/NamePhonology/GenderedNamesInKoreaBangladeshBritainUSA. The data for Bangladeshi and older Korean names were directly obtained from their respective governments, whereas the remaining datasets were sourced from the government websites referenced in the paper. The authors extend their gratitude to the research assistants who helped collect and analyze the data: Sharmin Sumayah for Bangladeshi names, Sumin Kim for British names, Hee-Jin Shin, and Sun Ah Kim for Korean and US names. An earlier version of this paper was presented at the 2023 annual meeting of the American Name Society.

** Jong-mi Kim is the first and corresponding author. U-ri Go and Sharon N. Obasi are the second and third authors.

Key Words: phon-onomastics, name phonology, vowels, gender, Korean names, English names, Bangladeshi names, cross-linguistic name analysis

1. Introduction

This study investigated three primary aspects. First, we delved into the gender specificity (masculine vs. feminine) of names, with particular emphasis on the significance of the final vowels. Second, a comprehensive cross-linguistic comparison was conducted, examining names across the English, Bengali, and Korean languages, while also exploring dialectal variations within English names observed in both the United States and Britain[1]. Third, our inquiry extended the examination of sound symbolism by considering variations within this shared phonological characteristic. These three focal points serve to build upon and expand on arguments that were initially posited by Barry and Harper (2000) for English names exclusively and Kim and Obasi (2023) for both English and Korean names.

1.1. Interplay of Phonology and Onomastics

Scholars have noted the interplay of phonology and onomastics. Generally, many scholars have supported the claim that vowels and sonorants symbolize femaleness (Barry & Harper, 1995, 1998, 2000, 2014; Cai & Zhao, 2019; Cassidy et al., 1999; Hough, 2000; Kim et al., 2022; Köhler, 1970; Sidhu & Pexman, 2015; Slater & Feinman, 1985). Barry and Harper (2000) observed that the final *a, e,* and *i* letters[2] identify the gender of the first names in English. Kim and Obasi (2023) extended this claim to include Korean names.

However, the extent and manner in which cross-linguistic phonological cues operate remain unclear in the existing literature. For instance, Mutsukawa (2014) revealed both similarities and dissimilarities in specific phonological patterns. He observed that longer

1) Britain encompasses England and Wales exclusively, Great Britain extends to include Scotland, whereas the United Kingdom encompasses all four: England, Wales, Scotland, and Northern Ireland (Barrow, n.d.). Thus, the term 'Britain' in this paper refers specifically to England and Wales. Names from Scotland and Ireland, with potentially more divergent naming patterns, were excluded. When comparing naming patterns with U.S. names, we focused on British names from Wales and England, assuming more shared properties than with those from Scotland and Ireland.

2) In literature, as well as in this study, vowel letters in orthography, rather than vowel sounds in the International Phonetic Alphabet (IPA), have been studied for consistency of data collection and analysis. This approach was chosen because vowel sounds vary diachronically throughout history and synchronically across languages and dialects, ensuring a more uniform and comparable dataset for analysis.

names tended to be perceived as masculine in Japanese but as feminine in English, whereas monosyllabic names were associated with masculinity in both Japanese and English. In contrast, Cai and Zhao (2019) presented cross-linguistic sound symbolism as a means of inferring name gender in an unfamiliar language. Their experiments involved native speakers of English and German identifying real personal names in Min, a South Chinese language to which they had not been exposed. Kim and Obasi (2023) also proposed that certain aspects of name phonology, referred to as "phon-onomastics" or "phono-onomastics"[3], may be universal. Their argument was based on the analysis of Korean and English data, which belong to unrelated language families.

This study aimed to expand the existing foundation by incorporating data from linguistically disparate and comparable names in Korea, Bangladesh, Britain, and the United States. Our objective was to investigate whether certain aspects of name phonology vary based on linguistic distance. The decision to focus on the *a, e,* and *i* vowels as our research scope[4] was motivated by the limited number of segments tested in previous studies that employed sound grouping across languages (Kim & Obasi, 2023).

1.2. Linguistically Disparate and Comparable Names

In our investigation, we selected English and Korean as the languages from unrelated families, with English belonging to the Indo-European language family and Korean being classified as an Altaic language[5].

Furthermore, we chose English and Bengali as languages from comparable families. Both languages originated from the Indo-European ancestor and developed into sister

3) We define "phon-onomastics" or "phono-onomastics" as the study of name phonology, encompassing the scientific examination of sound patterns in proper names, including their origin, history, and usage. This definition is constructed by combining the meanings of 'phonology,' which refers to the sound patterns of languages (adapted from Britannica), and 'onomastics,' which pertains to the study of the origin, history, and use of proper names (inspired by the work of the American Name Society).

4) This specific grouping of *a, e,* and *i* vowel-ending names is contrastive in Korean names, where male names exhibit considerably more *o* and *u* endings, as in *Minsu* and *Junho*, for reasons yet unknown. In addition, there is a striking similarity in English names, wherein *o* and *u* endings occur more frequently in male names, as evidenced by examples such as *Antonio, Olivero*, and *Beau* (Barry & Harper, 2000).

5) For the purpose of discussion, we provisionally categorized Korean within the Altaic language family, aligning with the perspective presented by Sohn (1999, pp. 18-25). It is noteworthy that alternative viewpoints, such as considering the Korean language as a member of the Koreanic family, remain consistent with our overarching argument that the Korean language is unrelated to the Indo-European language family.

families of Germanic languages for English,[6] and Indo-Iranian languages for Bengali[7].

In addition, we explored the comparable dialects of a single language, namely British and American English. This study extended the investigation into gendered phonology from our previous work by Kim and Obasi (2023), now including additional data from English and Bengali as comparable language families and British and American English as dialects of the same language.

1.3 Research Question

The new research question in the present study is the following.

(1) Research Question
In the name data from Korea, Bangladesh, Britain, and the United States, final vowel letters *a, e,* and *i* identify female names:
a. In all languages, regardless of their varied linguistic distance?
b. With varied degrees by linguistic distance?

Research question (1a) explores the possibility of universal phon-onomastics across varied linguistic distance including unrelated, related, and dialectal contexts. This question yields a positive result if female names in all four countries consistently end with *a, e,* and *i* letters considerably more often than male names do. If affirmed, the subsequent research question (1b) investigates whether phon-onomastics is sensitive to linguistic distance. A positive response to this question indicates that the gender specificity demonstrated by these final vowel letters varies across linguistic distances. For instance, a substantial difference may be observed in English and Korean names, belonging to unrelated language families (English as an Indo-European and Korean as an Altaic language), whereas a smaller difference may be found in British and American dialects of the same English language. Similarly, some differences may exist between Bengali and English, as they belong to daughter families within the Indo-European language family (English being Germanic and Bengali being Indo-Aryan).

6) English gendered names draw from Latin phonology (Hough, 2000; Cassidy et al., 1999; Slater & Feinman, 1985, p. 429), where vowel endings are more common for female than male names, and more syllables are typically found in female than male names.

7) Bengali belongs to the Indo-Aryan language family within the Indo-Iranian branch of the Indo-European languages. Indo-Aryan languages are spoken in India, Pakistan, Bangladesh, Nepal, and Sri Lanka (adapted from Britannica).

This paper is organized as follows. First, we discuss the method of data collection for the historically top 20 popular names per decade in Korean, Bengali, and two dialects of English (American and British). Subsequently, our main results are presented, demonstrating gender specificity through final vowel letters across these languages and dialects with some variation. Finally, we explore the universal sounds for gender specificity with typological geographical differences.

2. Method

To address the research questions pertaining to the phon-onomastic universality and variability of gender specificity, our methodology involved the systematic collection and analysis of the distribution of final vowels in comparable name data from reliable sources across languages with varying linguistic distances. This process included a comprehensive gathering and scrutiny of the top 20 names of each gender (male and female) per decade throughout history, sourced from governments representing different linguistic relationships, encompassing those belonging to unrelated language families (Korean and English), relative language families (Bengali and English), and the same language family (British and American English). The information derived from this historical examination of phon-onomastics and gender in these countries was utilized to determine the phonological similarities and dissimilarities of gendered personal names over time in comparable languages.

2.1. Collection of Name Data

We collected the top 20 names per decade and gender (male and female) registered in each of the four governments: Britain, the United States, Korea, and Bangladesh, where the national language serves as a linguistic contrast for our research purposes. Data were sourced from authentic channels, including direct interactions with government officials for Bangladeshi names and older Korean names predating 2010.[8] Data for other countries were sourced from government websites.

8) Recent Korean name data (Supreme Court of Korea, 2024) were publicly available, whereas older names were obtained through a petition for information release (#1301 on November 2, 2018, and #42 on January 28, 2019). These data, previously utilized in Kim and Go (2022) and Kim and Obasi (2023), have been updated and incorporated into this paper. We exclusively utilized birth names, aligning with the data consistency across the three other countries compared in this study.

The availability of data varied by country. The Bangladeshi government provided the top 20 male and female birth names from the 1990s to the 2020s (2020—2022). The Korean government (Supreme Court, 2024) facilitated data access by providing the top 20 preferred birth names for males and females from 1940 to 2022. The British government (Office for National Statistics, 2024) supplied data on popular birth names every ten years from 1904 to 1994. Starting in 2000, British government data were released annually until 2021. The US dataset includes names from the 1880s to the 2020s (2020—2022) from the Social Security Administration. Decadal data were collected through yearly summations, particularly when they were not directly available from each government.

The complete dataset comprises 1,640 personal names, encompassing the top 20 most popular names in each decade for their respective countries. Specifically, it included 160 Bangladeshi names (20 names, 4 decades, 2 sexes), 360 Korean names (20 names, 9 decades, 2 sexes), 520 British English names (20 names, 13 decades, 2 sexes), and 600 American English names (20 names, 15 decades, 2 sexes).

Table 1 displays a curated excerpt from the dataset representing the decade 2010 for all four countries. Romanization is provided in parentheses for names officially registered in non-Roman letters. The Official Romanization System[9] for Korean orthography and the National Library at Kolkata romanization[10] for Bengali orthography were employed to ensure consistency in name representation across languages.

Table 1. Top 20 Names by Decade in Korea, Bangladesh, England, and the United States (2010-2019)

Country	Top 20 male names	Top 20 female names
Korea	민준(Min-jun) 서준(Seo-jun) 도윤(Do-yun) 예준(Ye-jun) 주원(Ju-won) 시우(Si-u) 지호(Ji-ho) 지후(Ji-hu) 하준(Ha-jun) 준우(Jun-u) 건우(Geon-u) 준서(Jun-seo) 도현(Do-hyeon) 우진(U-jin) 현우(Hyeon-u) 지훈(Ji-hun) 선우(Seon-u) 서진(Seo-jin) 현준(Hyeon-jun) 유준(Yu-jun)	서윤(Seo-yun), 서연(Seo-yeon) 지우(Ji-u) 서현(Seo-hyeon) 하은(Ha-eun) 하윤(Ha-yun) 민서(Min-seo) 지유(Ji-yu) 채원(Chae-won) 윤서(Yun-seo) 다은(Da-eun) 소율(So-yul) 지아(Ji-a) 지윤(Ji-yun) 지민(Ji-min) 수아(Su-a) 은서(Eun-seo) 지안(Ji-an) 예원(Ye-won) 하린(Ha-rin)

9) Korea Ministry of Culture, Sports, and Tourism. (2000). *Romanization of Korean.*
10) Library of Congress, USA. (2024). *Bengali Romanization Table*.

Country	Top 20 male names	Top 20 female names
Bangladesh	আসিফ(Asif) সাব্বির(Sabbir) মারুফ(Maruf) শাহরিয়ার(Shahriyar) রিয়াজুল(Riyajul) ফাহিম (Fahim) আমিন (Amin) আমির (Amir) হামজা (Hamja) হাবিব (Habib) রাসেল (Rasel) সুমন (Sumon) রিয়াজুল (Riyajul) সজীব (Shajib) শরিফুল (Shoriful) রাকিব (Rakib) ফিরোজ (Firoz) মামুন (Mamun) আশরাফুল (Ashraful) তানজিল (Tanjil)	রোকেয়া (Rokeya) মাইশারা (Maishara) স্নেহা (Sneha) শারমিন (Sharmin) ফাতিমা (Fatima) মেহরীমা(Mehrima) সুমাইয়া(Sumaiya) মেহরিন(Mehrin) দিয়া (Diya) নাহিদা(Nahida) নাদিয়া (Nadiya) লিপি (Lipi) নিপা (Nipa) আফিয়া (Afiya) তাশফিয়া (Tashfiya) মারজিয়া (Marjia) ইরিন (Irin) আশা (Asha) সুমনা (Sumona) সামিরা (Samira)
Britain	Oliver Harry Jack George Charlie Jacob Thomas Alfie William Noah Joshua Oscar James Muhammad Leo Ethan Henry Archie Freddie Samuel	Amelia Olivia Emily Ava Lily Isla Isabella Mia Grace Poppy Sophie Evie Jessica Ella Sophia Ruby Freya Chloe Isabelle Charlotte
USA	Noah Liam Jacob William Mason Ethan Michael Alexander James Elijah Benjamin Daniel Aiden Logan Jayden Matthew Lucas David Jackson Joseph	Emma Olivia Sophia Isabella Ava Mia Abigail Emily Charlotte Madison Elizabeth Amelia Evelyn Ella Chloe Harper Avery Sofia Grace Addison

Note. Names in each cell are ordered by popularity rank from left to right, and top to bottom.

This table showcases the 20 most popular names by gender for the years spanning 2010 to 2019 in Korea, Bangladesh, England, and the United States. Our comprehensive dataset, which includes the names used in this table, is available on GitHub (see Footnote * on the title page), encompasses the entire range of historically available data from government records with public access.

2.2. Phon-onomastic Analysis of Multilingual Data

After romanizing non-English names from countries where English is not the primary language, we tallied the occurrences of names ending with the vowel letters *a, e,* and *i* for each gender in each decade. Examples include *Ye-na* (Korean), *Lipi* (Bangladeshi), and *Grace* (American). However, we excluded the British name *Stanley*, which concludes with the letter 'y' and lacks one of the final *a, e,* and *i* vowels. Subsequently, we compared the number of female and male names.

To evaluate the significant gender-based differences in name preferences within each country, we employed an independent *t*-test. This statistical test allowed us to examine whether notable disparities in name preferences based on gender existed within the respective countries. In addition, an ANOVA test was conducted to investigate the

potential differences in name preferences across the four countries for both genders. This analysis enabled us to determine whether significant variations in name choice existed across the countries considered. Through these statistical tests, we aimed to reveal meaningful insights into gender-based differences within individual countries and the potential variations in name preferences among the four countries. These two statistical analyses identified significant patterns or trends in the data, thereby enhancing our understanding of the factors influencing name preferences across various demographics and geographic locations.

3. Results

The results showed a consistent gender-specific phonological pattern related to vowels across all four countries, with persistent differences between male and female names from 1880 to 2022. However, the frequency and pattern of these phonological differences varied significantly depending on the historical trends unique to each country.

3.1. Comparison of Gendered Names by Country (1880–2022)

Historical analysis demonstrated that in all four countries (Bangladesh, Korea, Britain, and the United States), female names consistently featured a higher frequency of vowel letters than male names. Figure 1 presents the distribution of the final letters *a, e,* and *i* in names across genders for these four countries (N = 1,640 names, comprising 160 from Bangladesh, 360 from Korea, 520 from Britain, and 600 from the United States).

In Figure 1, male names across various languages demonstrate a notable scarcity of the final *a, e,* and *i* letters. In contrast, female names consistently exhibited abundant occurrences of these vowel letters. A distinct gender-based pattern emerged, revealing significant differences among Bangladeshi names (69%), Korean names (27%), British names (41%), and USA names (43%) in terms of the percentage difference between male and female names concluding with these final vowel letters. Overall, names ending with these vowels serve as effective gender identifiers across all languages, with the most prominent distinction observed in Bangladesh, followed by Britain, the United States, and to some extent Korea.[11)]

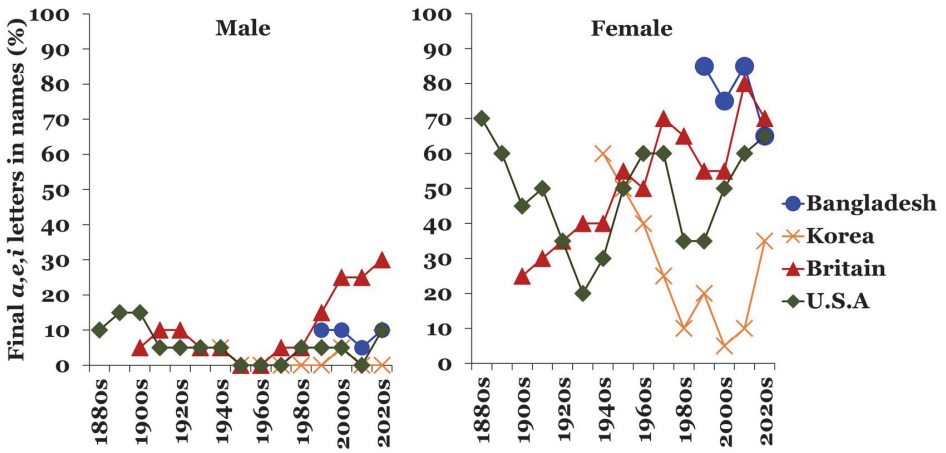

Figure 1. Comparison of Vowel Patterns in Names by Gender in the Countries of Bangladesh, Korea, Britain, and the United States (1880-2022)

Results from an independent sample *t*-test conducted on male and female names showed that female names consistently exhibited a greater frequency of the final vowel letters, *a, e,* and *i,* compared to male names across all four countries and decades. For instance, Bangladeshi female names (M = 0.78, SD = 0.42, N = 80) displayed a significantly higher number of occurrences for these vowel letters per name than male names (M = 0.09, SD = 0.28, N = 80, $t(138.8)$ = 12.12, $p < .001$, Glass's d = 1.64, 95% CI [0.58, 0.80], two-tailed).[12] Levene's test indicated unequal variances (F = 26.4, $p < .001$); hence, the degrees of freedom were adjusted from 158 to 138.8. This effect size of 1.64 is very large, suggesting a substantial difference between male and female names in Bangladesh in terms of the occurrence of the final vowels. Note that all numerical values of frequency in our statistical reports refer to the number of occurrences per name. Importantly, the statistical M for mean or average cannot be higher than 1. In addition, the same numerical values are presented in percentages in all figures to ensure presentational clarity.

11) The observed trend of gender specificity aligns with survey findings in Korea and the United States (Kim and Obasi, 2023). The survey respondents in their study consistently identified a high level of gender specificity for US names and a moderate level for Korean names. This pattern held true for both given names and changed names.

12) Glass's delta was employed due to the presence of distinct standard deviations within each group. Cohen's d yielded even larger effect sizes (1.93 for Bangladeshi, 0.82 for Korean, 0.99 for British, and 1.08 for US names).

Similar to the case of Bangladeshi names, Korean female names at birth (M = 0.28, SD = 0.45, N = 180) exhibited a significantly greater number of occurrences of these vowel letters compared to male names (M = 0.01, SD = 0.11, N = 180, t(198.3) = 7.87, p < .001, Glass's d = 0.60, 95% CI [0.20, 0.34], two-tailed). Levene's test indicated unequal variance (F = 543.1, p < .001); hence, the degrees of freedom were adjusted from 358 to 198.3. This effect size of 0.60 is also very large, suggesting a substantial difference between male and female names in Korea in terms of the occurrence of the final vowels.

This trend was consistent across British female names (M = 0.52, SD = 0.50, N = 260), which displayed a significantly higher frequency of occurrence of these vowel letters than male names (M = 0.11, SD = 0.31, N = 260, t(432.6) = 11.16, p < .001, Glass's d = 0.82, 95% CI [0.34, 0.48], two-tailed). Levene's test indicated unequal variances (F = 411.9, p < .001); hence, the degree of freedom was adjusted from 518 to 432.6. This effect size of 0.82 is also very large, suggesting a substantial difference between male and female names in Britain in terms of the occurrence of the final vowels.

Likewise, the US female names (M = 0.48, SD = 0.50, N = 300) displayed a significantly higher number of occurrences of these vowel letters compared to male names (M = 0.06, SD = 0.23, N = 300, t(421.4) = 13.40, p < .001, Glass's d = 0.84, 95% CI [0.36, 0.49], two-tailed). Levene's test indicated unequal variances (F = 1089.1, p < .001); hence, the degree of freedom was adjusted from 598 to 421.4. This effect size of 0.84 is also very large, suggesting a substantial difference between male and female names in the United States. in terms of the occurrence of the final vowels.

3.2. Comparison of Gendered Names by Linguistic Distance (1990–2022)

Historical analysis also showed varying degrees of genderedness based on the linguistic distance for female names with vowel letters compared with male names. Figure 2 shows the distribution of the final *a, e,* and *i* letters in names across genders for comparing countries: British and American English as the same language, Bengali and English as related languages, and Korean and English as unrelated languages (N = 780 names, comprising 80 from Bangladesh, 180 from Korea, 260 from Britain, and 260 from the United States)[13].

13) The total count of 780 names is derived from 20 names per decade across the following periods: Britain 1900-2020, USA 1900-2020, Bangladesh 1990-2020, and Korea 1940-2020.

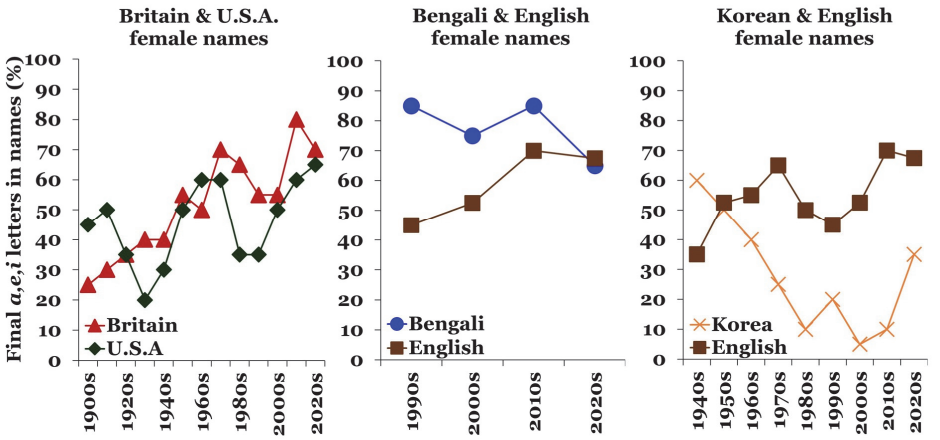

Figure 2. Comparison of Vowel Patterns in Female Names by Linguistic Distance Among Bengali, Korean, and British and American English (1900-2022)

The leftmost graph in Figure 2 illustrates a similar increasing trend in the names of Britain and the United States with a small average distance (6%) between them (52% vs 46%, respectively). This indicates a minor difference between British and US names, both belonging to the English language.

In contrast, the rightmost graph reveals a substantial difference between English and Korean names, originating from two unrelated language families (Indo-European vs. Altaic). The distinctive patterns of increase and decrease, along with an average distance of 27% (55% for English names and 28% for Korean names), underscore the significant divergence between these two linguistic groups.

The middle graph illustrates a somewhat similar pattern for Bengali and English names, both demonstrating frequent occurrences of these final vowel letters (78% vs 59%, respectively). This indicates a moderate difference between these languages, which belong to daughter language families within the same Proto-Indo-European family. Thus, we observed varying degrees of gender specificity based on linguistic distance.

To examine gender specificity in female names during the overlapping years in all four countries, Figure 3 presents varying degrees (N = 320 names, comprising 80 from each country: Bangladesh, Korea, Britain, and the United States).

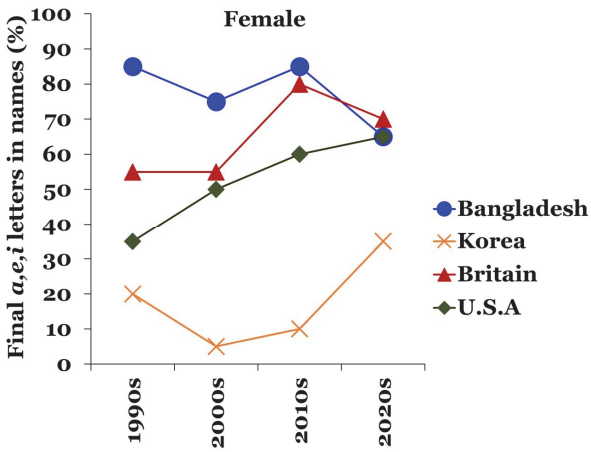

Figure 3. Comparison of Vowel Patterns in Female Names by Country in Bangladesh, Korea, Britain, and the United States (1990-2022)

In Figure 3, gendered naming by final *a, e,* and *i* letters in female names prevails across all Indo-European languages spoken in Bangladesh, Britain, and the United States, but is not as prominent in Korean, which is an unrelated language (78% for Bangladesh, 65% for Britain, 52% for the United States, and 18% for Korea). Among the countries speaking Indo-European languages, Britain and the United States exhibited similar patterns and frequencies (%) of these vowel letters. These similarities align with linguistic distance, with Britain and the United States being the closest among the four countries.

In the statistical analysis of the female names of all four countries (N = 820) using a one-way ANOVA test, a significant effect was found ($F(3, 816) = 20.4$, $p < .001$]). Post-hoc tests using the Scheffe test[14] revealed significant mean differences in the number of occurrences of final *a, e,* and *i* vowel letters among the female names from Bangladesh, Korea, Britain, and the United States; between Bangladesh (M = 0.78, SD = 0.42) and Korea (M = 0.28, SD = 0.45, $p < .001$, Cohen's d = 1.15); between Bangladesh (M = 0.78, SD = 0.42) and the United States (M = 0.48, SD = 0.50, $p < .001$, Cohen's d = 0.65; between Bangladesh (M = 0.78, SD = 0.42) and Britain (M = 0.52, SD = 0.50, p = .001, Cohen's d = 0.56); between Britain (M = 0.52, SD = 0.50) and Korea (M = 0.28, SD = 0.45,

14) Scheffe's test, chosen for its conservatism, yielded highly significant results in distinguishing dialectal from language variations across countries, given the absence of an assumption of equal population variances among the four countries. This test allows comparing more than just two means at once, unlike the Tukey post-hoc test.

$p < .001$, Cohen's $d = 0.50$); and between the United States ($M = 0.48$, $SD = 0.50$) and Korea ($M = 0.28$, $SD = 0.45$, $p < .001$, Cohen's $d = 0.42$).[15] However, the differences in the number of occurrences of these final vowel letters among the female names were statistically insignificant ($p = .89$) within the same language, that is, between British and American English.

3.3. Statistical Significance: Influence of Gender and Language Groups on Vowel-ending Name Preferences

Multiple two-way ANOVAs were performed to analyze the effects of gender and language groups (dialects, languages, sister families, and unrelated families) on preferences for names ending in *a, e,* and *i* ($N = 1{,}440$). Assumptions of normality, homogeneity of variance, and independence of the observations were confirmed. This particular set of statistical analyses aimed to determine how gender or language groups affect the preferences of such vowel-ending names and whether there was an interaction between the two factors (gender and language groups) on the preferences for such vowel-ending names.

For the first question on main effects, the gender factor exhibited a significant main effect on the vowel-ending name preferences, $F(1, 1632) = 434.14$, $p < .001$, $\eta^2 = 0.21$. Post-hoc analyses indicated a significantly higher preference for *a, e,* and *i* vowel-ending in female names ($M = 0.48$, $SD = 0.50$) than in male names ($M = 0.07$, $SD = 0.25$). As for the language group factor, of which each one of four language groups was embedded in each of the four two-way ANOVAs, significant main effects were demonstrated in three language groups (languages, sister families, and unrelated families), but not the dialectal group: $F(1, 1116) = 2.97$, $p = .09$, $\eta^2 = 0.002$ for the dialectal factor (American English and British English); $F(2, 1634) = 16.34\ 47.74$, $p < .001$, $\eta^2 = 0.04$ for the language factor (Bengali, English, and Korean languages); $F(1, 1276) = 17.85$, $p < .001$, $\eta^2 = 0.01$ for the related language factor (Bengali as an Indo-Iranian branch and English as a Germanic branch of the Indo-European family); $F(1, 1636) = 47.80$, $p < .001$, $\eta^2 = 0.03$ for the unrelated language factor (Bengali and English from the Indo-European family and Korean from the Altaic family). Therefore, the main effects in the four two-way ANOVAs revealed that both gender and language groups of the three languages, and two sister

15) Cohen's *d* was employed due to similar standard deviations among the groups. Glass's *d* yielded similar effect sizes (1.11 for Bengladesh and Korea, 0.71 for Bangladesh and the United States, 0.62 for Bangladesh and Britain, 0.53 for Britain and Korea, and 0.44 for the United States and Korea).

families, two unrelated families had a statistically significant effect on preferences for *a, e,* and *i* vowel-ending names ($p < .001$ for both the gender factor and language group factors). The effect sizes (η^2) offered further insight, revealing a large effect for the gender factor (0.21), a medium effect for the language factor (0.04), a small-to-medium effect for the unrelated language factor (0.03), a small effect for the related language factor (0.01), and insignificant difference for the dialectal factor (0.002).

For the second question on the interaction between the two factors (gender and language groups), a statistically significant interaction effect was determined between the effects of gender and each of the three language groups (languages, sister families, and unrelated families), but not the dialectal group (American English and British English). The significant interaction effects between the factors were: $F(2, 1634) = 16.34$, $p < .001$, $\eta^2 = 0.02$ for gender and languages (Bengali, English, and Korean); $F(1, 1276) = 16.10$, $p < .001$, $\eta^2 = 0.01$ for gender and sister families (Indo-Iranian and Germanic branches of the Indo-European family); and $F(1, 1636) = 15.05$, $p < .001$, $\eta^2 = 0.01$ for gender and unrelated families (Indo-European and Altaic families).

Overall, all three interaction effects between gender and each of the three language groups (but not the dialectal group) were statistically significant, suggesting meaningful relationships between gender and the specified factors (languages, sister families, and unrelated families). The effect sizes (η^2) provide additional context, indicating a larger effect (0.02) for the interaction between gender and the three languages, whereas the other two interactions (gender and sister families, gender and unrelated families) exhibited smaller effects (0.01). More specifically, only the factor for languages (Bengali, English, and Korean) has the practical significance of the findings beyond the statistical significance, among the four language group factors (dialects, languages, sister families, and unrelated families).

4. Discussion and Conclusion

In summary, our research findings affirmatively respond to our Research Question in (1), indicating that the presence of final *a, e,* and *i* letters effectively identifies female names in all four countries, albeit to varying degrees contingent on linguistic distance. More precisely, the cross-linguistic similarities of this particular gender specificity were demonstrated with diverse magnitudes, aligning with the distinctions in language families: unrelated, related, and identical.

4.1. Phon-onomastic Similarities in Gender Specificity

The first aspect of our Research Question in (1), pertaining to phon-onomastic similarities, yielded positive results, as illustrated in Figure 1. Male names exhibited few occurrences of the final *a, e,* and *i* letters, whereas female names displayed a higher frequency of these vowels across all four languages, reflecting different familial relationships: unrelated for English and Korean, related for English and Bengali, and identical for British and American English.

Our findings align with and extend previous research by Barry and Harper (2000) and Kim and Obasi (2023), indicating that the final *a, e,* and *i* letters identify female names in English and, to some extent, in Bengali and Korean as well. This cross-linguistic phon-onomastic pattern is consistent with the findings by Cai and Zhao (2019), who reported gender identification in an unknown language (Min) by foreign speakers (English and German). Cross-linguistic evidence of the genderedness of final vowel letters supports the assumption that some level of gender specificity in names may be universal.

4.2. Phon-onomastic Dissimilarities in Gender Specificity

The second aspect of Research Question (1) regarding phon-onomastic dissimilarities also yielded affirmative results, as illustrated in Figures 2 and 3, which show the varied degrees of gender specificity in names by linguistic distance. The patterns and frequencies of gender specificity were notably more dissimilar in unrelated languages than in related or the same languages. In Figure 2, a substantial difference is evident between the English and Korean names that belong to unrelated language families. In contrast, a marginal difference was observed between British and American English, which represent different dialects of the same language. Notably, there were some differences between Bengali and English, belonging to daughter families of the Proto Indo-European language family. Similarly, Figure 3 shows the cross-linguistic dissimilarities in gender specificity by linguistic distance, revealing a significant difference between Bengali and Korean names (unrelated language families), a minor difference between British and American English (different dialects of the same language), and a moderate difference between Bengali and English (daughter families of the Indo-European family).

4.3. Limitations and Future Studies

Our research is confined to examining the gender specificity of phon-onomastics across languages, focusing on the final *a, e,* and *i* letters and linguistic distance, encompassing only three languages and two dialects from the Indo-European and Korean language families. However, the generalizability of this particular gender specificity to other language families, such as Niger-Congo and Sino-Tibetan, remains questionable. The scope of future studies can be broadened by including a larger number of languages and exploring different sound groupings, such as consonants. Notably, specific sound grouping (vowels *a, e,* and *i*) for gender specificity was rigorously tested using selected cross-linguistic quantitative data in this study (four countries with a linguistically varied distance). Therefore, cross-cultural analyses involving a wider array of linguistically different countries that investigate how gender attribution in names varies across languages are warranted to validate and expand the observations made in the current study.

Although the inclusion of more countries and diverse sound groupings may enhance the external validity of our findings on cross-linguistic phon-onomastics, it would also be interesting to investigate regional variations within a country. This is particularly interesting because of the reported state differences in US naming patterns, such as namesaking practices (Obasi, 2016).[16] However, this topic presents two conflicting predictions. Regional differences may exist, similar to namesaking patterns, or they may not manifest, extending our study's suggestion that the final *a, e,* and *i* letters would exhibit even smaller differences in domestic regions than in different countries sharing the same language, as observed in the case of British and American English. Examination of regional variation in the gender attribution of names may also provide a mechanism for examining how societal changes may influence phonological patterns associated with gender within specific countries and regions.

16) "Namesaking" is the practice of naming a newborn after a specific family member. In Obasi (2016), namesaking patterns were assessed in rural, south-central Nebraska, examining 841 birth announcements printed in a local newspaper during the calendar years 1994 and 2014. The study reports that namesaking was more commonly practiced for first-borns than later-borns, more often for males than females, and more frequently after paternal relatives than maternal relatives.

4.4. Conclusion

In conclusion, our findings can be summarized in three key aspects. First, a cross-national and cross-cultural phonological pattern of names was shown. Our results establish that final vowel letters consistently identify female names across languages from different countries, even with varied linguistic distances. Second, our study indicates that certain phon-onomastic features may reflect the genetic distance between languages. The observed gender specificity in names demonstrates variations that align with the genetic distance between languages. Finally, our research implies that changes in name phonology can potentially serve as a mechanism for documenting broader linguistic and historical changes, such as those resulting from language contact and language family trees.

Revisiting the three focal points highlighted at the outset of our introduction, our study conclusively affirms the existence of gender specificity (masculine vs. feminine) in names, particularly regarding the final *a, e,* and *i* letters. In addition, the second point was validated, demonstrating that gender specificity transcends linguistic boundaries. This is evident in our cross-linguistic comparison of English, Bengali, and Korean names as well as in the dialectal examination of English names between the United States and Britain. Finally, our exploration confirmed the presence of a common sound symbolism with some variation. Gender specificity of sounds emerged as a shared characteristic of names across all four countries, showcasing nuanced degrees of variation influenced by linguistic distance.

Several unexplored questions within this area were addressed, shedding light on the universal similarity in how phonology suggests gender, language-specific peculiarities in names, phonology, and gender (Section 4.1), and the overall expansion of the study of phon-onomastics to interpret linguistic distance in familial relationships (Section 4.2).

However, phon-onomastic gender specificity based on linguistic distance remains an avenue for further exploration that is beyond the scope of our study. The findings in this study are confined to the observation that the gender specificity of the final *a, e,* and *i* vowels in female names exhibits a substantial difference between English or Bengali and Korean names (unrelated language families), a minimal difference between British and US names (same language), and a moderate difference between Bengali and English names (related language families).

References

Barrow, M. (n.d.). What is the difference between UK, Britain, Great Britain (GB) and the British Isles? Retrieved March 31, 2024, from https://projectbritain.com/britain/britain.html

Barry, H. Ⅲ., & Harper, A. S. (1995). Increased choice of female phonetic attributes in first names. *Sex Roles*, *32*(11-12), 809-819.

Barry, H. Ⅲ., & Harper, A. S. (1998). Phonetic differentiation between first names of boys and girls. In *Proceedings of the XIXth International Congress of Onomastic Sciences*, 40-46, University of Aberdeen.

Barry, H. Ⅲ., & Harper, A. S. (2000). Three last letters identify most female first names. *Psychological Reports*, *87*(1), 48-54.

Barry, H. Ⅲ., & Harper, A. S. (2014). Unisex names for babies born in Pennsylvania 1990–2010. *Names: A Journal of Onomastics*, *62*(1), 13–22.

Cai, Z. G., & Zhao, N. (2019). The sound of gender: Inferring the gender of names in a foreign language. *Journal of Cultural Cognitive Science*, *3*(1), 63–73.

Cassidy, K. W., Kelly, M. H., & Sharoni, L. J. (1999). Inferring gender from name phonology. *Journal of Experimental Psychology: General*, *128*(3), 362–381.

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

Kim, J.-M., & Go, U.-R. (2022). Historical evidence of phonological changes in Korean names, 1940–2021. *Studies in Phonetics, Phonology and Morphology*, *28*(3), 417–441.

Kim, J.-M., & Obasi, S. N. (2023). Phonological trends of gendered names in Korea and the U.S.A. *Names*, *71*(3), 36–46.

Kim, J.-M., & Obasi, S. N. (2023). Phono-onomastics of gendered names: Similarities and dissimilarities in Bangladesh, Korea, Britain, and the US. Paper presented at the American Name Society Annual Meeting (ANS 2023), Online, January 20–22.

Kim, S.-H., Mun, S., & Ko, E.-S. (2022). Sound symbolism in the gender of personal names: With focus on the game characters in Animal Crossing. *Studies in Phonetics, Phonology and Morphology*, 28(1), 33-55.

Köhler, W. (1970). Gestalt psychology: *An introduction to new concepts in modern psychology*. New York: Liveright.

Korea Ministry of Culture, Sports, and Tourism. (2000). *Romanization of Korean*. (Notice of Regulation, No. 2000-8, No. 2014-42). Available at National Institute of Korean Language, Republic of Korea. Retrieved January 31, 2024, from https://www.korean.go.kr/front_eng/roman/ roman_01.do

Library of Congress, USA. (2024). Bengali Romanization Table. Retrieved January 31, 2024, from https://www.loc.gov/catdir/cpso/romanization/ bengali.pdf

Mutsukawa, M. (2014). Phonological and semantic gender differences in English and Japanese given names. In *Proceedings of the XXIV International Congress of Onomastic Sciences*, 370-377. Retrieved January 31, 2024, from https://www.gencat.cat/llengua/BTPL/ICOS2011/041.pdf

Obasi, S. (2016). Naming patterns in rural south central Nebraska. *Names: A Journal of Onomastics*, *64*(3), 158–165. http://dx.doi.org/10.1080/00277738. 2016.1197644

Office for National Statistics, UK. (n.d.). Baby names in England and Wales: From 1996 and top 100 baby names: Historical data. Retrieved January 31, 2024, from https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/datasets/babynamesinenglandandwalesfrom1996

Office for National Statistics, UK. (n.d.). Top 100 baby names: Historical data. Retrieved January 31, 2024, from https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/datasets/babynamesenglandandwalestop100babynameshistoricaldata

Sidhu, D. M., & Pexman, P. M. (2015). What's in a name? Sound symbolism and gender in first names. *PLoS ONE*, *10*(5), 1–22.

Slater, A. S., & Feinman, S. (1985). Gender and the phonology of North American first names. *Sex Roles, 13*, 429–440. https://doi.org/10.1007/BF00287953

Social Security Administration, USA. (n.d.). *Popular names by birth year.* Retrieved November 24, 2023, from https://www.ssa.gov/oact/babynames/ index.html

Sohn, H.-M. (1999). *The Korean Language*. Cambridge University Press.

Supreme Court of Korea. (n.d.). Sangwi chulsaeng / gaemyeong singo irum hyeonhwang [Popular names registered with the government as given names and changed names]. *Electronic Family Registration System.* Retrieved January 28, 2024, from https://stfamily.scourt.go.kr/st/StFrrStatcsView.do?pgmId=090000000025

Jong-mi Kim
Professor
Department of English Language and Literature, College of Humanities
Kangwon National University
1 Kangwondaehak-gil, Chuncheon-si, Gangwon State, 24341, Republic of Korea
Phone: +82-33-250-8150
Email: kimjm@kangwon.ac.kr

U-ri Go
Lecturer
Department of English Language and Literature, College of Humanities
Kangwon National University
1 Kangwondaehak-gil, Chuncheon-si, Gangwon State, 24341, Republic of Korea
E-mail: urigo@kangwon.ac.kr

Sharon N. Obasi
Associate Professor
Department of Counseling, School Psychology and Family Science
University of Nebraska at Kearney
2504 9th Ave., Kearney, NE 68849, U. S. A.
E-mail: obasis2@unk.edu

Received on February 15, 2024
Revised version received on March 24, 2024
Accepted on March 31, 2024