

The Diverse Names Generator: An app for decreasing bias and promoting inclusion

Maura O’Leary, Rainey Williams & Mario Peng Lee*

Abstract. It has been well established that example sentences in linguistics use a remarkably non-diverse set of proper names in terms of gender, culture, and ethnicity (e.g., Macaulay & Brice 1997, Cépeda et al. 2021, and Kotek et al. 2021). Here, we introduce a new resource, the Diverse Names Generator (DNG), which provides randomly selected proper names with IPA transcriptions from a user-contributed, linguist-curated database of names from a wide range of languages and cultures. Generating names randomly helps users to overcome unconscious bias that may lead them to default to using Anglophone, male-gendered names. The DNG can be accessed both through a website interface and through a downloadable Android app, both with offline capabilities. This novel resource is the first of its kind and can be used both while preparing examples ahead of time and while generating examples live in the classroom.

Keywords. pedagogy; teaching; diversity; example sentences; proper names; resources; website; mobile app

1. Introduction. Example sentences play a crucial role in illustrating linguistic theories and phenomena. In many linguistics papers, example sentences constitute much of the data provided for the reader. Likewise, linguistics classrooms are filled with example sentences, both generated ahead of time and on the fly. In both contexts, many of the example sentences include proper names. Rarely is it the case that the phenomenon showcased in the example sentence requires a specific proper name, so, theoretically, any name could be chosen. When any name could be used without affecting the theoretical points being made, we might want to choose specific proper names for another reason – such as promoting inclusion and representation through the use of names that show gender and cultural diversity.

The linguistics community has become increasingly cognizant of the influence of unconscious biases in language use, particularly in example sentences. Recent studies examining the names used in example sentences over the last two decades have found that example sentences tend to display heteronormative gender roles and primarily feature Western characters, such as “John” and “Mary,” reflecting the dominant white culture to a greater extent than minority races and ethnicities (Pabst et al. 2018, Cépeda et al. 2021, and Kotek et al. 2021).

In an ideal scenario, the distribution of proper names in example sentences should reflect the diversity of our field and the languages we study. However, it is not fair to ask authors and instructors to think up an endless list of diverse names to use in their example sentences, nor to ask them to fully overcome the unconscious biases of the cultures they live in. This highlights the

* Many thanks go to Elise Bell and her students for submitting nearly 200 names to our database. Thank you also to our colleagues who acted as sounding boards, support systems, and beta testers, including Jessica Rett, Nathan Sanders, Pocholo Umbal, and Lex Konnelly. We are likewise grateful for the feedback of three anonymous LSA reviewers and audiences at the LSA 2023 Annual Meeting and the UCLA Bruin Linguists Society. Finally, thanks go to the UCLA Center for Advancement of Teaching; the UCLA Linguistics Department; and the UCLA Undergraduate Research Center for Humanities, Arts, and Social Sciences for their financial support. Authors: Maura O’Leary, Swarthmore College (moleary1@swarthmore.edu); Rainey Williams, University of California Los Angeles (raineywilliams@ucla.edu); & Mario Peng Lee, University of California Los Angeles (mariopeng@ucla.edu).

need for a resource that eliminates the influence of subjective biases while promoting diversity in example sentences.

In an effort to promote the integration of diverse proper names, this paper introduces the Diverse Names Generator, a new resource that generates names from a user-submitted, linguist-curated database through randomization – thereby removing unconscious bias from the name-choosing process.

§2 discusses a few of the works that have highlighted disparities in linguistic example sentences and mentions two resources that have already been created to help increase gender and cultural diversity in the proper names linguists choose. §3 introduces the Diverse Names Generator, a resource available via both website and app that allows instructors and authors to randomly generate names from a diverse database. §4 discusses the database itself, including the name submission, approval, and correction processes, as well as statistics about the diversity of the names present in the database at the time of publishing.

2. Background literature & the current state of affairs. Several studies have highlighted the gender bias prevalent in linguistics examples, both in textbooks and in peer-reviewed papers.

Macaulay & Brice (1997) analyzed 200 randomly selected example sentences from each of ten syntax textbooks published between 1969 and 1994 and found that male arguments outweighed female arguments in eight of the ten textbooks. In those eight books, which were largely written by male authors, the number of male subjects and agents outweighed the number of female subjects and agents with ratios varying from 8-to-1 to an extreme 22-to-1.

Additionally, the predicates which were paired with male or female arguments tended to reinforce gender stereotypes. Male-gendered names were used more frequently in examples involving cars, violence, or “intellectual” activities such as reading or writing. Female-gendered names were three times more likely to have their appearances described than males. In a case study of one syntax textbook, Macaulay & Brice (1994) found that female-gendered arguments tended to be portrayed as “excessively emotional” (1997:803), often appearing as the objects of ‘annoy’ and ‘please.’

Little has improved in the decades since Macaulay & Brice’s 1997 survey. Pabst et al. (2018) – findings later published in Cépeda et al. (2021) – found that textbooks published between 2005 and 2017 continued the same trends of privileging male arguments and promoting gender stereotypes. In the six textbooks surveyed, male arguments were almost twice as frequent as female ones and were much more likely to appear as subjects, agents, and experiencers. Male-gendered arguments were more likely to have occupations, handle books, and perpetrate violence. Meanwhile, female-gendered arguments were more likely to be direct or indirect objects, patients, and recipients and were more likely to exhibit negative emotions.

The issues discussed by Macaulay & Brice (1997), Pabst et al. (2018), and Cépeda et al. (2021) extend beyond the realm of textbooks. Kotek et al. (2021) examined 25,106 third-person arguments across 22,954 example sentences in 927 articles published in three major linguistics journals over the course of twenty years and found the same patterns present in textbooks. Male arguments appeared twice as often as female arguments, more often as subjects and agents, and were more often associated with having occupations or engaging in violence. Female arguments were “overrepresented as nonsubjects, especially as recipients” (2021:671) and were more often described with kinship terms.

In addition to the persistent gender imbalances and stereotypes, these studies have also re-

vealed a striking pattern of overrepresentation of specific proper names in educational materials. Cépeda et al. (2021) and Kotek et al. (2021) both highlight that the usage of two names in particular, “John” and “Mary,” eclipsed the usage of all others. Cépeda et al.’s analysis of textbooks revealed that “Mary” constituted 39% of the 241 female-gendered proper names, a usage rate nine times higher than the next most commonly used female name. Meanwhile, “John” was used in 37% of the 473 male-gendered proper names, a rate four times higher than the next most common male name. Similarly, Kotek et al. found that in the papers they examined, “30% of all male names are *John*” and “31% of all female names are *Mary*” (2021:663). The usage of these specific names reflect a strong Western bias when choosing proper names. Indeed, both papers noted that “a majority of the names . . . exhibited a bias that strongly favors Western, Caucasian-sounding proper names” (Cépeda et al. 2021:686) and that there was “a strong Western/Christian bias in authors’ choices, with little variability” (Kotek et al. 2021:663).

Diversity, or lack thereof, in the names chosen for use in example sentences can make a meaningful difference to students who encounter those sentences. Name bias in academia can cause students with non-Anglican names to feel like they don’t belong in the classroom. International students interviewed about the impact of heritage name mispronunciation report that their name is a symbol of both individual and cultural respect and that dismissing these names can result in a sense of underrepresentation and discrimination (Zhang & Noels 2021). Given that names can represent personal identity, familiar history, and ethnocultural identity, excluding names from certain cultural groups can affect students at an individual level, a family level, and a heritage level. The resulting feelings of exclusion and underrepresentation can contribute to the ‘leaky pipeline’ of academia, which affects both women (see, e.g., Goulden et al. 2011) and racial minorities (e.g., Holmes & Menachemi 2017).

One solution for this lack of representation is to integrate a more diverse set of names into classrooms, textbooks, and academic articles. By promoting the use of names that are conventionally underrepresented, we can help to create a learning environment where students feel represented, respected, and included. To quote Kotek et al., “better-constructed example sentences, using inclusive language, can send an important message to the field: inclusive language encourages participation from underrepresented groups, leading to a better community and therefore to better science, at the cost of just a little more thoughtfulness” (2021:672). They recommend: “When using names in example sentences, consider using diverse names, paying explicit attention to the distribution of gender in your examples. . . . [I]nclude non-Western names as an additional signal of diversity” (2021:673).

In efforts to help linguists use names that are more diverse in terms of both language of origin and gender, Konnelly et al. (2021) created the “Diverse Names Database,” which provides 72 diverse names, each coming from a different language. There are 26 names (one for each letter of the alphabet) for each of “feminine-leaning” names, “masculine-leaning” names, and “all-gender” names. Along with the name, gender, and language of origin, the database also provides a transcription for each name using the International Phonetic Alphabet (IPA). Thus, teachers and authors can pull names from a more diverse pool while still being able to pronounce those names.

In a similar vein, Kirby Conrod has compiled a list of names¹ that can be considered non-binary, as submitted on Twitter by non-binary people who use those names, again providing a simple method for confidently increasing diversity in example sentences.

¹ <https://docs.google.com/spreadsheets/u/1/d/1ILn1uBWqgwm5Qc71cr3OX2O3CoXsFqHOEUiYtRoAFo>

3. The Diverse Names Generator. When pulling from a well-curated list of diverse names – like those mentioned above or the homemade lists that many instructors keep – teachers and authors will likely use significantly more diverse names than they would if they were to merely pull names out of their head on the fly or if they succumb to the industry standards of “John” and “Mary.” However, actively choosing a name off of a list still requires the user to make choices about which name to use at any given moment, which leaves room for unconscious bias to creep in. Asking a teacher or author to constantly choose names off of a list may still result in a gender or cultural imbalance, even if the list they are pulling from is diverse and well balanced and even if the user is well-intentioned.

Our resource aims to make the process of using diverse names easier on two fronts: first, by providing an easily accessible, curated list of diverse names and second, by providing a mechanism by which to choose names that does not require the user to make a conscious choice, thereby avoiding issues of unconscious bias. We achieve the latter goal by randomly generating names from our database, allowing users to fill their example sentences with names that exhibit cultural and gender diversity without having to choose the names themselves.

The remainder of this section introduces this resource: the Diverse Names Generator (DNG). As is covered below, the resource exists as both a website and an app. Names are randomly generated from a database of user- and student-submitted names, but the selection pool may be narrowed by filters such as first letter, gender, presence of an example audio recording, or language-specific phonotactic constraints. However, outside of these optional filters, the names are generated randomly, allowing no space for unconscious bias to play a role in which names are chosen. Finally, each generated name minimally includes an IPA transcription, the gender most commonly associated with the name (“generally feminine,” “generally masculine,” or “gender neutral”), and at least one language the name is associated with. Name entries may also include information such as additional languages/cultures the name is associated with, the name written in non-Roman orthography, alternate spellings/forms/pronunciations, a recording, notes on famous people with the name (so that users have a better idea of what the name may conjure up for students and readers), and other notes (such as historical or cultural information).

§4 discusses the manner in which names are submitted, checked, and stored in the database.

3.1. USING THE DNG WEBSITE. The Diverse Names Generator website, which can be found at diversenamesgenerator.com,² was designed to provide a user-friendly interface that allows users to randomly generate proper names with minimal effort. On the homepage, names can be generated easily by pressing the “Generate” button (see the location marked with a (1) in Figure 1 below). Each name generated minimally includes the corresponding IPA, the gender commonly associated with this name (one of: “generally feminine,” “generally masculine,” or “gender neutral”), and the languages associated with this name (see location (2) of Figure 1).

² <https://diversenamesgenerator.com/>

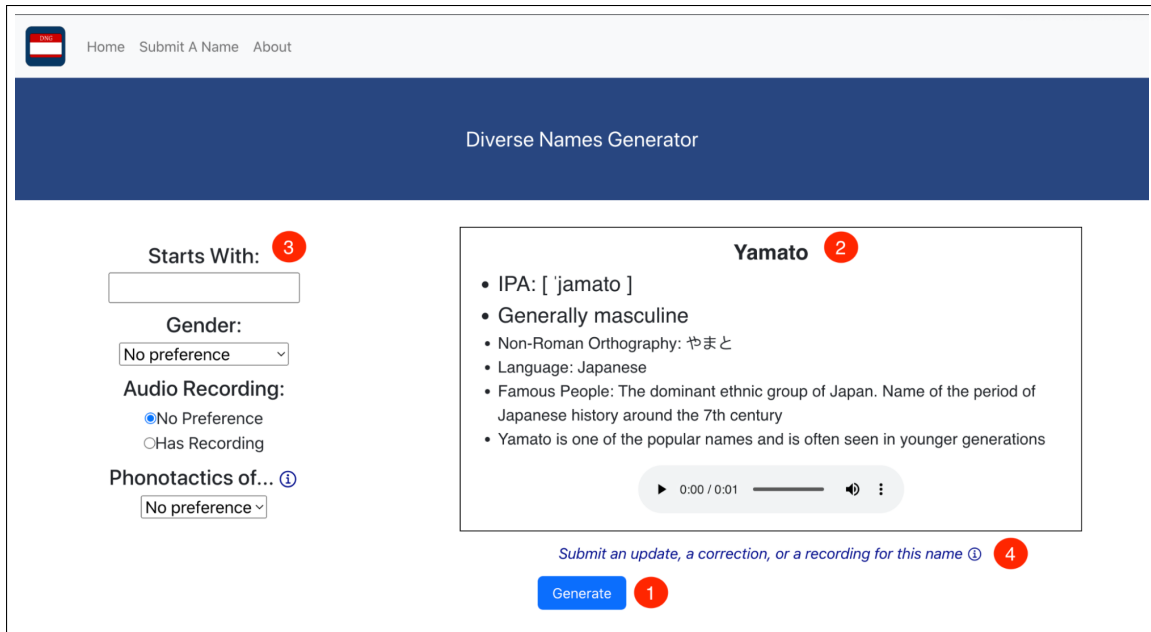


Figure 1. Homepage of the website

The examples in Figure 1 and Figure 2 showcase some of the additional pieces of information that may be provided by the submitter. Name entries may include the way the name is traditionally written in an alternate (non-Roman) orthography; a list of famous people with the given name, so that instructors and authors will have a better idea about what affiliations the name may bring up for students and readers; and a line containing a fun fact, historical information, or general notes about the name. Names may additionally come with an audio recording submitted as a reference for pronunciation, which the user can access by clicking the play arrow in the gray oval. Not pictured is that name entries may also list alternate spellings, alternate pronunciations, or shortened forms (e.g. “Sam” for “Samantha”) of the name. Finally, if a name submitter would like to receive recognition for their submission, their name will appear just underneath the name entry, as in Figure 2:

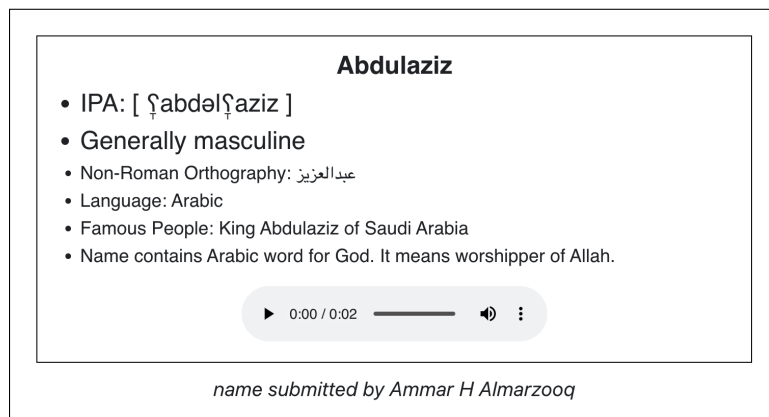


Figure 2. An example of a name with information provided by the submitter

3.1.1. USING THE DNG MOBILE APP. The mobile app offers an easy way to access diverse names quickly when they are needed at a moment's notice. Instructors can simply lay their phone or tablet on a desk or podium at the front of a classroom and then swipe across their screen to generate a new name in the exact moment that it is needed to complete an example sentence on the board. Names are written in a large font, with extra information available, but not cluttering the screen, so that the most important information for on-the-go usage is easily readable at a glance.

This portable version is inspired by index-card-style study apps; names appear on each “card” of a deck, and new names are generated by swiping the “top” card from the screen. Each card displays a name, an IPA transcription for the name, and the gender usage trends (“generally feminine,” “generally masculine,” or “gender neutral”), as provided by the submitter.

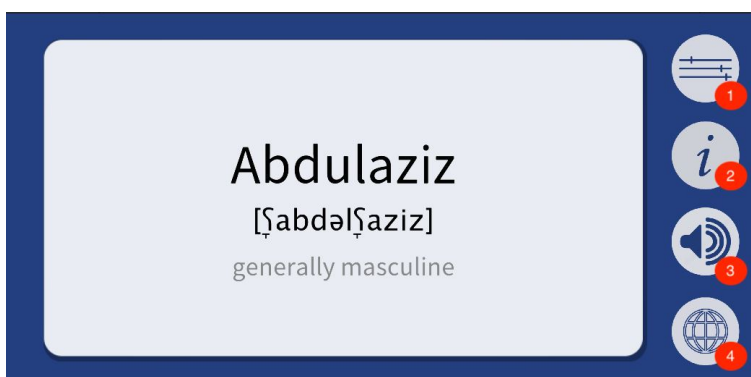


Figure 3. A name generated on the mobile app

Along the far right hand side of the app are four buttons. From top to bottom, the first button (location marked with a (1) in Figure 3) accesses filters that can be used to limit the pool from which names are pulled (see §3.3 for more information about the available filters). The second button (location (2) of Figure 3) will “flip” the index card to reveal additional information on the back (see Figure 4). Like the website, this information can include the name written in any traditional (non-Roman) orthography, language(s) of origin, alternate spellings or pronunciations, a list of famous people with the name, and any additional notes provided by the submitter. The third button will play an audio recording of the name’s pronunciation, if one has been submitted. (If no recording is available, this button will be grayed out.) The fourth and final button will take the user to the DNG website, so that they can access additional functions such as submitting or correcting a name or accessing additional information under the website’s “About” page.

At the time of publishing, the DNG mobile app is available on Android devices via the [Google Play Store](https://play.google.com/store/apps/details?id=com.DefaultCompany.DiverseNameGenerator).³ We intend to release an iOS version in the future; until that point, Apple users can access the website on their phones, which has been optimized for mobile use.

³ Our app is findable on the Google Play Store at the following link:

<https://play.google.com/store/apps/details?id=com.DefaultCompany.DiverseNameGenerator>
or by searching for “Diverse Names Generator” and clicking on the app bearing the following icon:



Diverse Names Generator
Diverse Names Generator

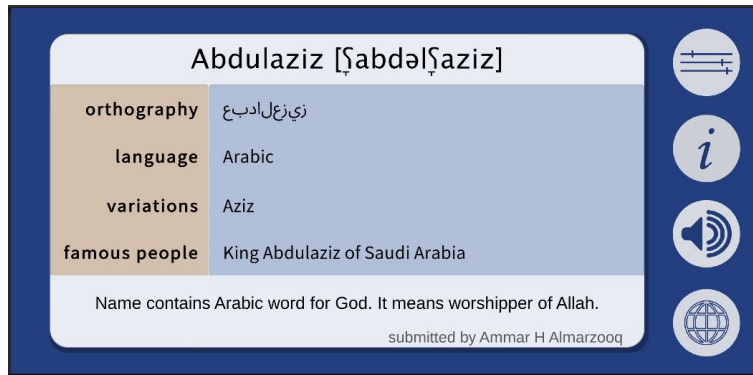


Figure 4. Information on the back of a name card

3.2. FILTERS. If desired, users can filter names so that the randomly generated names are pulled from a specified subset of the database. To access filters, website users should look to the right hand column of their screen (labeled (3) in Figure 1, repeated below), and app users can press the top button on the left hand side (labeled with (1) in Figure 3) to access the filter menu, which can be seen in Figure 5 below.

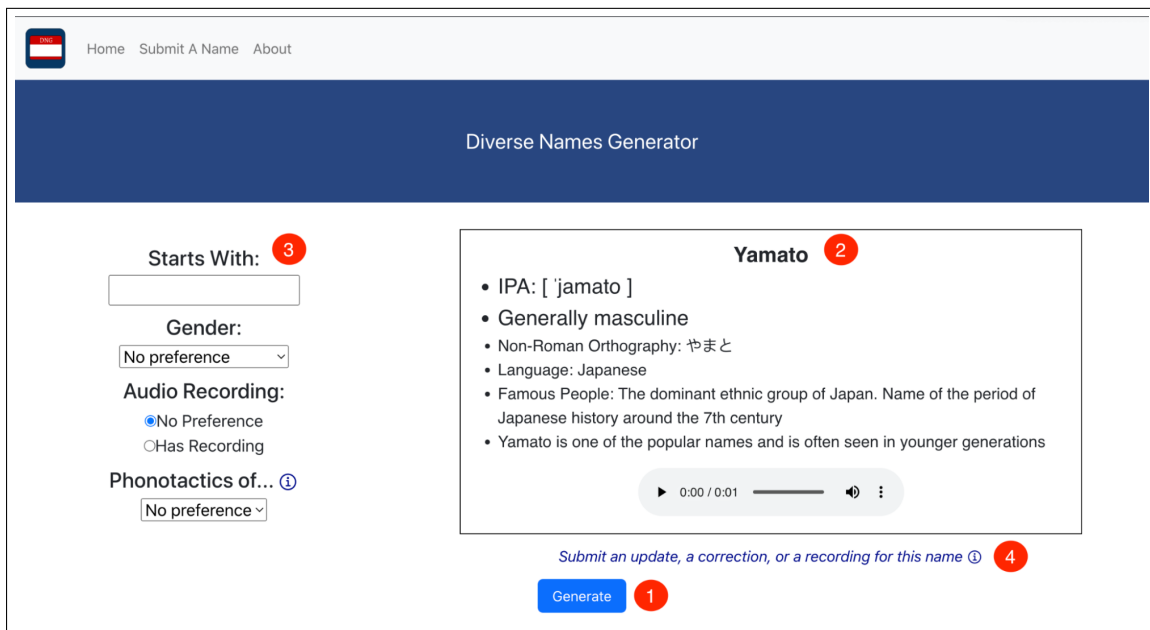


Figure 1. Homepage of the website

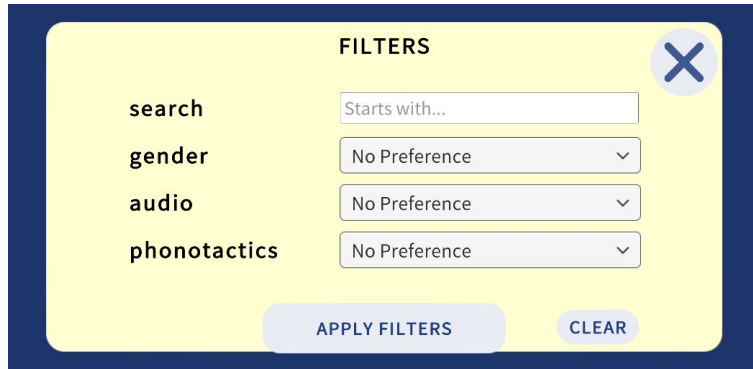


Figure 5. Filter menu on the mobile app

There are four available filters (all available on both the website version and the app).

First, the user can generate names that start with certain letters. This is particularly useful for semanticists, who often need to have no more than one name per letter of the alphabet, so that names will not be confused once they are reduced to a single initial in a predicate logic, lambda calculus, or other formal languages. This filter can also be useful for finding information about a specific name – just type the whole name into the search bar, and the pool of names will be limited to a single entry. (If there is only one name which fits all the filter criteria, a message will appear saying that only one name was found. If there are no names that match all specified criteria, a message will appear which states that no names were found and that the user should try using different filter options.)

Second, users can select for names that match a certain gender, which can be especially useful for syntacticians or morphologists who need to match certain phi features. This drop-down menu defaults to “no preference,” which will pull names from the entire database. “Generally feminine” or “generally masculine” will pull names which have been marked by name submitters as being primarily used for feminine or masculine arguments, respectively. “Gender neutral” will specifically pull names that submitters report as having equal or near-equal use between feminine and masculine. Each name in the database is only marked with one of these tags, which we view as close enough to mutually exclusive. We do not yet have a category for names that are recommended as non-binary and highly recommend the use of Kirby Conrod’s list of non-binary names.

The third filter allows users to exclude names that are not accompanied by an audio recording, which can be a useful tool for modeling the pronunciation of a given name.

Finally, the “Phonotactics of . . .” filter will generate names which use only the phonemes and phonotactics of a specific language (loosely construed). This filter was added to help ease the anxieties that some instructors might feel about producing unfamiliar phonemes or unfamiliar consonant clusters live in the classroom. This option allows those instructors to still pull names randomly from a diverse pool, without having to worry that the pronunciation of the names generated will be too far outside of their comfort zone. Staying within a certain level of comfort regarding pronunciation is also beneficial in that it curbs the mispronunciation of unfamiliar names, which, as discussed in §2, can be seen as disrespectful or othering to students from those cultures. That being said, we recommend that instructors who are comfortable pronouncing most of the IPA leave this filter set to “no preference” so that the pool of names remains maximally diverse.

At the time of publishing, we have currently tagged all of the names in our database for

whether they obey English phonotactic constraints. We are actively looking for volunteers to help tag names for other languages; if you, the reader, have this expertise and are willing to help tag names as binarily in or out of a specific language’s phoneme inventory and phonotactic constraints, please contact us at the email found on the “About” page of our website.

As a side note, we have considered using language processing to automatically tag submitted names; however, while it would be easy to tag names based on the phonemes they contain, it is beyond the scope of this project to create an automatic tagging system that can define syllable boundaries and check for pre-specified phonotactic constraints. Thus, we are relying on a human tagging system for the time being.

In the end, the filters we provide are meant to make the DNG more useful to the people we anticipate to be users. All else equal, we recommend generating names from the entire database, unfiltered. However, there are often many constraints at play when picking a name for an example sentence – be it initial letter, gender, or ease of pronunciation – and we have attempted to make the DNG such that users can still benefit from the random generation of diverse names even when additional constraints are at play.

3.3. OFFLINE CAPABILITIES. While the DNG requires an internet connection to fetch names from the database, users can pre-load a batch of names into their local device to use in offline settings (using either the website or the app). In order to use the DNG without internet connection, simply apply the desired filter constraints and generate at least one name. From there, the relevant batch of names will be locally stored and users can freely generate names without internet connection.

Offline usage comes with two limitations. First, because audio takes up more storage space, recordings of sample pronunciations are never downloaded to be stored locally on the user’s device. Thus, recordings will not be available when the website or app are used offline. Second, a user working offline will only have access to the batch of names that their online search generated. If the user closes or refreshes the webpage, or if they attempt to search using new filters while offline, access to the downloaded set of names will be lost.

3.4. TECHNICAL DEVELOPMENT. Both the mobile app and the website were designed following the standard software construction paradigm: the front end (user interface) connects to the back end (data management and algorithms).

The two implementations’ back end are both powered by the same database hosted at Google Firestore Database, an interface for large-scale cloud data storage and management. As a design choice, both technologies perform a singular API call to Google Firestore upon applying a filter and shuffling the names, as opposed to individual API calls for each name generated. This batch of filtered names is then stored locally in the user’s device, and the ‘generate’ function parses through this local list of names instead of calling for new names from the database. This enables increased performance, reduced API calls, and offline capabilities.

Names are generated through a random index generated via integrated pseudo-random algorithms. The random index for the local list is stored and compared to the previous random index to ensure that the same name entry is never repeated twice in a row. Audio clips are stored as references, and when a name with audio is displayed, the platform accesses the storage for the audio of that name only to avoid downloading unnecessary audio clips.

The front end of the website was built with React, a standard technology in modern web development. This design choice allows effortless integration of the database and responsive design

that is compatible with different screen sizes across all devices. The mobile app was developed with Unity Engine, a versatile toolbox for application development that facilitates portability to different platforms. We selected this technology due to the ample available resources and potential scalability; currently the app is only available on Android devices, but we intend to release an iOS version in the future.

4. The DNG database of names. As mentioned above, the names that are generated by the DNG app and website are pulled from our database (hosted on Google Firestore Database). All of the names in our database (including accompanying information and recordings) were submitted by users of the website or by students who received course credit for submitting names. This section discusses the name submission process (§4.1), including how to give students course credit for submitting names to the DNG database (§4.1.1), the name approval process (§4.2), and how to submit corrections or recording for names already in the database (§4.3). We also include a brief look at the diversity of the names that are in the database at the time of writing (§4.4).

4.1. NAME SUBMISSION. The DNG database is fully composed of submissions from users and students. To submit a name, users can go to the website and navigate to the “Submit a Name” tab, where they will be able to submit names into the database. We ask minimally for the name, IPA transcription, any common gender associations, and the languages the name is affiliated with. These questions are followed by a series of optional questions aimed at filling out a full picture of the name; a complete list of submission questions is given in an appendix.

Home Submit A Name About

Diverse Names Generator

Thank you for helping us to improve the Diverse Names Generator!

Please use the questions below to add a name to our database. We especially appreciate names from underrepresented cultures and linguistic groups. Please use names that you are familiar with and know how to pronounce. We are only looking for first/non-family names.

Are you filling this out for class credit? *

No Class Credit

What name are you submitting? Please use only Roman characters (diacritics are fine). Please only submit ONE name per form submission. *

Please fill out this form for only one name submission at a time

Please provide an IPA transcription for the name you are submitting, including stress and tone markers. Please use [this site](#) to transcribe, and then copy and paste your transcription. *

What gender is associated with this name? *

Figure 6. The ‘submit a name’ page of the DNG website

If submitters would like recognition for submitting a name, they can provide their own name when prompted. Figures 2 and 4 show examples of what recognition looks like on the website and app, respectively.

4.1.1. NAME SUBMISSION FOR COURSE CREDIT. As an additional feature, we have integrated a crediting system that allows instructors to encourage their students to submit names for extra

credit or participation points. Names submitted for class credit require a recording to be submitted so that we can check the IPA transcription – this step is optional for submitters not receiving course credit.

We hope this expansion of the pool of submitters beyond users of the site will result in more diverse names for our database, likely in ways that are representative of the student populations we serve. We hope that instructors teaching phonetics, phonology, sociolinguistics, and perhaps even introductory linguistics will (continue to) ask any students who have adequate experience with IPA to submit names as a small example of how their newly learned linguistics skills can make a positive difference in the world.

If you would like your students to receive course credit for submitting names, please contact us at the email listed on the “About” page of our website.

4.2. DATABASE AND APPROVAL. When a new name is submitted through the website, it is not immediately added to the pool from which names are randomly generated. Instead, submitted names (and accompanying information) are temporarily stored in a “needs approval” folder on the Google Firestore Database. All of the information submitted is hand-checked by a member of our team prior to being added to the pool from which the DNG randomly generates names.

Given that our team will continue to change over time, we have made a user-friendly desktop program to make the name approval process as easy as possible. The program shows all of the information submitted and allows the linguist checking the submission to correct any obvious errors. It should be noted, however, that we do not change the submitted content unless we find transcription errors or typographical mistakes, or if the entry contains content inappropriate for the DNG. (See §4.3 for information about correcting entries in the DNG.)

The screenshot shows a desktop application window titled "Intermediary". The interface is divided into several sections. On the left, there are input fields for "Name", "IPA", "Primary Language", "Native Orthography", "Famous People", and "Extra Notes". To the right of these are fields for "Alt Name", "Alt IPA", "Alt Languages", and "Any Gender" (a dropdown menu). Below these are fields for "No Audio Permissions" (a dropdown menu), "Any Phonotactic" (a dropdown menu), "extra credit Name", and "Submitted by". At the bottom left, there are fields for "extra credit Class" and a "Fetch Database" button. On the right side, there is a large grey area labeled "Button" at the top, and a "Final Approval" button at the bottom. In the middle of the right side, there is a "Play" button next to an "Audio Reference" field, and a red "Delete" button next to it.

Figure 7. “Intermediary” desktop program used for the approval of new name submissions

Using the desktop program means that changes are not made directly in the database entry, which, in turn, means that all of the originally submitted information can be fully preserved, even if we make small changes to the entry that is eventually sent to the DNG generation pool. After a name is approved, the edited entry is moved to an “approved” folder in the database and an original, unedited version is stored in a “backup” folder in case changes need to be reverted. The DNG pulls names exclusively from the “approved” folder.

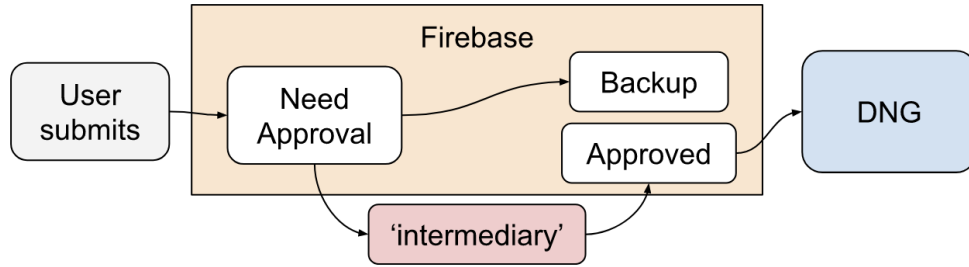


Figure 8. The name submission pipeline illustrated

4.3. NAME CORRECTION. Our database is completely composed of user- and student-submitted entries. As mentioned above, we do not require every possible field to be filled out in the name submission form, and we generally do not alter name submissions beyond what we believe to be unintentional mistakes. Thus, any given name entry may be missing some information or may contain errors. Likewise, we do not require submitters to provide a recording to be used on the website, so not all name entries are accompanied by recordings.

Any user can submit an update, a correction, or a recording for a given name by clicking on the text directly below the name entry which reads “Submit an update, a correction, or a recording for this name” (marked with a (4) in Figure 1). A text box is then provided where the user can write a message explaining either an addition or suggested modification to the currently displayed name (see Figure 9 below). For instance, if a name is listed as being associated with “Italian, Spanish, and Estonian,” a user may wish to write in the text box that the name is also associated with French and Finnish, so that that information can be added to the name entry.

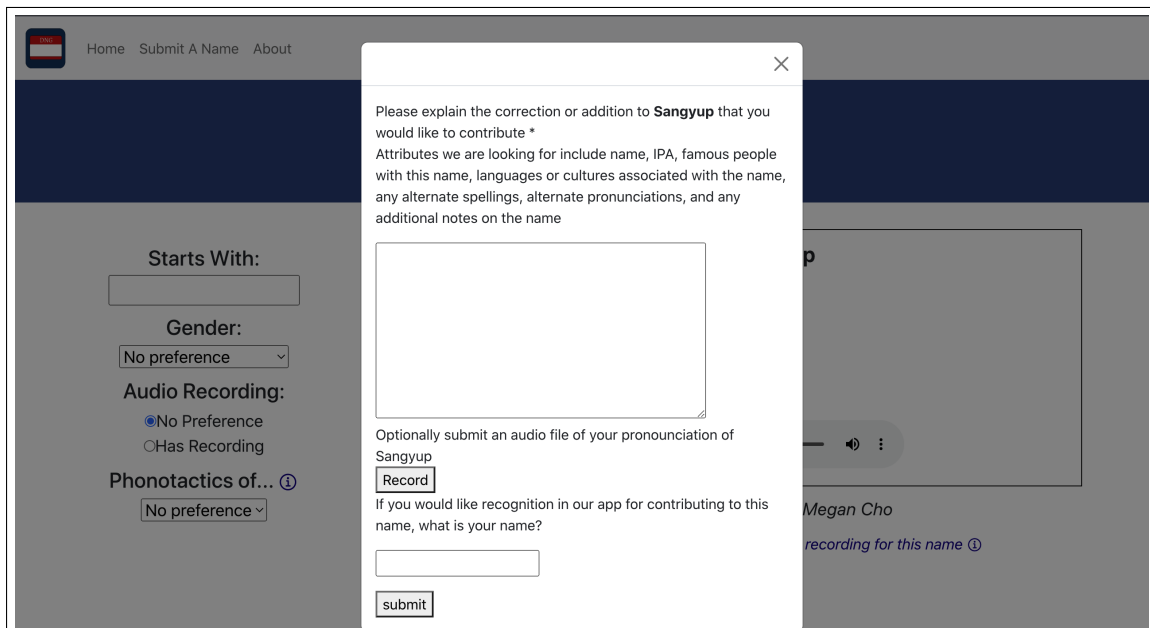


Figure 9. The textbox which appears when the user clicks “Submit an update, a correction, or a recording for this name”

Underneath the text box is a “Record” button where the user can submit an audio recording of their pronunciation of the name to be added to the name entry.

4.4. CURRENT STATISTICS ABOUT THE DATABASE. At the time of publishing, the database contains over 200 names from more than 30 different languages.

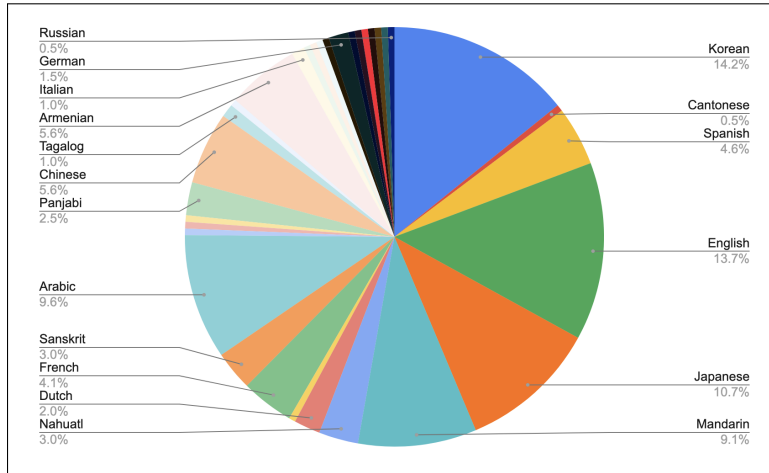


Figure 10. Distribution of language diversity among submitted names

This distribution is affected by the population that submitted the names, which has thus far been primarily undergraduate students studying phonetics and phonology at UCLA. We hope that in the future, with a larger and more varied population of submitters, we can get an even more diverse distribution of names.

The distribution of name submissions by gender is also pleasantly diverse, as shown in Figure 11:

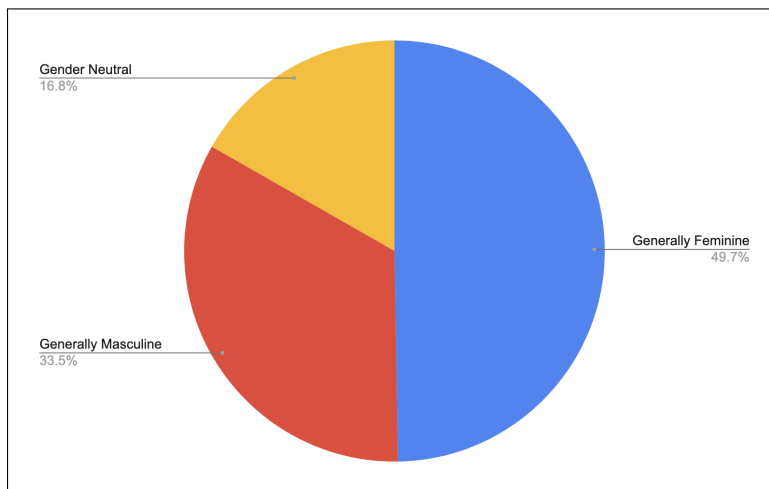


Figure 11. Distribution of gender diversity among submitted names

5. Conclusion. This paper has presented a novel resource for any instructor or author who uses proper names in example sentences. Given the importance of using names that represent the diversity of our field and our students, we hope to encourage instructors and authors to aim for

more diverse representation in the proper names that they use; the Diverse Names Generator makes it more feasible to do this without the effects of unconscious bias.

The generator, available via website or Android app, provides names randomly pulled from a user-submitted database, with the allowance for filtering by starting letter, gender, presence of recording, or phonotactic constraints. The generated names are minimally accompanied by IPA transcriptions, common gender associations, and language associations. Some names come with additional information, such as recordings, alternate forms of the name, the name written in a non-Roman orthography, or additional interesting facts. Users are able to enrich the database by directly submitting or correcting name entries.

We hope that easy access to diverse names via this resource will help to encourage authors and instructors to incorporate diverse names into their materials and help to promote diversity in our field, especially for the purpose of helping our students to see how they fit into the field of linguistics.

References

- Cépeda, Paola, Hadas Kotek, Katharina Pabst & Kristen Syrett. 2021. Gender bias in linguistics textbooks: Has anything changed since Macaulay & Brice 1997? *Language* 97(4). 678–702. <https://doi.org/10.1353/lan.0.0256>.
- Goulden, Marc, Mary Ann Mason & Karie Frasch. 2011. Keeping women in the science pipeline. *The ANNALS of the American Academy of Political and Social Science* 638(1). 141–162. <https://doi.org/10.1177/0002716211416925>.
- Holmes, Ann M. & Nir Menachemi. 2017. Retention of minority faculty: Does a leaky pipeline threaten achieving diversity in academic health administration? *Journal of Health Administration Education* 34(2). 229–242.
- Kohli, Rita & Daniel G. Solrzano. 2012. Teachers, please learn our names!: Racial microaggressions and the K-12 classroom. *Race Ethnicity and Education* 15(4). 441–462. <https://doi.org/10.1080/13613324.2012.674026>.
- Konnely, Lex, Pocholo Umbal & Nathan Sanders. 2021. The Diverse Names Database: A tool for creating more equitable, diverse, and inclusive linguistic example sentences. Presentation. Annual Conference of the Canadian Linguistics Association.
- Kotek, Hadas, Rikker Dockum, Sarah Babinski & Christopher Geissler. 2021. Gender bias and stereotypes in linguistic example sentences. *Language* 97(4). 653–677. <https://doi.org/10.1353/lan.2021.0060>.
- Macaulay, Monica & Colleen Brice. 1994. Gentlemen prefer blondes: A study of gender bias in example sentences. In Mary Bucholtz, A. C. Liang, Laurel A. Sutton, and Caitlin Hines (eds.), *Cultural performances: Proceedings of the Third Berkeley Women and Language Conference*, 449–61. Berkeley: University of California.
- Macaulay, Monica & Colleen Brice. 1997. Don't touch my projectile: Gender bias and stereotyping in syntactic examples. *Language* 73(4). 798–825. <https://doi.org/10.2307/417327>.
- Pabst, Katharina, Paola Cépeda, Hadas Kotek, Kristen Syrett, Katharine Donelson & Miranda McCarvel. 2018. Gender bias in linguistics textbooks: Has anything changed since Macaulay & Brice (1997)? Paper presented at the 92nd Annual Meeting of the Linguistic Society of America, Salt Lake City.
- Zhang, Ying Shan Doris & Kimberly Noels. 2021. The frequency and importance of accurate heritage name pronunciation for post-secondary international students in Canada. *Journal*

Appendix

Below are the questions on our name submission form. If you are filling out the form for course credit, questions (1)-(4), (6), and (7) are required. If you are not filling out the form for course credit, only (1)-(4) are required.

1. What name are you submitting? Please use only Roman characters (diacritics are fine).
2. Please provide an IPA transcription for the name you are submitting, including stress and tone markers.
3. What gender is associated with this name? (Pick from “Generally Masculine,” “Generally Feminine,” or “Generally Gender-Neutral.”)
4. What language(s) is this name associated with?
5. If this name is generally written using non-Roman characters, please supply the name in that orthography here.
6. Please record yourself pronouncing this name so that we can check your IPA transcription. (To record, press the record button below.)
7. We may offer an audio reference for some names. If so, do you consent for the recording(s) you submit to be used on our website/app? You are not obligated to let us use your recording(s), even if you are completing this for course credit. (Pick from “yes” or “no.”)
8. Are there any alternate spellings for this name? Are there any lengthened or shortened forms of this name? If so, please enter them here, separated by commas. If none, leave blank.
9. Are there any alternate pronunciations? If so, please enter them here, separated by commas. This is also a good place to enter a more broad/narrow transcription than you used above. If none, leave blank.
10. Please upload a recording of yourself pronouncing any alternate pronunciations so that we can check your IPA transcription.
11. What culture(s) (if not included in languages) is this name associated with? Type and separate with commas.
12. Are there any famous instances of this name? If none, leave blank.
13. Any other notes about this name? If none, leave blank.
14. If you would like recognition in our app for submitting this name, what is your name?
15. If you need course credit, what is your name?
16. If you need course credit, what is your school, course code, and term?